

# 基于随机森林-投影寻踪法的生物滞留系统多目标评价方法

程麒铭<sup>1</sup>, 陈 垚<sup>1,2</sup>, 刘 臻<sup>1,2</sup>, 唐颖辉<sup>1</sup>, 袁绍春<sup>1,2</sup>

(1. 重庆交通大学 河海学院, 重庆 400074; 2. 重庆交通大学 环境水利工程重庆市工程实验室, 重庆 400074)

**摘要:** 生物滞留系统性能受植物和介质土的影响显著,且各性能指标具有多样性和不相容性,而传统评价方法具有较强的主观性易导致评价结果出错。采用随机森林模型对原始数据进行特征筛选以降低数据维度,构建投影寻踪模型对不同植物和土壤的水力渗透性能和污染物去除性能进行多目标评价,并利用高鲁棒性的遗传算法(GA)和粒子群算法(PSO)进行模型求解。植物评价结果表明,风车草为生物滞留系统的最佳植物,且评价结果与层次分析模型和BP神经网络模型方法相似;介质土评价结果表明,RST2(9.8%壤砂土+88.2%细砂+2%蛭石)为生物滞留系统的最佳介质土配置方案,且评价结果与传统投影寻踪法相似。研究结果证实了随机森林-投影寻踪(RF-PP)模型适用于生物滞留系统多目标评价。

**关键词:** 生物滞留系统; 投影寻踪法; 随机森林; 多目标评价; 遗传算法; 粒子群算法

中图分类号: X52; TV877

文献标识码: A

文章编号: 1672-643X(2022)04-0085-06

## Multi-objective evaluation method of bioretention system based on random forest - projection pursuit method

CHENG Qiming<sup>1</sup>, CHEN Yao<sup>1,2</sup>, LIU Zhen<sup>1,2</sup>, TANG Yinghui<sup>1</sup>, YUAN Shaochun<sup>1,2</sup>

(1. School of River and Ocean Engineering, Chongqing Jiaotong University, Chongqing 400074, China;

2. Engineering Laboratory of Environmental Hydraulic Engineering of Chongqing Municipal Development and Reform Commission, Chongqing Jiaotong University, Chongqing 400074, China)

**Abstract:** The performance of the bioretention system is significantly affected by plants and soil media, and its various diversified performance indexes are incompatible. However, the conventional evaluation methods are flawed by their strong subjectivity, which could easily lead to the errors of evaluation results. Here, the random forest (RF) model was used to screen the original data so as to reduce the dimensionality of the data. Then, the projection pursuit (PP) model was constructed to evaluate the hydraulic permeability and pollutant removal performance of different plants and soil media for multi-objective evaluation, and the model was solved by genetic algorithm (GA) and particle swarm optimization (PSO). The results of plant evaluation showed that *Cyperus alternifolius* L. was the optimal plant for the bioretention system, which was similar to the results of the analytic hierarchy process model and the back propagation (BP) neural network model. The evaluation results of soil media showed that RST2 (9.8% loamy sand + 88.2% fine sand + 2.0% vermiculite) was the optimal soil media configuration for the bioretention system, which was similar to the conclusion of the conventional projection pursuit method. These findings indicate that RF-PP model is suitable for multi-objective evaluation of the bioretention system.

**Key words:** bioretention system; projection pursuit (PP); random forest (RF); multi-objective evaluation; genetic algorithm (GA); particle swarm optimization (PSO)

收稿日期:2021-10-19; 修回日期:2022-06-17

基金项目:重庆市研究生联合培养基地项目(JDLHPYJD2019);国家自然科学基金项目(51709024);重庆市建设科技计划项目(城科字2020第5-7);重庆市自然科学基金项目(cstc2020jcyj-msxmX0716、cstc2020jcyj-msxmX1000)

作者简介:程麒铭(1996-),男,重庆永川人,硕士研究生,研究方向为水污染防治技术研究。

通讯作者:陈 垚(1983-),男,浙江永康人,博士,教授,博士生导师,研究方向为水污染防治与城市雨洪管理。

## 1 研究背景

生物滞留系统是海绵城市建设的主要技术措施之一,其可通过植物吸收、填料过滤和微生物降解等多种作用实现对径流量的削减和污染物的控制<sup>[1-2]</sup>,植物与介质土均是组成生物滞留系统的重要部分<sup>[3]</sup>,可直接影响生物滞留系统的运行性能。因此,植物和介质土的筛选是生物滞留系统设计和运行的重要考量部分<sup>[4]</sup>。由于不同植物和介质土对生物滞留系统除污能力和水力性能的改善效果不尽相同,使得评价指标数据样本之间存在高度非线性关系,且原始实验数据通常具有较高的维度。研究表明<sup>[5]</sup>,高维度、非线性数据样本的数值分析容易造成维数灾难,导致计算工作量成倍增加。目前,国内外均未形成统一的评价方法,通常采用层次分析法、模糊综合评价法等<sup>[6]</sup>,但此类方法存在评价结果主观性较强,且模型假设和概化对评价结果影响较大的问题<sup>[7]</sup>。

