

文章编号: 2096-1618(2024)04-0416-06

# 基于 FRBPSO-RBF 神经网络的污水 BOD<sub>5</sub> 软测量方法

班慧琳, 李中志, 李斌勇, 王远

(成都信息工程大学网络空间安全学院, 四川 成都 610225)

**摘要:** 污水处理过程中污水 BOD<sub>5</sub> 难以实时准确测量, 故软测量方法逐渐被用于污水 BOD<sub>5</sub> 的预测, 其中 RBF 神经网络软测量方法应用广泛, 但存在训练过程易陷入局部极值等问题。为提高 RBF 神经网络的预测精度, 提出了基于适应度排名的粒子群算法 (fitness ranking based particle swarm optimization, FRBPSO), 根据适应度排名与迭代次数确定惯性权重的大小, 并根据粒子个体历史最优值的排名与迭代次数确定自我学习因子与社会学习因子的大小, 并将 FRBPSO 算法引入 RBF 神经网络的参数训练中。基于 13 个基准测试函数与其他 3 个粒子群优化算法对比, 实验结果显示 FRBPSO 算法的寻优能力相对较强。再将基于 FRBPSO 算法的 RBF 神经网络用于构建污水 BOD<sub>5</sub> 软测量模型, 仿真结果表明, 在测试数据中, FRBPSO-RBF 软测量模型的平均绝对误差比 PSO-RBF 软测量模型、DAIW-RBF 软测量模型、SCVPSO-RBF 软测量模型分别降低了 0.7178、0.2402、0.5851, 平均绝对百分比误差分别降低了 0.47%、0.15%、0.33%, 均方根误差分别降低了 0.0034、0.0015、0.0039。与其他 3 个基于 PSO 算法的 BOD<sub>5</sub> 软测量模型相比, FRBPSO-RBF 模型具有较高的 BOD<sub>5</sub> 预测精度。

**关键词:** RBF 神经网络; PSO 算法; 软测量模型; BOD<sub>5</sub> 软测量; 污水水质预测

**中图分类号:** TP391

**文献标志码:** A

**doi:** 10.16836/j.cnki.jcuit.2024.04.004

## 0 引言

在污水处理过程中, 生化需氧量 (biochemical oxygen demand, BOD) 是反映水中有机污染物含量的一个重要指标。BOD 是指在一定的条件下, 微生物分解水中可生化降解物质所进行的生物化学反应过程中消耗的溶解氧的数量。BOD 的测定需要较长的时间, 水样经过 5 d 的生化反应之后碳化阶段完成度能达到 60% ~ 70%<sup>[1]</sup>, 因此普遍规定使用 5 d 生化需氧量 BOD<sub>5</sub> 作为污水厂的监测指标。传统的 BOD<sub>5</sub> 测量方法耗时长, 难以及时检测, 且检测装置安装维护成本较高。因此, 寻找可靠的软测量方法是污水处理领域需要解决的问题。软测量技术的基本思想是利用易测变量通过构造模型计算出难以测量的待测变量, 本质上是一种建模的方法, 通过构造某种数学模型描述输入量、被控变量、扰动变量与待测量之间的函数关系<sup>[2]</sup>, 软测量技术已经被用于不同的行业<sup>[3]</sup>。近年来, 基于机器学习的污水水质参数软测量方法受到国内外学者的关注<sup>[4]</sup>。Qiao 等<sup>[5]</sup>提出一种基于混沌理论的 BOD 软测量方法。丁海旭等<sup>[6]</sup>提出一种自组织递归模糊神经网络的 BOD 软测量方法, 通过评估神经元的互信息关系和激活强度以增长或修剪规则层神经元。陈如清等<sup>[7]</sup>提出基于烟花算法的 RBF 神经网络 (radial basis function network,

RBF) 污水 COD 软测量方法。RBF 神经网络拓扑结构紧凑, 具有良好的非线性映射能力, 运算速度较快, 从而得到广泛研究与应用。在 RBF 神经网络参数优化方面, 不同的优化算法被应用于参数优化过程中。谢俊标等<sup>[8]</sup>利用引力搜索算法对传统人工鱼群算法进行改进, 用于优化 RBF 神经网络的权重、径向基函数中心和宽度。吴学梅等<sup>[9]</sup>采用鸟群算法优化传统麻雀搜索算法, 并用于优化 RBF-BP 神经网络。

