

DOI:10.3880/j.issn.1004-6933.2023.03.014

基于机器学习的太湖流域多层次防洪调度方案综合评价

高玮志¹,高华勇¹,王兆礼^{1,2},赖成光^{1,2}

(1. 华南理工大学土木与交通学院, 广东 广州 510641; 2. 人工智能与数字经济广东省实验室(广州), 广东 广州 510330)

摘要:为解决太湖流域多层次防洪调度方案在不同防洪层次目标下的评价问题,构建了流域、区域、城镇多层次防洪排涝调度方案综合评价指标体系,并基于 K 近邻(KNN)和随机森林(RF)算法构建调度方案综合评价模型。结果表明,联合KNN模型和RF模型实现了KNN-RF组合模型评价,其针对流域、区域与城镇3个层次防洪目标进行调度方案评价的平均相对误差和平均绝对误差分别降低至1.25%、0.82%、2.43%和0.511、0.342、1.380,最大相对误差和最大绝对误差得到改善,等级划分总体正确率高于95%;KNN-RF组合模型能筛选出各层次防洪目标下较优的调度方案,减少单一算法不确定性导致的异常评价误差,评价精度显著提高。

关键词:多层次防洪调度方案; K 近邻算法;随机森林算法;KNN-RF组合模型;太湖流域

中图分类号:TV87 **文献标志码:**A **文章编号:**1004-6933(2023)03-0118-08

Comprehensive evaluation of multi-level flood control operation schemes in the Taihu Lake Basin based on machine learning//GAO Weizhi¹, GAO Huayong¹, WANG Zhaoli^{1,2}, LAI Chengguang^{1,2} (1. School of Civil Engineering & Transportation, South China University of Technology, Guangzhou 510641, China; 2. Guangdong Artificial Intelligence and Digital Economy Laboratory(Guangzhou), Guangzhou 510330, China)

Abstract: In order to solve the evaluation problem of multi-level flood control operation schemes under different flood control level objectives, taking the the Taihu Lake Basin as study area, a comprehensive evaluation index system of multi-level flood control and waterlogging drainage operation schemes for basins, regions and cities was constructed. And a comprehensive evaluation model of operation schemes was constructed based on K -nearest neighbor (KNN) and random forest (RF) algorithms. The results show that the combination evaluation of KNN-RF is achieved by combining KNN model and RF model. The average relative error and average absolute error of KNN-RF combining model in the operation scheme evaluation for the three levels of flood control objectives, namely, basin, region and town, are reduced to 1.25%, 0.82%, 2.43% and 0.511, 0.342, 1.38, respectively. The maximum relative error and maximum absolute error have been improved, and the overall accuracy rate of grading is higher than 95%. The KNN-RF combining model can screen out optimal operation schemes for various levels of flood control objectives, reduce abnormal evaluation errors caused by the uncertainty of a single algorithm, and significantly improve the evaluation accuracy.

Key words: multi-level flood control operation scheme; K -nearest neighbor algorithm; random forest algorithm; KNN-RF combining model; Taihu Lake Basin

洪涝灾害已经成为全球范围内自然灾害中占发生次数最多的一种灾害^[1],世界各地河流洪水的规模和频率都在增加,并且未来几十年里全球遭受洪涝灾害的风险仍将上升^[2-3]。据《中国水旱灾害公报》统计,2006—2018年,我国洪涝灾情年均受灾面积超过900万 hm^2 ,年均直接经济损失接近2000亿元。

为应对洪涝灾害,几十年来水利工程建设高速推进,2013年第一次水利普查公告中^[4],我国已有10万座规模以上水闸近和424451座泵站,大量的闸门和泵站建设在经济发达、人口稠密的珠三角、长三角平原河网区^[5-6]。分布密集的闸泵群通过调度方案进行调度控制,为充分发挥闸门和泵站的防洪排涝效

基金项目:国家重点研发计划(2018YFC1508200,2021YFC3001000);国家自然科学基金(51879107);广东省科技计划(2020A0505100009)

作者简介:高玮志(1998—),男,硕士研究生,主要从事洪涝模拟研究。E-mail: 609163796@qq.com

通信作者:王兆礼(1979—),男,教授,博士,主要从事水文水资源研究。E-mail: wangzhli@scut.edu.cn

果,需要采用合适的调度方案,从而降低流域洪水风险。调度方案的评价和优选是防洪决策的关键^[7],但防洪调度是一个多目标非线性的复杂决策问题,各调度目标之间存在矛盾,且防洪效益难以度量,使得调度方案的决策优选成为研究难点^[8-9]。近年来,不少学者将改进的 TOPSIS^[10]、模糊综合评价^[11-12]、集对分析法^[13]、物元分析法^[14]、蚁群算法^[15]等评价方法应用于调度方案的评价优选研究中。这些研究多考虑水库和水库群的调度方案评价优选,针对平原河网区闸泵多层次调度方案的评价优选研究仍相对缺乏。平原河网区的多层次调度方案评价具有多指标、多层次、非线性的特点^[16]。如在太湖流域,城镇集中排涝导致江南运河水位陡涨陡落,区域防洪压力增大,进而又导致流域层面的外排压力增大,造成流域、区域和城镇的防洪排涝统筹上的困难和矛盾,因此调度方案的评价和选取需兼顾考虑流域、区域、城镇3个防洪层次。由于涉及的流域水系和闸泵数量众多,相比单一水库的调度方案评价,需要考虑更多水位、水量指标和流域、区域、城镇不同防洪侧重情景。

传统数学评价方法的运算步骤较为复杂,特别在指标数和方案数较多时,巨大的模糊矩阵的计算量会降低评价的效率,复杂的评价步骤也在一定程度上限制了其在平原河网多层次调度方案评价优选问题中的应用。近年来,人工智能算法发展迅速,其中机器学习算法能从经验数据中学习输入与输出之间的复杂关系,具有提取高维度数据特征和处理非线性数据的能力^[17-18]。随机森林(random forest, RF)算法和K近邻(K-nearest neighbor, KNN)算法是典型的机器学习算法,具有成熟、计算效率高、所需调节参数少、预测精度高、对训练集异常值鲁棒性好等优点^[19-20],目前已经被广泛应用在洪水风险^[21-23]、水资源管理^[24-26]、水生态^[27-30]等诸多领域的评价研究中。也有一些学者指出,由于单一机器学习算法的局限性,无论如何优化单一模型的参数,都可能会存在较大的误差^[31-32]。因此,部分学者尝试将机器学习模型进行组合,以突破单一模型的性能限制^[33],如 Hou 等^[32]将 KNN 算法和 RF 算法结合,显著提升了模型对淹没水深的预测精度。