为寻求更客观、高效的多目标综合评价方法,研究者们尝试采用神经网络模型(如 Elman 神经网络<sup>[8]</sup>、模糊神经网络<sup>[9]</sup>和 BP 神经网络<sup>[10]</sup>)的机器学习功能实现对水体水质的多目标评价,但结果显示评价具有偶然性而导致结果可信度不高。而采用贝叶斯模型<sup>[11]</sup>、基于改进萤火虫算法的投影寻踪模型<sup>[12-13]</sup>等统计学模型,则易发生“弃真”和“取伪”的错误<sup>[14]</sup>。同时,上述方法还存在典型的维数灾难问题,导致计算量呈指数增长,甚至造成计算机无法运行。闫佰忠等<sup>[15]</sup>采用随机森林模型的机器学习功能对安阳市 8 个地下水监测点的水质检测数据进行处理,证实基于随机森林模型的水质评价模型能够准确评价水质的同时,拥有更高训练效率与稳定性。但该方法主要基于模型的分类原理,并未达到降低数据维度的目的,导致计算工作量较大。而基于特征选择功能的随机森林模型可达到数据降维的目的,在保证数据分析结果准确性的同时大幅提高计算速度。为此,本文采用随机森林(random forest, RF)模型对不同植物、介质土下的水力性能和污染物去除率等原始数据进行特征筛选以降低数据维度,构建基于投影寻踪(projection pursuit, PP)模型的多目标综合评价模型(RF-PP),并通过遗传算法(genetic algorithm, GA)和粒子群算法(particle swarm optimization, PSO)对 RF-PP 模型进行求解,最后将评价结果与传统评价方法进行对比分析,以验证本文方法的准确性,从而提出一种适合生物滞

留系统多目标评价的方法。

## 2 数据来源与研究方法

### 2.1 数据来源

本研究所用数据分为两组,其中 A 组数据来自课题组 2017 年生物滞留系统植物筛选实验数据,滤柱采用口径 150 mm 的 PVC(滤料层)与亚克力材料(蓄水层)制成,种植层采用河沙、石英砂、沙壤土(质量比为 2:2:1)混合而成。该组实验暂不考虑不同介质土的影响,在相同种植层条件下种植不同植物进行实验,其余实验方法等详见文献[16]。A 组实验原始数据为 132 个实验样本,以及各植物对悬浮物(SS)、化学需氧量(COD)、氨氮(NH<sub>3</sub>-N)、硝氮(NO<sub>3</sub><sup>-</sup>-N)、总氮(TN)、总磷(TP)等 6 个水质污染物去除率百分比。B 组数据来自课题组 2020 年生物滞留系统介质土筛选实验数据,滤柱与 A 组实验相同,种植层采用不同介质土,而不考虑不同植物的影响,且在无植物条件下进行实验,其余实验方法等详见文献[17]。B 组原始数据为 72 个实验样本,以及各介质土对 COD、NH<sub>3</sub>-N、NO<sub>3</sub><sup>-</sup>-N、TN、TP 等 5 个水质污染物去除率百分比和水力渗透系数(hydraulic permeability coefficient, HPC)1 个水力性能指标(mm/h),原始数据情况详见表 1。

### 2.2 模型原理

(1)随机森林特征选择。随机森林(RF)算法属于机器学习算法的一种,最早由美国统计学家 Leo Breiman 提出,其可通过袋外数据误差估计(out of bag error estimation)衡量每个数据特征对模型拟合效果的重要性,从而实现研究数据的特征选择<sup>[18]</sup>,具体步骤详见文献[19]、[20]。

(2)投影寻踪模型。投影寻踪(PP)法是对采集到的样本数据进行高维处理和分析的新型统计学方法<sup>[21]</sup>,该方法采用计算机对低维空间进行分析,以描述高维空间的数据信息,通过对高维空间的数据再整合,并将数据引入低维子空间进行降维处理,得到指标量远小于样本数据量的投影指标,从而寻找出高维样本数据的整体特征<sup>[22-23]</sup>。在投影寻踪模型中,投影方向代表不同的高维数据结构或特征,其中最佳投影方向最能体现高维数据的结构特征,而求解最佳投影方向就是对投影指标函数进行最大化寻优,具体方法详见文献[24]。本研究首先利用 RF 算法对原始数据进行降维后利用 PP 方法进行模型构建,最终建立生物滞留系统多目标评价 RF-PP 模型。