1995 年, Kennedy 等<sup>[10]</sup>提出粒子群优化算法 (particle swarm optimization, PSO)。粒子群算法参数少、收敛快, 因此得到了广泛应用, 但粒子群算法容易陷入局部最优解。徐浩天等<sup>[11]</sup>以正态分布曲线作为惯性权重的衰减策略曲线, 并引入控制因子对粒子的位置进行改善。王生亮等<sup>[12]</sup>提出种群进化离散度的概念, 考虑到粒子之间进化差异, 赋予粒子不同的惯性权重。Zheng 等<sup>[13]</sup>采用适应度-距离平衡的方法构建一个替代全局最优粒子位置的向导粒子。Zhao 等<sup>[14]</sup>提出一种精英-普通协同粒子群优化算法, 精英粒子通过自我学习来保持种群的多样性, 普通粒子则朝着一个统一的目标飞行。Yu 等<sup>[15]</sup>提出一种基于新颖性的向导粒子, 两类具有高适应度和高新颖性的向导粒子分别保存在两个档案库中, 粒子随机选择两个向导粒子来更新其速度。

虽然上述改进使 PSO 算法的寻优能力获得一定程度的提升, 但 PSO 算法仍有很大的改进空间。在普通的时变参数调整策略中, 惯性权重、学习因子的变化

收稿日期: 2023-05-25

基金项目: 四川省科技计划资助项目 (2021JDR0046)

通信作者: 李中志. E-mail: lizz@cuit.edu.cn

与迭代次数有关,较少考虑到与粒子的适应度的关系。本文算法将粒子的适应度进行排序,结合粒子的适应度排名以及迭代次数,在每一次新的迭代中更新每个粒子的惯性权重和学习因子。RBF神经网络适用于非线性优化及预测建模等领域,但用于污水处理等复杂非线性过程的软测量建模时,存在参数难以确定,训练过程易陷入局部极值等问题。针对这类问题,本文提出一种基于FRBPSO的RBF神经网络参数训练方法,并建立污水BOD<sub>5</sub>浓度软测量模型,实验结果验证了该模型的有效性。

## 1 基础理论

### 1.1 RBF神经网络

RBF神经网络<sup>[16]</sup>是使用径向基函数作为激活函数的人工神经网络,RBF神经网络结构一般由输入层、隐含层及输出层构成,其结构如图1所示。

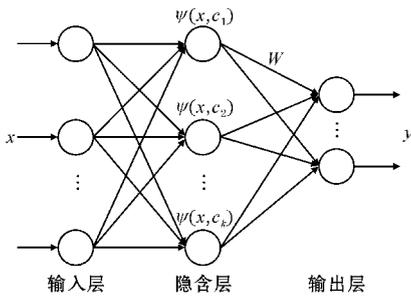


图1 RBF神经网络结构

第一层为输入层,输入向量  $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ , 通过输入层直接传递到隐含层中。

第二层为隐含层,常用高斯函数作为RBF神经网络的激活函数,隐含层第  $k$  个神经元的输出:

$$\psi_k(\mathbf{x}, c_k) = \exp\left(-\frac{\|\mathbf{x} - c_k\|^2}{2\sigma_k^2}\right) \quad (1)$$

式中,  $\mathbf{x}$  为输入向量,  $c_k$  表示第  $k$  个隐含层神经元的中心,  $\sigma$  表示第  $k$  个隐含层神经元的宽度。

第三层为输出层,输出层将隐含层的输出经过线性加权求和之后,得到输出结果。输出层的第  $i$  个输出:

$$y'_i = \sum_{j=1}^k W_{ij} \psi_j(\mathbf{x}, c_j) \quad (2)$$

式中,  $W_{ij}$  表示第  $j$  个隐含层节点到第  $i$  个输出层节点的权值,  $\psi_j(\mathbf{x}, c_j)$  表示第  $j$  个隐含层节点的输出。

RBF神经网络的参数中心值  $c$ 、宽度  $\sigma$ 、权值  $W$  决定神经网络模型的效果,因此需要对这些参数进行优化。RBF神经网络的参数个数与网络的拓扑结构有关,参数个数:

$$n = b(a+c+1) \quad (3)$$

式中,  $a$  表示输入层节点个数,  $b$  表示隐含层节点个数,  $c$  表示输出层节点个数。

### 1.2 基本粒子群算法

在粒子群优化算法中,每个个体被命名为一个“粒子”,每个粒子都被视为  $D$  维空间中的一个点,每个粒子都代表了一个问题的解决方案。基本粒子群算法的粒子速度与位置更新公式如下<sup>[10]</sup>:

$$\mathbf{V}_i(t+1) = w\mathbf{V}_i(t) + c_1 r_1 (\text{pbest}_i(t) - \mathbf{X}_i(t)) + c_2 r_2 (\text{gbest}(t) - \mathbf{X}_i(t)) \quad (4)$$