鉴于此,本文以太湖流域超标洪水情景的大量调度方案及其评分为例,结合 KNN 模型和 RF 模型构建调度方案 KNN-RF 组合评价模型,利用该模型对调度方案进行评价,旨在为流域、区域与城镇不同防洪侧重情景下调度方案的快速评价提供参考。

1 研究区概况与数据来源

1.1 研究区概况

太湖流域地处长三角城市群核心区域,总面积 3.69 万 km²,是我国经济最发达、人口最集中的地区之一^[34-35],在国家经济社会发展中具有重要战略地位。但受平原地势低洼、河道坡降小和潮汐顶托等影响,流域河网排水速度慢,洪涝灾害频繁。经过多年的治理,太湖流域已形成流域、区域、城镇3个层次的防洪格局和工程体系(图1)。流域骨干控制工程包括望虞河、太浦河,区域控制工程包括武澄锡西控制线、沿太湖控制线等,城市防洪排涝工程包括苏锡常三大防洪包围圈。防洪大包围的建设令城镇排涝动力激增,强降雨时城镇排涝造成江南运河水位陡涨陡落,增大了区域防洪压力^[36],进而又会增大流域层面太湖、太浦河、望虞河的外排压力,流域、区域、城镇防洪排涝矛盾凸显,极大增加了统筹调度的难度^[37],对调度方案的防洪协同性提出了更高的要求。本文主要针对江南运河江苏段沿线区域,包括湖西区、武澄锡虞区和阳澄淀泖区3个水利分区。

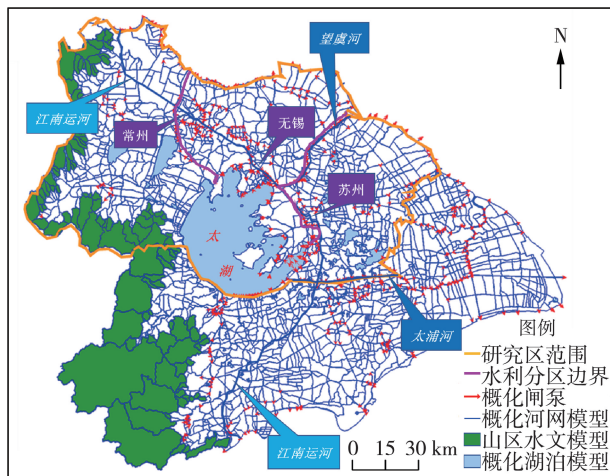


图1 太湖流域多层次防洪格局和防洪排涝工程调度仿真模型
Fig.1 Multi-level flood control pattern and flood control and drainage engineering scheduling model in the Taihu Lake Basin

1.2 数据来源

本文所用数据来源于国家重点研发项目“高度城镇化地区防洪排涝实时调度关键技术与示范”。基于项目研发的太湖流域防洪排涝工程调度仿真模型^[38],设置超标洪水情景、降雨情景,采用1999年和2016年的5—9月(汛期)的实测逐小时降雨数据,水位边界设置采用相应的实测逐小时潮位数据,共计对157个调度方案的调度情况进行模拟,得到每个调度方案下太湖流域的各区域水位、水量等数据。

调度方案设置依据《太湖流域洪水与水量调度方案》《太湖流域 2016 年超标准洪水应对方案》《2020 年太湖超标洪水应急调度方案》和《苏南运河区域洪涝联合调度方案》。提炼研究区防洪排涝基础调度方案,以 5 cm 为一个梯度、30 cm 或 50 cm 为总变化幅度分别调整流域骨干工程、区域控制工程、城镇大包围工程调度规则的控制水位,并对调度规则进行组合优化,构建流域、区域、城镇不同层次的防洪排涝调度方案。

模型在流域内概化得到 2 083 条河流及 144 个湖泊,概化了沿江沿杭州湾外边界、环太湖、流域主要控制线、城市大包围等重点闸泵工程共 863 座,在本研究区内有 406 座闸泵,工程位置如图 1 所示。模型经率定验证表明整体可信度高,能较为准确地模拟不同调度方案下流域各地区的水位、水量情况,可为本文调度方案的评价优选提供基础。

2 研究方法

2.1 KNN 模型

KNN 算法是 Cover 和 Hart 于 1968 年提出的一种非参数回归的机器学习算法^[39],其基本思想是使用邻域算法在训练集中找到与目标样本最接近的 K 个样本,并使用这 K 个样本进行估计。基于 KNN 算法构建的 KNN 模型具有数据挖掘的能力,借助历史数据确定输入与输出之间的关系,能较好地保持原始数据的特性。KNN 模型具有成熟、简单、对训练集噪声鲁棒性好的优点。

2.2 RF 模型

RF 算法是 Breiman 于 2001 年提出的一种以决策树为基础的高效的集成算法,由一系列树形预测器组成^[40]。基于 RF 算法构建的 RF 模型,可以利用 bootstrap 重抽样方法构建多个样本,对每个样本构建决策树,每棵树对输入变量进行评分,最终的回归结果由全部决策树的结果均值进行决定。RF 模型具有强大的数据挖掘能力,对解决多变量、非线性问题有很好的效果,对异常值和噪声容忍度高,且不易过拟合。

2.3 KNN-RF 组合模型

为降低由单一机器学习算法本身存在的异常误差,提高调度方案评价模型的预测精度和可靠性,模型组合是一种有效方法^[41]。根据训练集的确定系数 R^2 对 RF 模型和 KNN 模型的模拟评价结果进行加权再分配,得到 KNN-RF 组合模型评价结果,公式为

$$R_{\text{fin}} = \frac{M_{\text{KNN}}R_{\text{KNN}} + M_{\text{RF}}R_{\text{RF}}}{M_{\text{KNN}} + M_{\text{RF}}} \quad (1)$$

式中: R_{fin} 为 KNN-RF 组合模型评分; M_{KNN} 和 M_{RF} 分别为 KNN 和 RF 模型的确定系数; R_{KNN} 和 R_{RF} 分别为 KNN 和 RF 模型的评分。

选取 R^2 、均方根误差 (root mean square error, RMSE)、平均绝对误差 (mean absolute error, MAE) 和平均相对误差 (mean relative error, MRE) 4 个常用的机器学习性能评估指标用于验证模型的可靠性。