表 1 本研究所采用的原始数据情况

数据名	研究对象	评价指标	样本数
数据 A	植物: 美人蕉、梭鱼草、千屈草、芦竹、风车草、鸢尾、香根草、纸莎草、细叶芒、狗牙根	SS、COD、NH <sub>3</sub> -N、NO <sub>3</sub> <sup>-</sup> -N、TN、TP	132
	介质土: RST1(10.0% 壤砂土 + 90.0% 细砂)		
	RST2(9.8% 壤砂土 + 88.2% 细砂 + 2.0% 蛭石)		
	RST3(9.8% 壤砂土 + 88.2% 细砂 + 2.0% 生物炭)		
数据 B	RST4(9.8% 壤砂土 + 88.2% 细砂 + 2.0% 珍珠岩)	COD、NH <sub>3</sub> -N、NO <sub>3</sub> <sup>-</sup> -N、TN、TP、HPC	72
	ST1(20.0% 砂土 + 80.0% 细砂)		
	ST2(18.4% 砂土 + 73.6% 细砂 + 8.0% 蛭石)		
	ST3(18.4% 砂土 + 73.6% 细砂 + 8.0% 生物炭)		
	ST4(18.4% 砂土 + 73.6% 细砂 + 8.0% 珍珠岩)		

### 2.3 模型求解

由于不同评价指标具有不同的量纲和取值范围,为消除奇异样本数据导致的不良影响,通常采用归一化方法(如最大最小归一化、Z-score 归一化、非线性归一化等)对原始数据样本进行处理。根据生物滞留系统运行性能越大越优的目标,应采用最大最小归一化法:

$$S_{mn}^* = \frac{S_{mn} - S_{n,\min}}{S_{n,\max} - S_{n,\min}} \quad (1)$$

式中:  $m$  为各组实验样本数;  $n$  为样本指标数;  $S_{mn}$  为第  $m$  个样本的第  $n$  个指标;  $S_{n,\max}$  为所有样本第  $n$  个指标的最大值;  $S_{n,\min}$  为所有样本第  $n$  个指标的最小值;  $S_{mn}^*$  为第  $m$  个样本的第  $n$  个指标归一化结果。

由于投影寻踪模型属于非线性优化问题,通常可采用鱼群算法、粒子群算法、萤火虫算法、遗传算法等群体智能算法求解这类多极值点的复杂优化问题<sup>[25-26]</sup>。其中,遗传算法(GA)是受生物界进化规律启发,从初始群体中随机选择个体,将该个体作为父级生成下一代子级,通过一系列的交叉、变异等重复步骤后,对获得的样本进行适应性验证,最终演化为模型最优解,已广泛应用于各领域求解有约束或无约束的非线性优化问题<sup>[27-28]</sup>。粒子群算法(PSO)是受鸟群捕食行为的启发,将鸟群作为一个粒子群,将每只鸟作为一个粒子,根据各粒子自身所处的位置,将寻找食物的路径分享给群体,通过比较路径距离而不断改变路径,最终获得最优路径即模型最优解<sup>[29-31]</sup>。为了避免单一算法造成的偶然性结果,本研究采用 GA 和 PSO 两类算法求解构建的生物滞留系统多目标评价 RF-PP 模型,并对多目

标评价结果进行优劣排序。

## 3 结果分析与讨论

### 3.1 随机森林法结果分析

根据随机森林原理编写 Python 程序,运行后得到 A、B 两组数据中相关指标重要性程度的得分情况如图 1 所示,两组数据指标各自的相关系数如图 2 所示。

由图 1 可知,A 组数据重要程度最低的指标为 NH<sub>3</sub>-N(0.070)和 NO<sub>3</sub><sup>-</sup>-N(0.115),而重要程度最高的指标为 COD(0.340)。这是由于 TN 包括可溶性氮和悬浮颗粒携带的氮,其含量包含 NH<sub>3</sub>-N 和 NO<sub>3</sub><sup>-</sup>-N;且通过对原始数据进行相关性分析发现,TN 与 NH<sub>3</sub>-N、NO<sub>3</sub><sup>-</sup>-N 的相关性较高,相关系数  $R^2$  分别可达 0.55 与 0.83(图 2)。因此,随机森林算法认为 NH<sub>3</sub>-N 和 NO<sub>3</sub><sup>-</sup>-N 指标的重要程度较低,而由于 COD 与其他指标的相关性均较低,则表现出 COD 指标的重要程度较高。为降低数据维数,可剔除 A 组数据中评分最低的两组指标(NH<sub>3</sub>-N 和 NO<sub>3</sub><sup>-</sup>-N)。