$$\mathbf{X}_i(t+1) = \mathbf{X}_i(t) + \mathbf{V}_i(t+1) \quad (5)$$

式中,  $c_1$  和  $c_2$  分别代表自我认知因子与社会认知因子,是两个正数,一般取为2。  $r_1$  和  $r_2$  是区间  $[0, 1]$  的两个随机数。  $\mathbf{V}_i(t)$  表示第  $i$  个粒子在第  $t$  次迭代的速度向量,  $\mathbf{X}_i(t)$  表示第  $i$  个粒子在第  $t$  次迭代的位置向量,  $\text{pbest}_i(t)$  表示第  $i$  个粒子在第  $t$  次迭代的历史最佳位置,对应的适应度值为  $\text{pbestScore}_i(t)$ ,  $\text{gbest}(t)$  表示整个粒子群体在第  $t$  次迭代的最佳位置,对应的适应度值为  $\text{gbestScore}(t)$ 。  $w$  表示惯性权重,如果取固定值,一般在区间  $[0.9, 1.2]$ 。惯性权重越大,则粒子的全局搜索能力越强,局部搜索能力越弱;惯性权重越小,粒子的局部搜索能力越强,全局搜索能力越弱<sup>[18]</sup>。

## 2 FRBPSO 算法

### 2.1 基于适应度排名与迭代次数的惯性权重

在粒子群中,优秀粒子的周边区域通常存在其他优秀的位置。为防止粒子离开优秀区域,此时粒子的惯性权重应设置为较小的值;适应度差的粒子通常离优秀位置较远,此时粒子的惯性权重应设置为较大的值。此外,粒子在种群迭代前期,侧重于开发与探索,因此惯性权重应设置为较大的值;在种群迭代后期,侧重于最优解的局部搜索,因此惯性权重应设置为较小的值。综合考虑以上原因,本文提出一种基于适应度排名与基于余弦迭代次数的惯性权重改进方法。在同一次迭代中,对适应度较优的粒子,设置较小的惯性权重,对适应度较差的粒子设置较大的惯性权重;而在不同的迭代中,相同排名的粒子的惯性权重随着迭代次数的增加非线性递减。

在每一轮新的迭代中,将每个粒子按照适应度从优到劣进行排序,即粒子适应度越优,粒子排名越靠前,再结合当前的迭代次数,计算惯性权重。粒子的惯性权重:

$$w = 0.4 \times \left[ \frac{r}{M} + \cos\left(\frac{\pi t}{2T}\right) \right] + 0.1 \quad (6)$$

式中,  $r$  为粒子的排名,  $M$  为粒子总数,  $T$  为最大迭代次数,  $t$  为当前迭代次数。

## 2.2 基于适应度排名与迭代次数的认知因子

$c_1$  是自我认知因子, 自我认知因子越大, 粒子群则更偏向全局探索  $c_2$  是社会认知因子, 社会认知因子越大, 粒子群则向当前整个粒子群的全局最优位置  $g_{best}$  靠拢。本文提出一种基于每个粒子的历史最优值的排名与迭代次数的认知因子, 在每一轮新的迭代中, 将所有粒子的历史最优值从优到劣进行排序,  $c_1$  和  $c_2$  如下:

$$c_1 = c_{\max} - \frac{\text{pbestSort} - (c_{\max} - c_{\min})t}{2M} \quad (7)$$

$$c_2 = c_{\min} + \frac{\text{pbestSort} - (c_{\max} - c_{\min})t}{2M} \quad (8)$$

式中,  $c_{\max}$  取值为 2.5,  $c_{\min}$  取值为 1,  $\text{pbestSort}$  表示该粒子历史最优值的排名。在同一轮迭代中, 粒子自身的  $\text{pbestSort}$  越小, 则粒子的自我认知因子越大, 社会认知因子越小; 在粒子的排名不变的情况下, 迭代次数越大, 粒子的自我认知因子越小, 社会认知因子越大。

## 2.3 FRBPSO 算法流程

本文提出的 FRBPSO 算法具体步骤如下:

步骤 1: 种群初始化, 设置最大迭代次数  $T$ , 粒子总数  $M$ , 随机初始化所有粒子的位置,  $t$  初始化为 0。

步骤 2: 根据粒子的位置计算适应度, 当前迭代次数  $t = t + 1$ 。

步骤 3: 分别计算出每个粒子个体的历史最佳适应度  $\text{pbestScore}$  与整个粒子群体的最佳适应度  $\text{gbestScore}$ , 并记录位置。

步骤 4: 将所有粒子当前的适应度从优到劣进行排序, 记录下每个粒子的适应度排名  $r$ , 并根据式(6)计算惯性权重。

步骤 5: 将所有粒子的  $\text{pbest}$  的适应度从小到大进行排序, 记录下每个粒子对应的排名, 根据式(7)、(8)分别计算自我认知因子和社会认知因子。