3 评价指标体系构建

3.1 评价指标体系

平原河网区的防洪排涝调度是实践性很强的生产工作,涉及数量众多的河道、水利分区、水工建筑物等,因此选取的指标主要根据太湖流域防洪排涝调度的相关文件并咨询专家得到。根据《太湖流域片水情年报》《苏南运河区域洪涝联合调度方案》等初步选取具有代表性的水位、水量作为评价指标;然后通过咨询专家,对指标体系进行增补、删减和合并,并给出下一步的修改意见;结合修改意见进行对指标体系进行修改,并重复上述咨询专家的过程。经过多次专家咨询和指标体系修改,建立了适用于太湖流域的“流域-区域-城镇”多层次防洪排涝调度方案评价指标体系(图 2),包括水位、水量两大类共 55 个指标变量。本文将指标分为 4 类:流域型、区域型、运河沿线型和城镇型。 A_1 指标中 $A_{1,1} \sim A_{1,12}$ 为区域型指标, $A_{1,13} \sim A_{1,15}$ 为流域型指标, $A_{1,16} \sim A_{1,20}$ 为运河沿线型指标; A_2 指标中 $A_{2,1} \sim A_{2,12}$ 为区域型指标, $A_{2,13} \sim A_{2,15}$ 为流域型指标, $A_{2,16} \sim A_{2,20}$ 为运河沿线型指标, $A_{2,21} \sim A_{2,23}$ 为城镇型指标; B_1 指标中 $B_{1,1} \sim B_{1,3}$ 为区域型指标, $B_{1,4}$ 为流域型指标; B_2 指标中 $B_{2,1} \sim B_{2,2}$ 、 $B_{2,4}$ 为区域型指标, $B_{2,3}$ 、 $B_{2,5}$ 为流域型指标; B_3 指标中 $B_{3,1} \sim B_{3,3}$ 为城镇型指标。

3.2 权重设置

各指标权重的设置需要综合考虑调度方案评价指标体系的构成以及不同层次的调度目标。因此在指标权重设置中遵循的原则为:①考虑指标体系的构成时,需要协调指标体系内水量、水位、水位超警时长等各类指标的权重分布,应对实际调度决策工作中更关注的指标赋予相对更大的权重;②考虑到多层次跨区域的调度目标时,即调度目标分别侧重于流域、区域、城镇不同层次时,要按照指标的不同类型,根据调度目标适当提高相应类型指标的权重。目前实际调度工作中方案的决策主要依赖调度专家集体协商,因此本文指标权重的设置采用专家赋权方法。

当调度目标侧重于流域层次时,得到 A_1 指标中 $A_{1,1} \sim A_{1,12}$ 权重为 0.01, $A_{1,13} \sim A_{1,14}$ 权重为 0.05, $A_{1,15}$

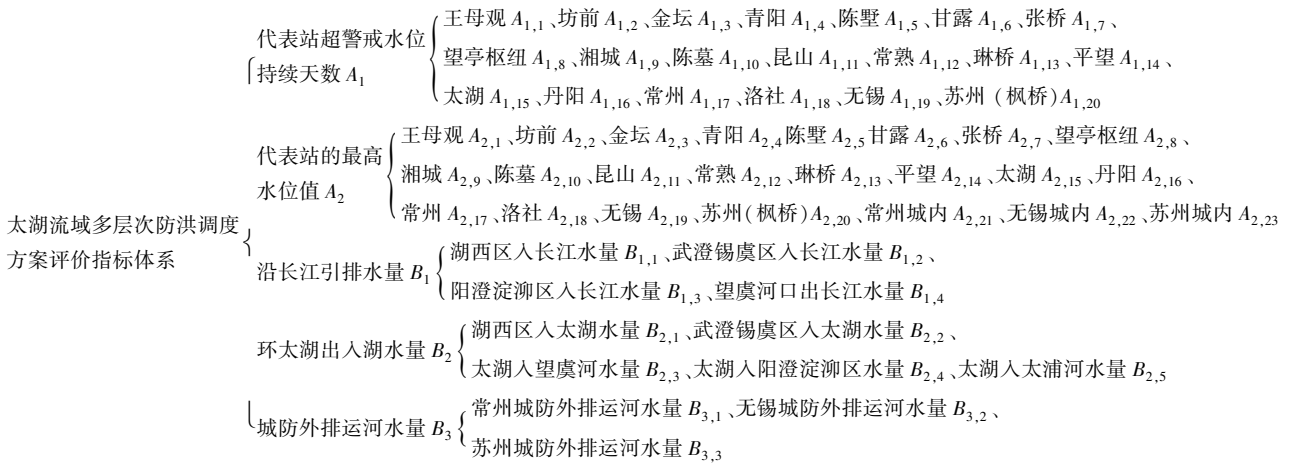


图2 太湖流域多层次防洪调度方案评价指标体系

Fig.2 Evaluation index system of multi-level flood control operation schemes in the Taihu Lake Basin

权重为 0.15, $A_{1,16} \sim A_{1,20}$ 权重为 0.01; A_2 指标中 $A_{2,1} \sim A_{2,12}$ 权重为 0.01, $A_{2,13} \sim A_{2,14}$ 权重为 0.05, $A_{2,15}$ 权重为 0.15, $A_{2,16} \sim A_{2,20}$ 权重为 0.01, $A_{2,21} \sim A_{2,23}$ 权重为 0.01; B_1 指标中 $B_{1,1} \sim B_{1,4}$ 权重为 0.01; B_2 指标中 $B_{2,1} \sim B_{2,5}$ 权重为 0.01; B_3 指标中 $B_{3,1} \sim B_{3,3}$ 权重为 0.01。

当调度目标侧重于区域层次时,得到 A_1 指标中 $A_{1,1} \sim A_{1,12}$ 权重为 0.016, $A_{1,13} \sim A_{1,15}$ 权重为 0.016, $A_{1,16} \sim A_{1,20}$ 权重为 0.03; A_2 指标中 $A_{2,1} \sim A_{2,12}$ 权重为 0.016, $A_{2,13} \sim A_{2,15}$ 权重为 0.016, $A_{2,16} \sim A_{2,20}$ 权重为 0.03, $A_{2,21} \sim A_{2,23}$ 权重为 0.016; B_1 指标中 $B_{1,1} \sim B_{1,4}$ 权重为 0.016; B_2 指标中 $B_{2,1} \sim B_{2,5}$ 权重为 0.016; B_3 指标中 $B_{3,1} \sim B_{3,3}$ 权重为 0.016。