B 组数据重要程度最低的指标为 NH<sub>3</sub>-N(0.130)和 TN(0.130),而重要程度最高的指标为 HPC(0.290),这是由于 NH<sub>3</sub>-N、TN 与 NO<sub>3</sub><sup>-</sup>-N 的相关性较高,相关系数  $R^2$  分别为 0.62 与 0.85(图 2),而 HPC 与其他指标的相关性均较低,使得随机森林算法给出了与 A 组数据相似的结论。同样,剔除 B 组数据中评分最低的两组指标 TN 与 NH<sub>3</sub>-N。

### 3.2 投影寻踪法结果分析

利用 MATLAB 遗传算法工具箱函数,并改写 PSO 工具箱函数分别运行 GA 算法和 PSO 算法以求

解本研究构建的 RF - PP 模型的优化问题<sup>[32]</sup>,从而得到各指标的最优投影系数  $a = a(a_1, a_2, \dots, a_m)$ ,

如图3所示。投影系数  $a$  越大,则表明该系数对应指标对评价结果的权重越大。

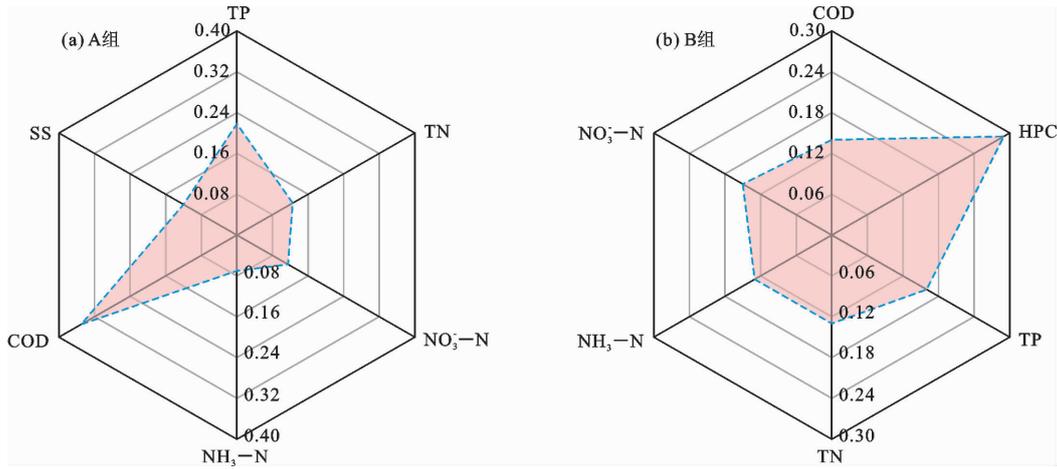


图1 A、B两组数据相关指标的重要程度得分值

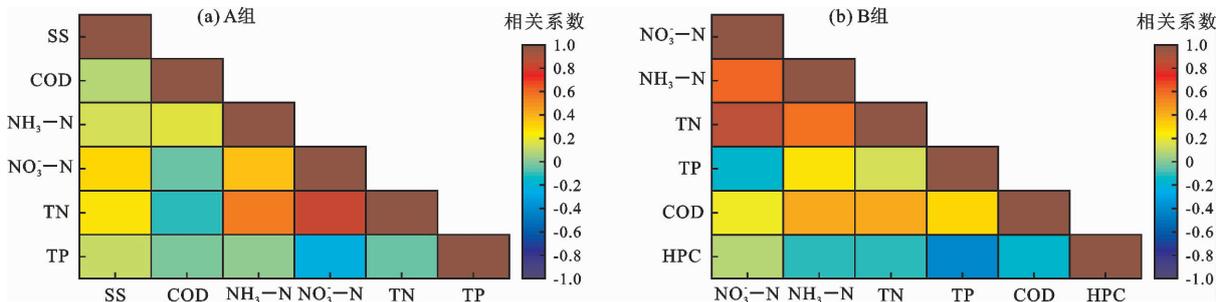


图2 A、B两组数据指标各自的相关系数

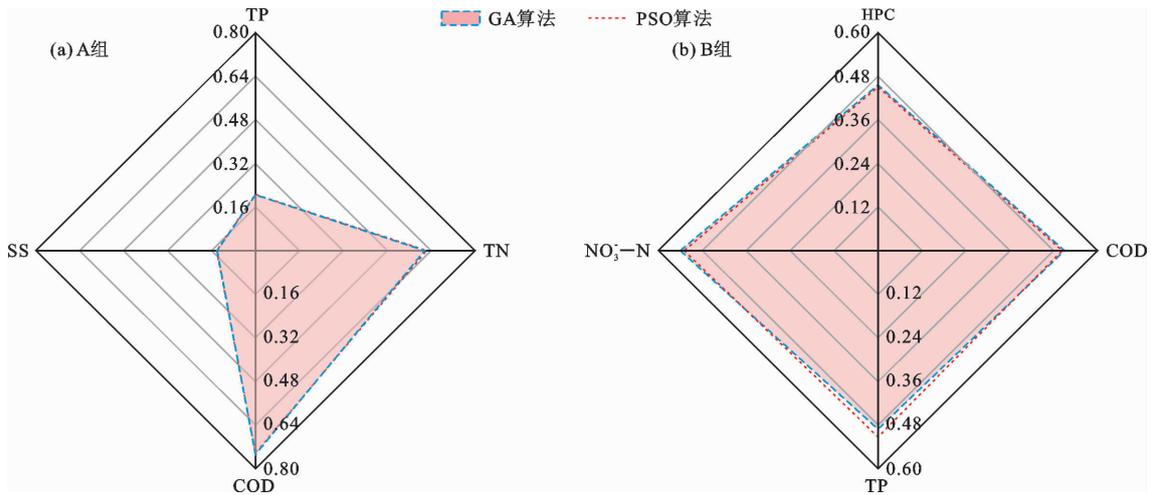


图3 A、B两组数据指标的最优投影系数

A、B两组数据分别采用 GA 算法和 PSO 算法求解投影寻踪模型的结果一致,均得出 A 组数据中 COD 指标的权重最大,其次为 TN 指标,而 TP 和 SS 指标的权重均小于 0.2;B 组数据中各指标的权重较均匀且小于 0.6,明显低于 A 组数据中 COD 和 TN 指标的权重,其中  $\text{NO}_3^- - \text{N}$  指标的权重最大,而