步骤 6: 根据式(4)、(5)更新粒子的速度与位置。

步骤 7: 若  $t$  未超过最大迭代次数  $T$ , 则返回步骤 2, 若超过最大迭代次数  $T$ , 则算法结束, 输出此时全局最佳位置  $\text{gbest}(T)$  与全局最佳适应度  $\text{gbestScore}(T)$ 。

## 2.4 算法寻优能力测试

### 2.4.1 基准测试函数

为检验本文 FRBPSO 算法的寻优能力, 使用 Matlab 2016 进行编程, 选取 13 个基准测试函数<sup>[18]</sup>, 用各 PSO 算法分别求出 13 个基准测试函数的最小值, 对算法的寻优能力进行测试。如表 1 所示, 已确定各个函

数自变量的范围, 其中  $F_1 \sim F_7$  为单峰测试函数,  $F_8 \sim F_{13}$  为多峰测试函数。

表 1 基准测试函数

测试函数	$x$ 范围
$F_1 = \sum_{i=1}^D x_i^2$	$[-100, 100]$
$F_2 = \sum_{i=1}^D  x_i  + \prod_{i=1}^D  x_i $	$[-10, 10]$
$F_3 = \sum_{i=1}^D (\sum_{j=1}^D x_j)^2$	$[-100, 100]$
$F_4 = \max_i \{  x_i , 1 \leq i \leq D \}$	$[-100, 100]$
$F_5 = \sum_{i=1}^{D-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	$[-30, 30]$
$F_6 = \sum_{i=1}^D ( x_i + 0.5 )^2$	$[-100, 100]$
$F_7 = \sum_{i=1}^D ix_i^4 + \text{random}[0, 1)$	$[-1.28, 1.28]$
$F_8 = -\sum_{i=1}^D [x_i \sin(\sqrt{ x_i })]$	$[-500, 500]$
$F_9 = \sum_{i=1}^D [x_i^2 - 10 \cos(2\pi x_i) + 10]$	$[-5.12, 5.12]$
$F_{10} = -20 \exp(-0.2 \sqrt{\frac{1}{30} \sum_{i=1}^D x_i^2}) - \exp(\frac{1}{30} \sum_{i=1}^D \cos 2\pi x_i^2) + 20 + e$	$[-32, 32]$
$F_{11} = \frac{1}{4000} \sum_{i=1}^D x_i^2 - \prod_{i=1}^D \cos(\frac{x_i}{\sqrt{i}}) + 1$	$[-600, 600]$
$F_{12} = \sum_{i=1}^D \mu(x_i, 10, 100, 4) + \frac{\pi}{30}  10 \sin^2(\pi y_1) \cdot \sum_{i=1}^{D-1} (y_i - 1)^2 \times [1 + 10 \sin^2(\pi y_{i+1})] + (y_D - 1)^2 $ $y_i = 1 + (x_i + 1) / 4$	$[-50, 50]$
$F_{13} = 0.1 \{ \sin^2(3\pi x_1) + \sum_{i=1}^{D-1} (x_i - 1)^2 [1 + \sin^2(3\pi x_{i+1})] + (x_D - 1)^2 [1 + \sin^2(2\pi x_D)] \} + \sum_{i=1}^D \mu(x_i, 5, 100, 4)$ $\mu(x_i, a, k, m) = \begin{cases} k(x_i - a)^m, & x_i > a \\ 0, & -a \leq x_i \leq a \\ k(-x_i - a)^m, & x_i < -a \end{cases}$	$[-50, 50]$

### 2.4.2 算法效果对比

选取了基本粒子群算法 PSO<sup>[10]</sup>、非线性动态自适应惯性权重 PSO 算法 (DAIW)<sup>[12]</sup>、总结性自适应变异的粒子群算法 (SCVPSO)<sup>[19]</sup> 与本文 FRBPSO 算法做对比实验。统一设置粒子总数  $M$  为 200, 最大迭代次数  $T$  为 2000, 粒子的维度  $D$  为 40。分别求 13 个基准测试函数的最小值, 实验结果越小表示算法效果越优良。为降低实验偶然性, 各个粒子群算法在每个基准测试函数上独立运行 30 次, 并计算平均值, 实验结果如表 2 所示。从实验结果知, FRBPSO 算法的寻优能力在 8 个测试函数中优于其他 3 种对比算法, 在 8 个测试函数中能找到最小值, 在 4 个测试函数中能找到次小值, 总体寻优能力较好。