当调度目标侧重于城镇层次时,得到 A_1 指标中 $A_{1,1} \sim A_{1,12}$ 权重为 0.008, $A_{1,13} \sim A_{1,15}$ 权重为 0.008, $A_{1,16} \sim A_{1,20}$ 权重为 0.008; A_2 指标中 $A_{2,1} \sim A_{2,12}$ 权重为 0.008, $A_{2,13} \sim A_{2,15}$ 权重为 0.008, $A_{2,16} \sim A_{2,20}$ 权重为 0.008, $A_{2,21} \sim A_{2,23}$ 权重为 0.2; B_1 指标中 $B_{1,1} \sim B_{1,4}$ 权重为 0.008; B_2 指标中 $B_{2,1} \sim B_{2,5}$ 权重为 0.008; B_3 指标中 $B_{3,1} \sim B_{3,3}$ 权重为 0.008。

4 评价方法

4.1 评价流程

利用 KNN-RF 组合模型对太湖流域多层次防洪调度方案进行评价,过程为:①基于已构建的太湖流域多层次防洪调度方案评价指标体系和模糊数学评价方法,将太湖流域 1999 年和 2016 年超标洪水情景的调度方案作为评价实例;②将已有评分的调度方案分为训练集和测试集,利用训练集训练 RF 和 KNN 模型,并建立 KNN-RF 组合模型;③利用测试集评估 KNN、RF 以及 KNN-RF 组合模型的调度方案评价精度,评价精度验证通过后得到 KNN-RF 组合模型。综合评价流程如图 3 所示。

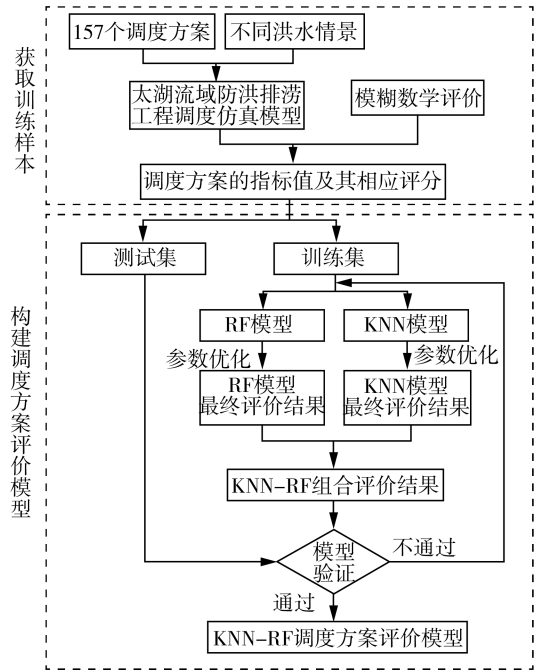


图3 综合评价流程

Fig.3 Evaluation flow

4.2 方案评分

基于太湖流域多层次防洪调度方案评价指标体系,结合模糊数学方法,对 1999 年和 2016 年超标洪水情景下,流域、区域和城镇 3 个层次的太湖流域调度方案进行评分及分级。优、良、一般、较差、差 5 个等级的评分标准分别为 >74 、 $>58 \sim 74$ 、 $>42 \sim 58$ 、 $>26 \sim 42$ 、 ≤ 26 。1999 年共 77 个调度方案,2016 年共 80 个调度方案,当调度目标侧重于不同的层次时,同一调度方案有不同的评分和等级,具体结果不再赘述。

4.3 模型参数选取

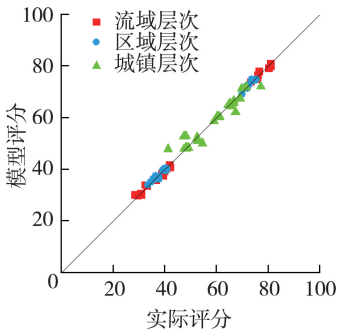
4.3.1 构建训练集和测试集

将太湖流域 2016 和 1999 年超标洪水情景共 157 个调度方案的指标值及其在流域、区域、城镇 3

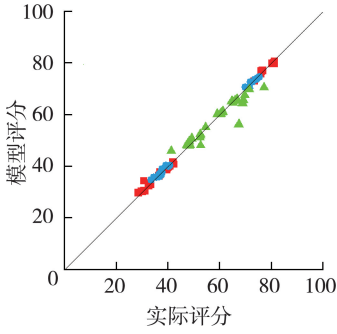
表 1 各模型评价精度指标值

Table 1 Evaluation accuracy index values of each model

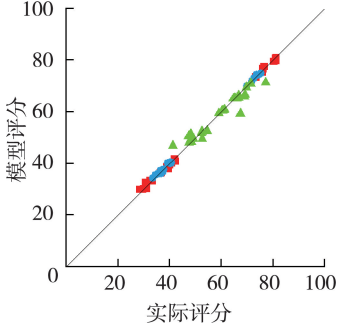
模型	调度目标	MAE	RMSE	MRE/%	R^2
KNN 模型	流域	0.581	0.823	1.372	0.998
	区域	0.448	0.577	1.048	0.998
	城镇	1.469	2.436	2.671	0.927
RF 模型	流域	0.575	0.912	1.467	0.998
	区域	0.443	0.621	1.001	0.998
	城镇	1.600	2.854	2.674	0.900
KNN-RF 组合模型	流域	0.512	0.725	1.252	0.999
	区域	0.342	0.440	0.820	0.999
	城镇	1.384	2.424	2.430	0.928



(a) KNN 模型



(b) RF 模型



(c) KNN-RF 组合模型

图 4 实际评分与模型评分的散点图

Fig.4 Scatter plots of actual score and model evaluation score

采用上述指标进行模型可靠性评估时,可能会忽略模型对个别调度方案的异常评价误差,因此将各模型对测试集进行评价的相对误差和绝对误差绘制为箱线图,如图 5 所示。由图 5 可知,3 个模型对 75% 以上调度方案评分的相对误差小于 3.5%、绝

种类型权重下的评分数据作为样本集,各调度方案指标值为输入,评分为输出。将训练集和测试集按 8:2 的比例进行划分,即训练集包括 125 个调度方案,测试集包括 32 个调度方案。训练集样本用于建立 KNN 模型和 RF 模型,测试集用于检验模型的评价性能。调度方案中水位、水量指标量纲不同,为避免对评价结果造成影响,对指标数据进行归一化处理。