HPC 指标的权重最小。分析认为,生物滞留系统对 COD 和 TN 的去除受植物种类的影响显著<sup>[33]</sup>,而对 SS 和 TP 的去除能力则主要取决于介质土类型<sup>[34]</sup>。同时 A 组数据中植物还受温度、生长特性等因素影响,造成 COD 和 TN 两个指标的方差远大于 TP 和 SS 指标(表 2),从而表现出 A 组数据中 COD 和 TN

两个指标的权重明显高于其他两个指标。B 组数据由于生物滞留系统未栽种植物,对各指标的影响主要受控于介质土类型,运行性能较为稳定,同时各指标方差的差异性较小(表 2),从而表现出 B 组数据各指标权重较均衡。

将 GA 和 PSO 两种方法求得的最优投影系数分别代入投影寻踪模型中,可分别得到 A 组和 B 组不同实验对象的投影值,如图 4 所示。A、B 两组数据

不同评价方法的评价结果分别见表 3、4。

表 2 A、B 两组数据指标的方差

A 组	方差值	B 组	方差值
SS	0.90	NO <sub>3</sub> <sup>-</sup> -N	3.89
COD	24.27	TP	1.98
TN	12.69	COD	3.41
TP	3.23	HPC	4.73

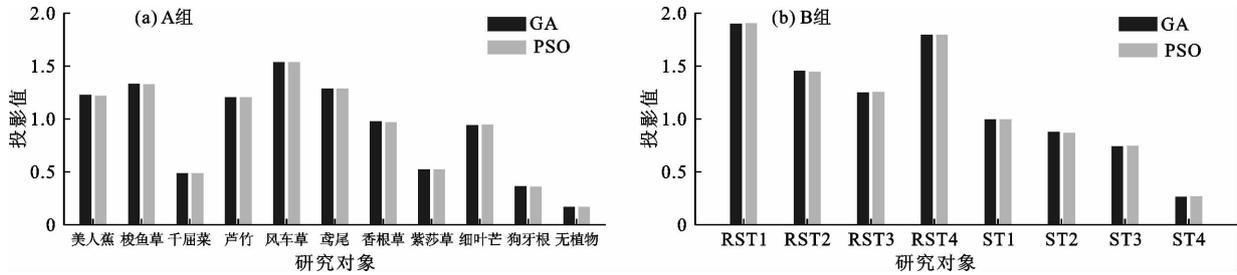


图 4 A、B 两组实验数据投影寻踪模型结果

表 3 A 组数据不同评价方法的评价结果

评价方法	评价排名
层次分析法	风车草 > 美人蕉 > 鸢尾 > 梭鱼草 > 香根草 > 千屈菜 > 芦竹 > 细叶芒 > 狗牙根 > 纸莎草 > 无植物
BP 神经网络法	风车草 > 美人蕉 > 鸢尾 > 梭鱼草 > 香根草 > 千屈菜 > 芦竹 > 狗牙根 > 纸莎草 > 细叶芒 > 无植物
本文研究	风车草 > 梭鱼草 > 鸢尾 > 美人蕉 > 芦竹 > 香根草 > 细叶芒 > 纸莎草 > 千屈菜 > 狗牙根 > 无植物