表 2 各算法在 13 个测试函数上的实验结果

测试函数	PSO	DAIW	SCVPSO	FRBPSO
$F_1$	2.40E+02	7.33E-09	2.31E-35	1.33E-33
$F_2$	6.20E+00	8.12E-11	5.93E-21	2.18E-12
$F_3$	1.86E+03	1.09E+02	7.32E+00	9.75E-05
$F_4$	8.43E+00	1.26E+00	5.97E-01	7.10E-02
$F_5$	1.06E+04	5.23E+01	5.69E+01	3.82E+01
$F_6$	2.22E+02	1.43E-09	1.82E-32	2.99E-27
$F_7$	6.57E-02	8.92E-03	6.23E-03	2.91E-03
$F_8$	-8.40E+03	-9.77E+03	-9.51E+03	-9.89E+03
$F_9$	1.43E+02	5.54E+01	5.23E+01	3.76E+01
$F_{10}$	4.68E+00	3.69E-12	1.43E-14	4.74E-14
$F_{11}$	2.53E+00	5.17E-03	6.52E-03	2.23E-02
$F_{12}$	2.42E+00	9.51E-12	3.89E-03	1.70E-23
$F_{13}$	6.86E+00	1.21E-03	2.19E-03	4.78E-30

### 3 基于 FRBPSO 的 RBF 神经网络的 BOD<sub>5</sub> 软测量

#### 3.1 FRBPSO 优化 RBF 神经网络参数

本文提出基于 RBF 神经网络污水 BOD<sub>5</sub> 软测量模型,RBF 神经网络具有很强的非线性逼近能力。将

RBF 网络用于软测量建模,就是利用 RBF 神经网络构造,采用常规传感器方便测量的辅助变量与难以直接测量的待测变量之间的网络模型,实现对待测变量的预测。利用 FRBPSO 算法优化 RBF 神经网络参数,其中包括隐含层神经元中心值  $c$ 、宽度  $\sigma$ ,隐含层到输出层的权值  $W$ 。将粒子的位置作为 RBF 神经网络的参数,即每个粒子的位置代表一组 RBF 神经网络的参数值。本文将 RBF 神经网络误差评价指标平均绝对误差(mean absolute error, MAE) 作为粒子的适应度函数,适应度越小,表示当前粒子所对应的网络参数越优。MAE 的计算公式如下:

$$MAE = \frac{1}{N} \sum_{a=1}^N |y_a - y'_a| \quad (9)$$

式中,  $y_a$  为实际值,  $y'_a$  为预测值,  $N$  为数据样本个数。

在软测量模型中,粒子适应度的计算公式:

$$fitness = \frac{1}{N} \sum_{a=1}^N \left| y_a - \sum_{j=1}^k W_j \exp\left(-\frac{\|\mathbf{x}_a - c_j\|^2}{2\sigma_j^2}\right) \right| \quad (10)$$

式中,  $\mathbf{x}$  为输入向量,  $c_j$  表示第  $k$  个隐含层神经元的中心,  $\sigma_j$  表示第  $j$  个隐含层神经元的宽度。  $W_j$  表示第  $j$  个隐含层节点到输出层节点的权值。

FRBPSO 优化 RBF 神经网络参数的流程如图 2 所示。

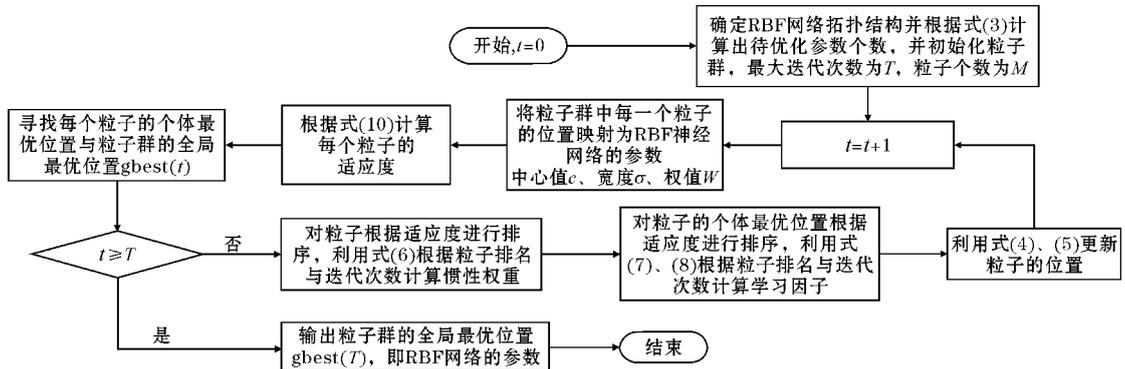


图 2 FRBPSO 算法优化 RBF 神经网络参数的流程

#### 3.2 样本数据处理

为验证基于 FRBPSO 算法的 RBF 神经网络 BOD<sub>5</sub> 软测量方法的有效性,本文使用来自国际水协会的污水处理厂进水数据,其中 400 组数据用于训练神经网络,272 组数据用于测试软测量模型的效果。本文选取污水的 7 个水质参数作为网络的输入变量,包括可溶性易生物降解有机物 SS、活性异养生物量 XBH、不可溶性可生物降解有机物 XS、不可溶性不可生物降解有机物 XI、氨氮 SNH、可溶性可生物降解有机氮 SND、不可溶性可生物降解有机氮 XND, BOD<sub>5</sub> 为输出变量。由于不同的水质指标具有不同的量纲,为消除不同量纲的影响,将水质参数数据进行归一化处理,映射到区间[0,1]。