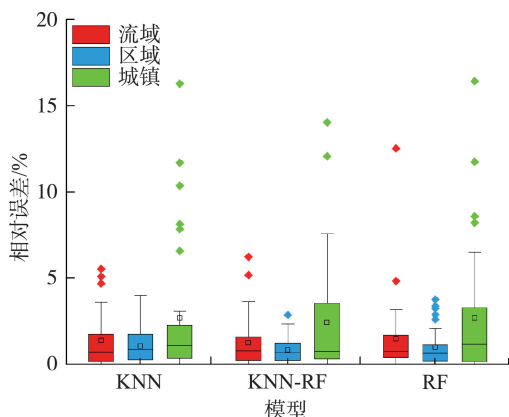
4.3.2 参数选取

本文在开源的 python 平台上进行 RF 和 KNN 模型的构建^[42]。构建 RF 模型需调节的参数主要包括决策树数量 $n_estimators$ 、分割节点考虑的最大特征数 $max_features$ 和最大深度 max_depth 。根据相关研究^[43-44], $max_features$ 不宜太小;对参数 $n_estimators$,可增加决策树棵树,直到得到稳定的结果;最大深度 max_depth 的常见取值为 10 ~ 100。基于 scikit-learn 框架的网格搜索算法^[45]对 RF 模型的参数进行优化,网格搜索算法是一种穷举算法,可以模拟不同参数的组合多次训练模型,并通过交叉验证对比误差从而得到当前训练数据的最佳参数组合,最终得到参数 $max_features$ 取 40, $n_estimators$ 取 2 000, max_depth 取 20。

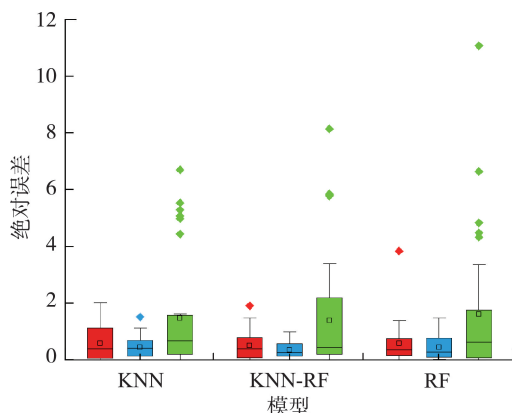
构建 KNN 模型需要设置的参数主要为 K 值、距离公式和搜索算法。其中, K 值是最重要的参数, K 值太大表示模型复杂度较低,可能引起欠拟合,不同的距离公式和搜索算法也会影响模型的评价精度^[20,24]。利用网格搜索算法进行参数优化选取,最终得到参数为 K 值取 3,距离公式为欧氏距离,搜索算法选择 $ball_tree$ 。

5 结果与分析

针对流域、区域和城镇不同层次防洪侧重,基于已构建的 KNN 模型、RF 模型和 KNN-RF 组合模型对测试集的调度方案进行评价,结果如表 1 所示;基于测试集的实际评分与各模型评分作散点图,如图 4 所示。结合表 1 和图 4 可知,在使用相同数据集的情况下,各个模型的 R^2 均在 0.9 以上,说明模型能较好地拟合数据,从训练集学习评价规律并实现测试集未知评分调度方案的评价。当调度目标侧重于流域和区域层次时,3 个模型的 MAE、RMSE、RME 均较小,MAE 在 0.5 左右,MRE 均控制在 1.5% 以内,而 R^2 能达到 0.99,表明模型评分接近实际评分,模型能很好地评价调度方案在流域层面和区域层面的协同性优劣;相比前两者,当调度目标侧重于城镇层次时,各模型的评价精度略低,但也在合理的误差范围内。KNN-RF 组合模型的各项评价精度指标均为最佳,在调度方案评价上整体表现最好。



(a) 相对误差



(b) 绝对误差

图5 各模型评价的相对误差和绝对误差

Fig.5 Relative error and absolute error of each model evaluation

对误差小于 2.5,表明模型的确能较好地评价大多数调度方案在各层次调度目标下的优劣。为进一步分析各模型的稳定性,重点关注模型的最大评价误差。相比于流域、区域层次,当调度目标侧重于城镇层次时,模型对个别调度方案评价误差较大,如 KNN 模型最大相对误差为 16.28%,最大绝对误差为 6.7;而 RF 模型在调度目标侧重于流域和城镇层次时,最大绝对误差和最大相对误差均高于 KNN 模型,说明 RF 模型总体评价性能略弱于 KNN 模型。需要指出的是,当调度目标侧重于城镇层次时,与 KNN 模型相比,KNN-RF 组合模型的最大绝对误差略有增加,这是因为 KNN 模型在这方面评价性能明显优于 RF 模型,导致 KNN-RF 组合模型的误差校正作用不明显,这和 Hou 等^[32]的研究类似。但是,结合其他指标整体来看,KNN-RF 组合模型在城镇层次的评价性能并未明显弱于 KNN 模型,针对城镇防洪侧重目标,KNN-RF 组合模型评价的最大相对误差相比 KNN 模型降低了 2.23%。且针对区域防洪侧重目标,KNN-RF 组合模型对调度方案评价的最大相对误差和最大绝对误差均得到了改善,分别

降低至 2.87% 和 0.99;针对流域防洪侧重目标,KNN-RF 组合模型评价的最大绝对误差降低至 1.9,模型组合后的总体评价性能得到提高。

根据调度方案的等级划分标准和 KNN-RF 组合模型给调度方案的评分,对测试集调度方案划分等级,与实际评级进行对比,进一步验证 KNN-RF 组合模型的评价性能。KNN-RF 组合模型在侧重流域、区域、城镇层次的评级误判率分别为 3.125%、3.125%、6.25%,总体正确率为 95.83%,说明 KNN-RF 模型能够对调度方案进行较为准确的等级划分。可见,KNN-RF 组合模型能够进行快速、可靠的多层次防洪调度方案综合评价。

6 结 论

a. 构建了太湖流域针对流域、区域、城镇不同层次的防洪调度指标体系,结合 KNN 模型和 RF 模型,构建 KNN-RF 组合模型对调度方案进行评价。

b. KNN-RF 组合模型对流域、区域与城镇 3 个不同层次的防洪侧重目标,MRE 和 MAE 分别降低至 1.25%、0.82%、2.43% 和 0.511、0.342、1.380,其评价精度较单一算法模型有显著提高。

c. 与单一算法模型相比,KNN-RF 组合模型能改善调度方案评价的最大相对误差和最大绝对误差;KNN-RF 组合模型评价的总体正确率达到了 95.83%。

参考文献:

- [1] ZHOU Zhengzheng, LIU Shuguang, ZHONG Guihui, et al. Flood disaster and flood control measurements in Shanghai [J]. Natural Hazards Review, 2017, 18 (1): B5016001.
- [2] PATTISON I, LANE S N, HARDY R J, et al. The role of tributary relative timing and sequencing in controlling large floods [J]. Water Resources Research, 2014, 50 (7): 5444-5458.
- [3] OPPERMAN J J, GALLOWAY G E, FARGIONE J, et al. Sustainable floodplains through large-scale reconnection to rivers [J]. Science, 2009, 326(5959): 1487-1488.
- [4] 中华人民共和国水利部, 中华人民共和国国家统计局. 第一次全国水利普查公报 [J]. 中国水利, 2013(7): 1-3.
- [5] 钟桂辉, 刘曙光, 娄厦, 等. 苏锡常排涝对江南运河产生的洪涝转移特征 [J]. 水利水电科技进展, 2020, 40 (5): 40-47. (ZHONG Guihui, LIU Shuguang, LOU Sha, et al. Flood dispatching characteristics of Jiangnan Canal caused by drainage from Suzhou, Wuxi and Changzhou [J]. Advances in Science and Technology of Water Resources, 2020, 40(5): 40-47. (in Chinese))
- [6] 叶建春, 章杭惠. 太湖流域洪水风险管理实践与思考 [J]. 水利水电科技进展, 2015, 35 (5): 136-141. (YE

- Jianchun, ZHANG Hanghui. Practices and thinking of flood risk management in Taihu Lake Basin[J]. Advances in Science and Technology of Water Resources, 2015, 35(5):136-141. (in Chinese)
- [7] 李俊,武鹏林. 基于改进 TOPSIS 多目标水库防洪调度方案决策[J]. 人民黄河, 2017, 39(6):34-37. (LI Jun, WU Penglin. Decision-making of flood control scheme for multipurpose reservoir based on improved TOPSIS[J]. Yellow River, 2017, 39(6):34-37. (in Chinese))
- [8] 周建中, 顿晓晗, 张勇传. 基于库容风险频率曲线的水库群联合防洪调度研究[J]. 水利学报, 2019, 50(11):1318-1325. (ZHOU Jianzhong, DUN Xiaohan, ZHANG Yongchuan. Study on joint flood control dispatching of reservoir groups based on reservoir capacity risk frequency curve[J]. Journal of Hydraulic Engineering, 2019, 50(11):1318-1325. (in Chinese))
- [9] 黄显峰, 吴志远, 李昌平, 等. 基于改进粒子群-逐次逼近法的水库调度图多目标优化[J]. 水利水电科技进展, 2021, 41(2):1-7. (HUANG Xianfeng, WU Zhiyuan, LI Changping, et al. Multi-objective optimization of reservoir operation chart based on IPSO-DPSA [J]. Advances in Science and Technology of Water Resources, 2021, 41(2):1-7. (in Chinese))
- [10] 关宏艳, 李宗坤, 葛巍, 等. 于加权广义马氏距离的 TOPSIS 方法在水库防洪调度决策中的应用[J]. 天津大学学报(自然科学与工程技术版), 2016, 49(12):1276-1281. (GUAN Hongya, LI Zongkun, GE Wei, et al. TOPSIS method based on weighted generalized mahalanobis distance: an application to reservoir flood control operation [J]. Journal of Tianjin University (Science and Technology), 2016, 49(12):1276-1281. (in Chinese))
- [11] 吴成国, 王义民, 金菊良, 等. 基于三角模糊数的水库防洪调度模糊综合风险分析[J]. 水力发电学报, 2011, 30(4):30-35. (WU Chengguo, WANG Yimin, JIN Juliang, et al. Comprehensive risk analysis based on triangular fuzzy number for reservoir flood control operation [J]. Journal of Hydroelectric Engineering, 2011, 30(4):30-35. (in Chinese))
- [12] 董增川, 陈牧风, 倪效宽, 等. 考虑模糊区间的水库群优化调度决策方法[J]. 河海大学学报(自然科学版), 2021, 49(3):233-240. (DONG Zengchuan, CHEN Mufeng, NI Xiaokuan, et al. Decision method of optimal operation of reservoir group considering fuzzy interval [J]. Journal of Hohai University (Natural Sciences), 2021, 49(3):233-240. (in Chinese))
- [13] 卢有麟, 陈金松, 祁进, 等. 基于改进熵权和集对分析的水库多目标防洪调度决策方法研究[J]. 水电能源科学, 2015, 33(1):43-46. (LU Youlin, CHEN Jinsong, QI Jin, et al. Multi-objective decision making method for reservoir flood operation based on improved entropy weight and set pair analysis [J]. Water Resources and Power, 2015, 33(1):43-46. (in Chinese))
- [14] 王丽萍, 叶季平, 苏学灵, 等. 基于可拓学理论的防洪调度方案评价研究与应用[J]. 水利学报, 2009, 40(12):1425-1431. (WANG Lingping, YE Jiping, SU Xueling, et al. Evaluation of flood control operation program based on extenics theory and its application [J]. Journal of Hydraulic Engineering, 2009, 40(12):1425-1431. (in Chinese))
- [15] 王磊之, 胡庆芳, 戴晶晶, 等. 面向金泽水库取水安全的太湖浦多目标联合调度研究[J]. 水资源保护, 2017, 33(5):61-68. (WANG Leizhi, HU Qingfang, DAI Jingjing, et al. Research on multi-objective joint operation of Taihu River oriented to water supply safety of Jinze Reservoir[J]. Water Resources Protection, 2017, 33(5):61-68. (in Chinese))
- [16] 王本德, 周惠成, 卢迪. 我国水库(群)调度理论方法研究应用现状与展望[J]. 水利学报, 2016, 47(3):337-345. (WANG Bende, ZHOU Huicheng, LU Di. The state-of-the-art review in the theory research and its application of multi-reservoir system operation in China[J]. Journal of Hydraulic Engineering, 2016, 47(3):337-345. (in Chinese))
- [17] TAHMASEBI P, KAMRAVA S, BAI T, et al. Machine learning in geo- and environmental sciences: from small to large scale[J]. Advances in Water Resources, 2020, 142:103619.
- [18] XU Tianfang, LIANG Feng. Machine learning for hydrologic sciences: an introductory overview[J]. Wiley Interdisciplinary Reviews: Water, 2021, 8(5):e1533.
- [19] TYRALIS H, PAPACHARALAMPOUS G, LANGOUSIS A. A brief review of random forests for water scientists and practitioners and their recent history in water resources[J]. Water, 2019, 11(5):910.
- [20] COSTACHE R, PHAM Q B, SHARIFI E, et al. Flash-flood susceptibility assessment using multi-criteria decision making and machine learning supported by remote sensing and GIS techniques[J]. Remote Sensing, 2020, 12(1):106.
- [21] WANG Zhaoli, LAI Chengguang, CHEN Xiaohong, et al. Flood hazard risk assessment model based on random forest[J]. Journal of Hydrology, 2015, 527:1130-1141.
- [22] 张珂, 牛杰帆, 李曦, 等. 洪水预报智能模型在中国半干旱半湿润区的应用对比[J]. 水资源保护, 2021, 37(1):28-35. (ZHANG Ke, NIU Jiefan, LI Xi, et al. Comparison of artificial intelligence flood forecasting models in China's semi-arid and semi-humid regions[J]. Water Resources Protection, 2021, 37(1):28-35. (in Chinese))
- [23] LU Y, QIN X S, XIE Y J. An integrated statistical and data-driven framework for supporting flood risk analysis