表 4 B 组数据不同评价方法的评价结果

评价方法	评价排名
PP-GA 法	RST2 > ST2 > RST4 > ST4 > RST3 > ST3 > RST1 > ST1
本文研究	RST2 > RST1 > RST3 > RST4 > ST2 > ST3 > ST4 > ST1

由图 4 可知,采用 GA 和 PSO 两种算法对生物滞留系统运行性能的多目标评价(RF-PP 模型)的优劣排序结果几乎一致。其中对于 A 组数据,研究结果表明风车草对污染物的综合去除性能最佳,而无植物的空白对照组的综合除污性能最差,与课题组前期采用层次分析法<sup>[35]</sup>和 BP 神经网络法<sup>[16]</sup>得到的最佳植物与最差植物(空白对照组)完全相同,但由于算法的差异,导致各指标权重不一致,从而造成其他植物的综合性能评价排名略有不同(表 3)。其中,差异较大的植物为其余方法排名第 2 的美人蕉与本文方法排名第 2 的梭鱼草。经过查看原始数据发现,梭鱼草、美人蕉对 TN 和 COD 两个权重较大的指标去除率较高,但梭鱼草对 COD 的去除率高于美人蕉,而对 TN 的去除率低于美人蕉,但本文算法给出的 COD 权重大于 TN,故得到梭鱼草综合排名大于美人蕉。除此之外,梭鱼草在 SS 与 TP 两个

指标去除率上也优于美人蕉。对于 B 组数据,GA 和 PSO 两种算法求解结果均认为介质土 RST2 的综合运行性能最佳,而介质土 ST1 的综合运行性能最差,这与课题组前期采用的传统投影寻踪模型<sup>[17]</sup>得到的最佳和最差介质土类型完全一致(表 4)。由于本研究对实验数据进行了特征选择,介质土 ST4、ST2 除污性能最高的除氮性能指标中正好剔除了 NH<sub>3</sub>-N 和 TN 两个指标,而 RST1 则恰好相反,导致 ST4、ST2 和 RST1 的综合评价排名与传统 PP 模型获得的结果有较大差异,其余研究对象的评价结果仅有微小差异。同时,NO<sub>3</sub><sup>-</sup>-N、NH<sub>3</sub>-N 和 TN 3 个指标之间相关性较强(图 2),传统 PP 模型未剔除这些显著相关的重复指标,容易造成多目标评价时的权重偏大。由此可见,本研究提出的方法更客观,结果更真实。

根据本研究算法得到的结果,植物与介质土单独实验时,可以采用风车草作为生物滞留系统的植物,采用 RST2 作为生物滞留系统的介质土;但由于植物与介质土可能存在耦合效应,因此,对于生物滞留系统中植物与介质土的最优组合仍需进一步通过实验进行验证。

## 4 结 论

(1) 针对生物滞留系统多目标评价问题的复杂性,提出了 RF-PP 模型,且通过对原始数据采用随机森林进行特征筛选降维后,在一定程度上提升了模型求解速度;但受实验数据样本数所限,RF-PP 模型尚不能体现出在避免维数灾难上的优势。同时利用鲁棒性高、求解速度快、参数不敏感的 GA 与 PSO 算法确定的权重指标,可有效避免传统求解方法参数难以确定、易陷入局部最优解等问题,进而得到更加合理的指标权重,实现方案优劣排序。

(2) 植物评价结果表明,风车草为生物滞留系统的最佳植物,且评价结果与层次分析模型和 BP 神经网络模型方法相似;介质土评价结果表明,RST2(9.8% 壤砂土 + 88.2% 细砂 + 2.0% 蛭石)为生物滞留系统的最佳介质土配置方案,且评价结果与传统投影寻踪法相似。本研究利用 GA 和 PSO 算法求解 RF-PP 模型具有一定的可行性和适用性,可为生物滞留系统多目标评价研究提供一定的参考和借鉴。