#### 3.3 仿真实验

为验证 FRBPSO-RBF 算法对 BOD<sub>5</sub> 的预测精度,本文算法将与 PSO-RBF 神经网络、DAIW-RBF 神经网络、SCVPSO-RBF 神经网络进行对比。实验使用 Matlab 2016 进行编程,首先构建 RBF 神经网络污水 BOD<sub>5</sub> 浓度软测量模型,其中 RBF 神经网络隐含层节点个数的选取与网络的输入输出维度相关,根据文献[7]的设置方法,设网络的输入层节点数为  $a$ ,将隐含层节点个数设置为  $b=2a+1$ ,为确保对比实验的公平性,各模型的网络拓扑结构相同。由于本实验有 7 个输入的水质参数,故网络拓扑结构统一设置为 7-15-1,根据式(3)计算出待优化网络参数个数为 135 个。

统一设置 4 个污水 BOD<sub>5</sub> 软测量模型的粒子个数  $M$  为 100,最大迭代次数  $T$  为 100,由于每个粒子的位置代表 RBF 神经网络的一组参数,故每个粒子的维度与待优化网络参数个数相同,为 135,粒子的适应度为 RBF 神经网络在训练数据中的平均绝对误差 MAE。为降低实验偶然性,各个粒子群算法优化的 RBF 神经网络软测量模型独立运行 50 次,并求出平均值。

统一设置 4 个污水 BOD<sub>5</sub> 软测量模型的粒子个数  $M$  为 100,最大迭代次数  $T$  为 100,由于每个粒子的位置代表 RBF 神经网络的一组参数,故每个粒子的维度与待优化网络参数个数相同,为 135,粒子的适应度为 RBF 神经网络在训练数据中的平均绝对误差 MAE。为降低实验偶然性,各个粒子群算法优化的 RBF 神经网络软测量模型独立运行 50 次,并求出平均值。

4个污水 BOD<sub>5</sub> 软测量模型分别独立运行 50 次, 50 次的平均适应度变化情况如图 3 所示。

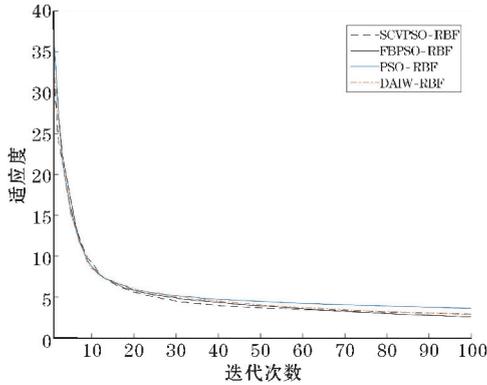


图3 适应度的变化情况

从图 3 可以看出, 本文 FRBPSO-RBF 算法的最终适应度略优于其他 3 个对比算法。本文算法在迭代前期侧重于开发与探索, 所以收敛速度并不是最快的; 在迭代后期, 优良粒子的局部寻优对寻找更优秀的解发挥了重要作用。因此, 与其他 3 个算法相比, 本文算法的最终适应度最小, 求解精度较高。

本文 FRBPSO-RBF 算法、其他 3 个对比算法对 272 个测试样本的 BOD<sub>5</sub> 浓度预测值与 BOD<sub>5</sub> 实际值如图 4 所示。

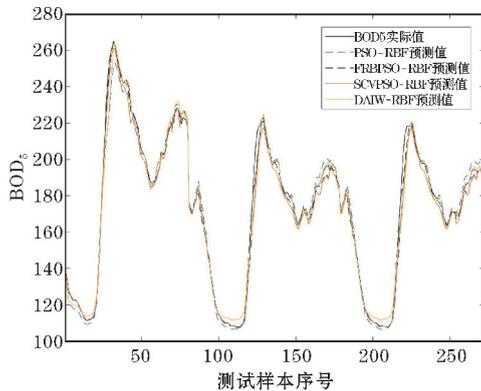


图4 4个软测量模型的预测结果与真实结果

从图 4 的 BOD<sub>5</sub> 预测结果可以看出, 本文 FRBPSO-RBF 算法对测试样本的 BOD<sub>5</sub> 浓度测量精度略优于其他 3 个对比算法, 本文所构建的软测量模型具有良好的 BOD<sub>5</sub> 预测效果。