- under climate change[J]. Journal of Hydrology, 2016, 533:28-39.
- [24] NAGHIBI S A, VAFAKHAH M, HASHEMI H, et al. Water resources management through flood spreading project suitability mapping using frequency ratio, K-nearest neighbours, and random forest algorithms[J]. Natural Resources Research, 2020, 29:1915-1933
- [25] 王盼, 陆宝宏, 张瀚文, 等. 基于随机森林模型的需水预测模型及其应用[J]. 水资源保护, 2014, 30(1):34-37. (WANG Pan, LU Baohong, ZHANG Hanwen, et al. Water demand prediction model based on random forests model and its application [J]. Water Resources Protection, 2014, 30(1):34-37. (in Chinese))
- [26] 吕向林, 姬世保, 仇亚琴, 等. 汾河流域上游基流特征及其影响因素[J]. 水资源保护, 2022, 38(2):147-153. (LYU Xianglin, JI Shibao, QIU Yaqin, et al. Characteristics of base flow in upstream of Fen River Basin and its influencing factors [J]. Water Resources Protection, 2022, 38(2):147-153. (in Chinese))
- [27] 崔东文, 金波. 基于随机森林回归算法的水生态文明综合评价[J]. 水利水电科技进展, 2014, 34(5):56-60. (CUI Dongwen, JIN Bo. Comprehensive evaluation of water ecological civilization based on random forests regression algorithm [J]. Advances in Science and Technology of Water Resources, 2014, 34(5):56-60. (in Chinese))
- [28] 闫佰忠, 孙剑, 安娜. 基于随机森林模型的地下水水质评价方法[J]. 水电能源科学, 2019, 37(11):66-69. (YAN Baizhong, SUN Jian, AN Na. Assessment of groundwater quality based on random forest model [J]. Water Resources and Power, 2019, 37(11):66-69. (in Chinese))
- [29] 林威伟, 钟登华, 胡炜, 等. 基于随机森林算法的土石坝压实质量动态评价研究[J]. 水利学报, 2018, 49(8):945-955. (LIN Weiwei, ZHONG Denghua, HU Wei, et al. Study on dynamic evaluation of compaction quality of earth rock dam based on random forest [J]. Journal of Hydraulic Engineering, 2018, 49(8):945-955. (in Chinese))
- [30] BERKHAHN S, FUCHS L, NEUWEILER I. An ensemble neural network model for real-time prediction of urban floods[J]. Journal of Hydrology, 2019, 575:743-754.
- [31] ZHANG Kequan, QU Zongxi, DONG Yunxuan, et al. Research on a combined model based on linear and nonlinear features: a case study of wind speed forecasting [J]. Renewable Energy, 2019, 130:814-830.
- [32] HOU Jingming, ZHOU Nie, CHEN Guangzhao, et al. Rapid forecasting of urban flood inundation using multiple machine learning models [J]. Natural Hazards, 2021, 108:2335-2356.
- [33] 嵇晓燕, 杨凯, 陈亚男, 等. 基于 ARIMA 和 Prophet 的水质预测集成学习模型[J]. 水资源保护, 2022, 38(6):111-115. (JI Xiaoyan, YANG Kai, CHEN Yanan, et al. An ensemble learning model for water quality forecast based on ARIMA and Prophet [J]. Water Resources Protection, 2022, 38(6):111-115. (in Chinese))
- [34] 林鹏, 陈启慧, 李琼芳, 等. 环太湖各水资源分区入出湖水量及贡献分析[J]. 水资源保护, 2021, 37(3):66-73. (LIN Peng, CHEN Qihui, LI Qiongfang, et al. Analysis on inflow and outflow of water resources zones around Taihu Lake and its contribution [J]. Water Resources Protection, 2021, 37(3):66-73. (in Chinese))
- [35] 赵静, 胡庆芳, 王腊春, 等. 基于 MSWEP 数据的太湖流域降水特性分析[J]. 水资源保护, 2020, 36(2):27-33. (ZHAO Jing, HU Qingfang, WANG Lachun, et al. Analysis of precipitation characteristics in Taihu Lake Basin based on MSWEP [J]. Water Resources Protection, 2020, 36(2):27-33. (in Chinese))
- [36] 张春松, 宋玉, 陶娜麒, 等. 江苏省苏南运河沿线地区联合调度实践与思考[J]. 中国防汛抗旱, 2018, 28(3):4-6. (ZHANG Chunsong, SONG Yu, TAO Naqi, et al. Practice and thinking on joint dispatching of South Jiangsu Canal in Jiangsu Province [J]. China Flood & Drought Management, 2018, 28(3):4-6. (in Chinese))
- [37] 梅青. 太湖流域 2016 年防汛抗洪工作及 2017 年工作展望[J]. 中国防汛抗旱, 2017, 27(1):45-48. (MEI Qing. Flood control work in Taihu Basin in 2016 and work prospects in 2017 [J]. China Flood & Drought Management, 2017, 27(1):45-48. (in Chinese))
- [38] 徐天奕, 刘克强, 李琛, 等. 太湖流域大尺度洪涝淹没仿真模型的建立及应用[J]. 水利水电科技进展, 2021, 41(4):40-45. (XU Tianyi, LIU Keqiang, LI Chen, et al. Establishment and application of large-scale flood inundation simulation model for Taihu Basin [J]. Advances in Science and Technology of Water Resources, 2021, 41(4):40-45. (in Chinese))
- [39] KELLER J M, GRAY M R, GIVENS J A. A fuzzy K-nearest neighbor algorithm [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1985, 15(4):580-585.
- [40] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45(1):5-32.
- [41] ZOUNEMAT-KERMANI M, BATELAAN O, FADAEI M, et al. Ensemble machine learning paradigms in hydrology: a review [J]. Journal of Hydrology, 2021, 598:126266.
- [42] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, et al. Scikit-learn: machine learning in python [J]. Journal of Machine Learning Research, 2011, 12:2825-2830.
- [43] BIAU G, SCORNET E. A random forest guided tour [J]. Test, 2016, 25(2):197-227.