(3) 本研究得出的最佳植物与最佳介质土组合后可能存在耦合效应,未来可进一步通过植物与介质土组合实验研究,选取更佳的组合方式。

### 参考文献:

- [1] 陈 垚,李欣芮,郑 爽,等. 前期干旱天数对生物滞留系统除氮性能的影响 [J]. 环境科学,2021,42(1):263-273.
- [2] 李亚娇,赵凌志,李家科,等. 生物滞留系统对城市降雨径流病原微生物的去除效果及机理研究进展 [J]. 水资源与水工程学报,2020,31(1):14-22.
- [3] 穆 莹,王金丽,郑华清,等. 北方盐碱区域生物滞留设施植物优选研究 [J]. 环境科学与技术,2020,43(7):111-116.
- [4] 陈 垚,杨 威,王健斌,等. 雨水生物滞留设施中植被的设计与养护 [J]. 中国给水排水,2017,33(12):6-11.
- [5] 马千驰,余国先,钟鸿鹏. 一种增强的局部保持投影方法 [J]. 计算机工程与应用,2010,46(10):119-121.
- [6] 姜启源,谢金星,叶 俊. 数学模型(第四版) [M]. 北京:高等教育出版社,2010.
- [7] 白玉娟,殷国栋. 地下水水质评价方法与地下水研究进展 [J]. 水资源与水工程学报,2010,21(3):115-119+123.
- [8] 邢立文,魏新平. 基于 Elman 神经网络模型的西河水库水质评价 [J]. 人民长江,2020,51(7):64-70.
- [9] 胡明星,郭达志. 湖泊水质富营养化评价的模糊神经网络方法 [J]. 环境科学研究,1998,11(4):42-44.
- [10] 李海涛,邵泽东. 基于头脑风暴优化算法与 BP 神经网络的海水水质评价模型研究 [J]. 应用海洋学学报,2020,39(1):57-62.
- [11] 陆 丹,耿昭克,闵 敏,等. 基于区间型贝叶斯模型的湟水干流水质评价 [J]. 水利水运工程学报,2020(2):15-21.
- [12] 白雁翎,崔晨风,童山琳. 基于改进萤火虫算法的投影寻踪太湖水质综合评价 [J]. 节水灌溉,2017(5):72-76+79.
- [13] 巩奕成,张永祥,丁 飞,等. 基于萤火虫算法的投影寻踪地下水水质评价方法 [J]. 中国矿业大学学报,2015,44(3):566-572.
- [14] POETER E, ANDERSON D. Multimodel ranking and inference in ground water modeling [J]. Ground Water, 2005, 43(4): 597-605.
- [15] 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.
- [16] 余雪花,陈 垚,任萍萍,等. 生物滞留系统植物筛选与综合评价 [J]. 环境工程学报,2019,13(7):1634-1644.
- [17] 朱嘉运. 土壤特性对生物滞留系统运行性能的影响研究 [D]. 重庆:重庆交通大学,2021.
- [18] 孙傲涵,李建柱,冯 平. RF-SVR 降尺度模型在滦河流域的适用性分析 [J]. 水资源与水工程学报,2021,32(2):31-37.
- [19] 吴奉亮,霍 源,高佳南. 基于随机森林回归的煤矿瓦斯涌出量预测方法 [J]. 工矿自动化,2021,47(8):102-107.
- [20] 陈 娟,王献雨,罗玲玲,等. 缺失值填补效果:机器学习与统计学习的比较 [J]. 统计与决策,2020,36(17):28-32.
- [21] FRIEDMAN J H, TUKEY J W. A projection pursuit algorithm for exploratory data analysis [J]. IEEE Transactions on Computers, 1974, 19(4): 881-890.
- [22] 王 硕. 系统预测与综合评价方法 [M]. 合肥:合肥工业大学出版社,2006.
- [23] 潘思成,崔东文. 新型萤火虫算法-投影寻踪模型在区域人水和谐评价中的应用 [J]. 水资源与水工程学报,2020,31(1):124-130.
- [24] 李 琪,赵志怀. 基于投影寻踪与内梅罗指数组合模型的地下水水质评价 [J]. 水电能源科学,2019,37(11):70-73.
- [25] 肖专文,徐日庆,龚晓南. 求解复杂工程优化问题的一种实用方法 [J]. 水利学报,1999,30(2):25-29.
- [26] 史春天,曾艳阳,侯守明. 群体智能算法在图像分割中的应用综述 [J]. 计算机工程与应用,2021,57(8):36-47.
- [27] 由睿鹏. 计算机网络优化设计中遗传算法的原理及应用 [J]. 电子技术与软件工程,2020(20):14-15.
- [28] 刘 佳,王 杰. 无人水面艇避障路径规划算法综述 [J]. 计算机应用与软件,2020,37(8):1-10+20.
- [29] 刘卫林,刘丽娜. 抚河水环境质量的投影寻踪回归分析 [J]. 中国农村水利水电,2011(5):39-42+49.