为降低偶然性, 更加公平地比较软测量模型的效果, 计算各个粒子群算法优化的 RBF 神经网络独立运行 50 次的平均测试误差, 其中包括平均绝对误差、平均绝对百分比误差 (mean absolute percentage error, MAPE)、均方根误差 (root mean square error, RMSE)。测试误差越小, 则表示软测量模型预测效果越好。MAPE 与 RMSE 的公式如下:

$$\text{MAPE} = \frac{100\%}{N} \sum_{a=1}^N \left| \frac{y_a - y'_a}{y_a} \right| \quad (11)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{a=1}^N (y_a - y'_a)^2} \quad (12)$$

式中,  $y_a$  为实际值,  $y'_a$  为预测值。

各算法独立运行 50 次的平均训练与测试误差如表 3 所示。

表3 4个 BOD<sub>5</sub> 软测量模型的误差

算法	训练 MAE	测试 MAE	测试 MAPE/%	测试 RMSE
PSO-RBF	3.3237	3.3823	1.93	0.0255
DAIW-RBF	2.7853	2.9047	1.61	0.0236
SCVPSO-RBF	2.9475	3.2496	1.79	0.0260
FRBPSO-RBF	2.4422	2.6645	1.46	0.0221

实验表明, 与另外 3 种 RBF 神经网络 BOD<sub>5</sub> 软测量模型相比, 本文基于 FRBPSO-RBF 神经网络 BOD<sub>5</sub> 软测量模型在测试数据中具有更小的 MAE、MAPE 与 RMSE, 再次验证了本文算法的有效性, 测试样本的 BOD<sub>5</sub> 浓度实际值与软测量模型输出值之间的误差较小, 可以较准确地预测 BOD<sub>5</sub> 浓度。在测试数据中, FRBPSO-RBF 软测量模型的 MAE 分别比 PSO-RBF 软测量模型、DAIW-RBF 软测量模型、SCVPSO-RBF 软测量模型降低了 0.7178、0.2402、0.5851, MAPE 分别降低了 0.47%、0.15%、0.33%, RMSE 分别降低了 0.0034、0.0015、0.0039。

## 4 结束语

BOD<sub>5</sub> 是评价污水水质的一项重要参数, 传统的 BOD<sub>5</sub> 测量耗时较长, 为实现 BOD<sub>5</sub> 的实时准确测量, 软测量方法开始逐渐应用。RBF 神经网络软测量方法应用广泛, 但参数训练过程易陷入局部极值。为优化 RBF 神经网络的参数, 本文提出 FRBPSO 算法, 并将该算法用于 RBF 神经网络的参数训练中, 提出一种基于适应度排名与迭代次数的惯性权重与学习因子更新策略, 并根据粒子个体历史最佳位置的适应度排名确定自我学习因子  $c_1$  与社会学习因子  $c_2$ 。上述策略有效地平衡了粒子群算法在迭代过程中全局搜索与局部搜索的矛盾。

实验表明, 与其他 3 种算法相比, FRBPSO 算法的寻优效果较好。将该算法引入 RBF 神经网络 BOD<sub>5</sub> 软测量模型中, 实验表明该软测量模型测量精度较高, MAE、MAPE 与 RMSE 分别为 2.6645、1.46% 和 0.0221, 均低于文中的其他 3 个对比算法。相对于其他 3 个基于 PSO 算法的 BOD<sub>5</sub> 软测量模型, FRBPSO-RBF 模型具有较高的 BOD<sub>5</sub> 预测精度, 为污水处理等复杂过程中关键变量的测量提供了一种可行的方法。后续还可以新增相关的水质参数, 作为 BOD<sub>5</sub> 软测量模型的辅助变量, 为软测量模型提供更多的数据支持。