(下转第 236 页)

- 学报,2018,12(8):2193-2201. (ZHONG Xing,ZHANG Wei,TIAN Xiaochen,et al. Runoff pollutants purification effect of bioretention media improved by black-odorous river sediments [J]. Chinese Journal of Environmental Engineering,2018,12(8):2193-2201. (in Chinese))
 - [11] LIU Mingwei, LIU Xiao, WANG Weiguo, et al. Effect of SiO₂ and Al₂O₃ on characteristics of lightweight aggregate made from sewage sludge and river sediment[J]. Ceramics International,2018,44(4):4313-4319.
 - [12] 肖继波,陈玉莹,瞿倩,等. 高效除磷型底泥陶粒对水体中磷的去除特性[J]. 环境污染与防治,2018,40(10):1122-1125. (XIAO Jibo, CHEN Yuying, QU Qian, et al. Phosphorus removal properties of sediment-ceramsite with high phosphorus removal efficiency [J]. Environmental Pollution & Control, 2018, 40 (10): 1122-1125. (in Chinese))
 - [13] 陈玉莹. 底泥陶粒的研制及其在生物滤池中的应用研究[D]. 温州:温州大学,2017.
 - [14] XIA Ting, LI Yue, XIAO Zhixing, et al. A study on ceramsite production using dredging sea mud and its biofilm formation capacity evaluation [J]. Advances in Civil Engineering,2018,2018:2401984.
 - [15] 张炜,王浩远,赵玉华,等. 基于 hydrus-1d 生物滞留设施雨水径流水文调控模拟[J]. 水资源保护,2022,38(3):102-108. (ZHAN Wei, WANG Haoyuan, ZHAO Yuhua, et al. Hydrological regulation simulation of rainwater runoff of bioretention based on HYDRUS-1D [J]. Water Resources Protection,2022,38(3):102-108. (in Chinese))
 - [16] ZHANG Hongwei, AHMAD Z, SHAO Yalu, et al. Bioretention for removal of nitrogen: processes, operational conditions, and strategies for improvement [J]. Environmental Science and Pollution Research, 2021,28(9):10519-10535.
 - [17] ZHU T, JENSSEN P D, MAHLUM T, et al. Phosphorus sorption and chemical characteristics of lightweight aggregates (LWA): potential filter media in treatment wetlands[J]. Water Science and Technology, 1997, 35 (5):103-108.
 - [18] 杜晓丽,韩强,于振亚,等. 海绵城市建设中生物滞留设施应用的若干问题分析[J]. 给水排水,2017,43(1):54-58. (DU Xiaoli, HAN Qiang, YU Zhenya, et al. Some issues of the bioretention application in sponge city construction[J]. Water & Wastewater Engineering,2017, 43(1):54-58. (in Chinese))
 - [19] 王海邻,曹雪莹,任玉芬,等. 北京城市主干道降雨径流污染负荷分析[J]. 环境科学学报,2019,39(6):1860-1867. (WANG Hailin, CAO Xueying, REN Yufen, et al. Analysis on pollution load of rainfall runoff in Beijing urban arterial roads [J]. Acta Scientiae Circumstantiae, 2019,39(6):1860-1867. (in Chinese))
 - [20] 欧阳威,王玮,郝芳华,等. 北京城区不同下垫面降雨径流产污特征分析[J]. 中国环境科学,2010,30(9):1249-1256. (OUYANG Wei, WANG Wei, HAO Fanghua, et al. Pollution characterization of urban stormwater runoff on different underlying surface conditions [J]. China Environmental Science, 2010, 30 (9): 1249-1256. (in Chinese))
 - [21] OSMAN M, WAN Y K, TAKAIJUDIN H, et al. A review of nitrogen removal for urban stormwater runoff in bioretention system [J]. Sustainability, 2019, 11 (19): 5415-5436.
 - [22] 刘芳. 秸秆生物炭生物滞留改良填料研制与参数优化 [D]. 西安:西安理工大学,2020.
 - [23] MARVIN J T, PASSEPORT E, DRAKE J. State-of-the-art review of phosphorus sorption amendments in bioretention media: a systematic literature review [J]. Journal of Sustainable Water in the Built Environment, 2020, 6 (1):03119001.
 - [24] 张飞剑,彭云清,章茹. 潮汐流人工湿地处理南方农村黑臭水体实验研究[J]. 水处理技术,2022,48(4):114-118. (ZHANG Feijian, PENG Yunqing, ZHANG Ru. Experimental study on the black and odorous water body in southern rural areas by tidal flow constructed wetland [J]. Technology of Water Treatment,2022,48(4):114-118. (in Chinese))
- (收稿日期:2022-06-21 编辑:熊水斌)
- +++++
- (上接第 125 页)
- [44] PROBST P,BOULESTEIX A L. To tune or not to tune the number of trees in random forest[J]. Journal of Machine Learning Research, 2018, 18:6673-6690.
 - [45] 周赓,侯精明,陈光照,等. 基于机器学习的山洪灾害快速预报方法[J]. 水资源保护,2022,38(2):32-40. (ZHOU Nie, HOU Jinming, CHEN Guangzhao, et al. A rapid forecasting method for mountain flood disaster based on machine learning algorithm [J]. Water Resources Protection, 2022,38(2):32-40. (in Chinese))
- (收稿日期:2022-04-03 编辑:王芳)