(下转第 96 页)

- [13] STRUPCZEWSKI W G, SINGH V P, FELUCH W. Non-stationary approach to at-site flood frequency modeling I. Maximum likelihood estimation [J]. *Journal of Hydrology*, 2001, 248(1): 123 – 142.
- [14] RIGBY R A, STASINOPOULOS D M. Generalized additive models for location, scale and shape[J]. *Journal of the Royal Statistical Society: Series C (Applied statistics)*, 2005, 54(3): 507 – 554.
- [15] VILLARINI G, SMITH J A, SERINALDI F, et al. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin [J]. *Advances in Water Resources*, 2009, 32(8): 1255 – 1266.
- [16] 叶长青, 陈晓宏, 张家鸣, 等. 具有趋势变异的非一致性东江流域洪水序列频率计算研究[J]. *自然资源学报*, 2013, 28(12): 2105 – 2116.
- [17] CHEN Xiaohong, YE Changqing, ZHANG Jiaming, et al. Selection of an optimal distribution curve for non-stationary flood series[J]. *Atmosphere*, 2019, 10(1): 31.
- [18] 余涛, 谢平, 桑燕芳, 等. 基于 LWHM - LUCC 模型的澜沧江流域非一致性年径流频率计算与分析[J]. *武汉大学学报(工学版)*, 2021, 54(12): 1085 – 1093.
- [19] HECHT J S, LACOMBE G, ARIAS M E, et al. Hydro-power dams of the Mekong River basin: A review of their hydrological impacts [J]. *Journal of Hydrology*, 2019, 568: 285 – 300.
- [20] MANN H B. Nonparametric test against trend[J]. *Econometrica*, 1945, 13(3): 245 – 259.
- [21] 李海川, 王国庆, 郝振纯, 等. 澜沧江流域水文气象要素变化特征分析[J]. *水资源与水工程学报*, 2017, 28(4): 21 – 27 + 34.
- [22] LI Yungang, WANG Zhaoxia, ZHANG Yueyuan, et al. Drought variability at various timescales over Yunnan Province, China: 1961 – 2015 [J]. *Theoretical and Applied Climatology*, 2019, 138(1 – 2): 743 – 757.
- [23] HAN Zhongying, LONG Di, FANG Yu, et al. Impacts of climate change and human activities on the flow regime of the dammed Lancang River in Southwest China[J]. *Journal of Hydrology*, 2019, 570: 96 – 105.
- [24] 陈翔, 赵建世, 赵铜铁钢, 等. 发电调度对径流情势及生态系统的影响分析——以小湾、糯扎渡水电站为例[J]. *水力发电学报*, 2014, 33(4): 36 – 43.
- [25] LI Dongnan, LONG Di, ZHAO Jianshi, et al. Observed changes in flow regimes in the Mekong River basin[J]. *Journal of Hydrology*, 2017, 551: 217 – 232.
- [26] 李杨, 王婕, 唐雄朋, 等. 变化环境下澜沧江 - 湄公河流域径流演变及年内分配特征[J]. *水土保持研究*, 2021, 28(3): 141 – 148.
- [27] 丁凯熙, 张利平, 余敦先, 等. 全球升温 1.5°C 和 2.0°C 情景下澜沧江流域极端降水的变化特征[J]. *气候变化研究进展*, 2020, 16(4): 466 – 479.
- [28] WANG Wei, LU Hui, LEUNG L R, et al. Dam construction in Lancang - Mekong River Basin could mitigate future flood risk from warming-induced intensified rainfall [J]. *Geophysical Research Letters*, 2017, 44(20): 10378 – 10386.
- [29] YUN Xiaobo, TANG Qihong, LI Jiabo, et al. Can reservoir regulation mitigate future climate change induced hydrological extremes in the Lancang - Mekong River Basin? [J]. *Science of the Total Environment*, 2021, 785: 147322.
- [30] 侯时雨, 田富强, 陆颖, 等. 澜沧江 - 湄公河流域水库联合调度防洪作用[J]. *水科学进展*, 2021, 32(1): 68 – 78.

(上接第 90 页)

- [30] 郝宗睿, 李超, 任万龙, 等. 基于改进粒子群算法的喷水推进泵叶片优化设计 [J]. *排灌机械工程学报*, 2020, 38(6): 566 – 570.
- [31] 王玉鑫, 任帅. 基于遗传 - 粒子群混合优化算法的拆卸序列规划方法研究 [J]. *机电工程*, 2021, 38(3): 337 – 342.
- [32] 高飞. *MATLAB 智能算法超级学习手册* [M]. 北京: 人民邮电出版社, 2014.
- [33] YANG Xiaohua, MEI Ying, HE Jun, et al. Comprehensive assessment for removing multiple pollutants by plants in bioretention systems [J]. *Chinese Science Bulletin*, 2014, 59(13): 1446 – 1453.
- [34] GUO H, LIM F Y, ZHANG Y, et al. Soil column studies on the performance evaluation of engineered soil mixes for bioretention systems [J]. *Desalination and Water Treatment*, 2015, 54(13): 3661 – 3667.
- [35] 任萍萍. 山地城市雨水生物滞留系统植物筛选与除氮机理研究 [D]. 重庆: 重庆交通大学, 2018.