## 参考文献:

- [1] 张钧, 杨文武. 水和废水中五日生化需氧量测定全程质量控制[J]. 环境科学与管理, 2011, 36

- (7):139-145.
- [2] 黄道平,刘乙奇,李艳.软测量在污水处理过程中的研究与应用[J].化工学报,2011,62(1):1-9.
- [3] Perera Yasith S,Ratnaweera D A A C,Dasanayaka Chamila H, et al. The role of artificial intelligence-driven soft sensors in advanced sustainable process industries: A critical review [J/OL]. <https://doi.org/10.1016/j.engappai.2023.105988>,2023-03-21.
- [4] Wang Gongming, Jia Qingshan, Zhou Mengchu, et al. Artificial neural networks for water quality soft-sensing in wastewater treatment: a review[J]. Artificial Intelligence Review,2021,55(1):565-587.
- [5] Qiao Junfei, Hu Zhiqiang, Li Wenjing. Soft Measurement Modeling Based on Chaos Theory for Biochemical Oxygen Demand (BOD) [J]. Water, 2016,8(12):581-581.
- [6] 丁海旭,李文静,叶旭东,等.基于自组织递归模糊神经网络的BOD软测量[J].计算机与应用化学,2019,36(4):331-336.
- [7] 陈如清,于志恒.基于TentFWA-GD的RBF神经网络COD在线软测量方法[J].电子测量与仪器学报,2022,36(3):53-60.
- [8] 谢俊标,江峰,杜军威,等.基于改进人工鱼群算法与RBF神经网络的股票价格预测[J].计算机工程与科学,2022,44(11):2080-2090.
- [9] 吴学梅,牟莉.基于改进RBF-BP神经网络的预测方法研究[J].国外电子测量技术,2022,41(2):105-110.
- [10] Kennedy J,Eberhart R. Particle Swarm Optimization [C]. Proceedings of ICNN'95-International Conference on Neural Networks. IEEE, 1995: 1942-1948.
- [11] 徐浩天,季伟东,孙小晴,等.基于正态分布衰减惯性权重的粒子群优化算法[J].深圳大学学报:理工版,2020,37(2):208-213.
- [12] 王生亮,刘根友.一种非线性动态自适应惯性权重PSO算法[J].计算机仿真,2021,38(4):249-253.
- [13] Zheng Kaitong, Yuan Xianfeng, Xu Qingyang, et al. Hybrid particle swarm optimizer with fitness-distance balance and individual self-exploitation strategies for numerical optimization problems [J]. Information Sciences,2022,608:424-452.
- [14] Zhao Shicun, Wang Da. Elite-ordinary synergistic particle swarm optimization [J]. Information Sciences,2022,609:1567-1587.
- [15] Yu Fei, Tong Lei, Xia Xuewen. Adjustable driving force based particle swarm optimization algorithm [J]. Information Sciences,2022,609:60-78.
- [16] Moody J, Darken C. Learning with localized receptive fields [C]. Proceedings of the connectionist models summer school. Yale Univ, 1988:133-143.
- [17] SHI Y, EBERHART RC. Parameters selections in particle swarm optimization [C]. Evolutionary Programming VII, 1998:591-600.
- [18] Yao Xin, Liu Yong, Lin Guangming. Evolutionary programming made faster. [J]. IEEE Trans. Evolutionary Computation, 1999, 3(2):82-102.
- [19] 陈博文,邹海.总结性自适应变异的粒子群算法[J].计算机工程与应用,2022,58(8):67-75.

## Soft Sensing Method of BOD<sub>5</sub> in Sewage based on FRBPSO-RBF Neural Network

BAN Huilin, LI Zhongzhi, LI Binyong, WANG Yuan

(College of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, China)

**Abstract:** During the wastewater treatment process, it is difficult to accurately measure the BOD<sub>5</sub> of wastewater in real time. Therefore, soft measurement methods are gradually being used to predict BOD<sub>5</sub> in wastewater. Among them, the RBF neural network soft measurement method is widely used, but there are problems such as the training process being easily trapped in local extremes. In order to improve the prediction accuracy of the RBF neural network, a fitness ranking based particle swarm optimization algorithm (FRBPSO) is proposed. The size of the inertia weight is determined according to the fitness ranking and the number of iterations, and the size of the self-learning factor and social learning factor are determined according to the ranking of the individual's historical optimal value and the number of iterations. Then, the FRBPSO algorithm is introduced into the parameter training of the RBF neural network. Based on 13 benchmark test functions and comparison with other 3 particle swarm optimization algorithms, experimental results show that the FRBPSO algorithm has relatively strong optimization ability. Then, the RBF neural network based on FRBPSO algorithm is used to construct a soft measurement model for wastewater BOD<sub>5</sub>. Simulation results show that in test data, compared with PSO-RBF soft measurement model, DAIW-RBF soft measurement model and SCVPSO-RBF soft measurement model, the average absolute error of FRBPSO-RBF soft measurement model decreased by 0.7178, 0.2402 and 0.5851 respectively; The average absolute percentage error decreased by 0.47%, 0.15%, and 0.33% respectively; The root mean square error decreased by 0.0034, 0.0015 and 0.0039 respectively. Compared with other three BOD<sub>5</sub> soft measurement models based on PSO algorithm, FRBPSO-RBF model has higher BOD<sub>5</sub> prediction accuracy.

**Keywords:** radial basis function neural network; particle swarm optimization; soft sensing model; 5-day BOD biochemical oxygen demand soft sensing; sewage quality prediction