

Research
Environmental Engineering—Article

基于可控参数的前馈神经网络出水总氮预测模型研究

赵子豪^a, 王子昊^a, 袁家洛^a, 马骏^a, 何哲灵^a, 徐一兰^d, 沈晓佳^e, 朱亮^{a,b,c,*}^a Institution of Environment Pollution Control and Treatment, Department of Environmental Engineering, Zhejiang University, Hangzhou 310058, China^b Zhejiang Province Key Laboratory for Water Pollution Control and Environmental Safety, Hangzhou 310058, China^c Zhejiang Provincial Engineering Laboratory of Water Pollution Control, Hangzhou 310058, China^d Haining Water Investment Group Co., Ltd, Haining 314400, China^e Haining Capital Water Co., Ltd, Haining 31440, China

ARTICLE INFO

Article history:

Received 12 April 2020

Revised 8 May 2020

Accepted 15 July 2020

Available online 1 December 2020

关键词

前馈神经网络 (FFNN)

算法

可控参数

序批式活性污泥工艺

总氮

摘要

当前,出水总氮(TN)稳定达标是我国城镇污水处理厂面临的关键问题。本文基于我国城镇污水处理厂主流工艺——序批式活性污泥工艺(SBR),构建了一种基于可控参数的前馈神经网络(FFNN)出水总氮预测模型。与已有预测模型相比,本模型具备以下两个特点:①采用可控参数(表面气速与缺氧段时长)代替溶解氧作为模型主要输入参数,明显提高模型可用可控性;②采用算法优化的前馈神经网络构建模型,显著提高模型预测精准度。研究表明:量化共轭梯度法优化的FFNN模型预测精准,其拟合相关系数(R)明显高于其他相关模型;优化后的FFNN模型可根据进水与关键控制参数实现出水TN精准预测,有望实现城镇污水处理厂总氮稳定去除、系统节能降耗,助力国家“3060”碳目标实现。

© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. 引言

近年来我国工业化城镇化进程不断加快,污水排放量攀升至每年669.6亿t以上[1]。为削弱污水排放所带来的环境不利影响,我国陆续采取新建或改扩建污水处理厂、修订污水排放标准等一系列措施。截至2018年年底,我国共有5370座污水处理厂处于正常运行状态,日处理能力达 $2.01 \times 10^8 \text{ m}^3$,总用电量达 $1.972 \times 10^{10} \text{ kW} \cdot \text{h}^{-1}$ [2]。据调查,目前我国多数城镇污水处理厂存在出水水质难以稳定达标、运行药耗能耗偏高、自动化控制程度较低等问题,亟需科技支撑予以解决。

序批式活性污泥工艺(SBR)因其工艺简便、操作方式灵活、抗负荷性能优异等特点,已成为我国应用较为广泛的污水处理工艺之一[3]。然而,SBR工艺对系统操控自动化水平要求较高,极易产生出水水质指标(尤其是总氮)波动大、运行能耗过高等问题[4]。而在SBR工艺中,脱氮是十分复杂的,包括好氧条件下的硝化反应及缺氧条件下的反硝化反应[5]。已有研究表明,溶解氧(DO)控制是实现污水处理系统总氮(TN)有效去除的重要因素,充足的DO可保证有机物降解及硝化反应顺利进行,但过量的DO则会导致污水处理能耗过高、反硝化效率降低、污泥性能恶化[6]。同时,根

* Corresponding author.

E-mail address: felix79cn@zju.edu.cn (L. Zhu).

据现有工程运行情况来看, 污水处理系统DO难以实现精准控制, 主要与DO监测设备准确度低、响应慢, 传统拟合模型和控制理论相对落后有关[7,8]。在大多数污水处理厂, 操作人员往往根据经验调整表面气速、缺氧段时长等控制参数, 从而实现体系DO浓度的相应调控, 这样的操作方式也比DO精准控制更为简便。

鉴于污水生物处理系统参数众多、过程复杂, 传统的生化模型难以拟合其高维的非线性数据, 进而难以应用到实际污水生物处理工艺中[9,10]。而人工神经网络(ANN)作为一种自学习算法模型, 理论上能够拟合任意非线性函数, 因而其可作为预测污水生物处理这种复杂非线性系统的有效工具[11-13]。从神经元拓扑结构来看, ANN可分为反馈神经网络和前馈神经网络(FFNN), 其中, FFNN模型理论上可以任意精度逼近任意连续函数, 具有较强的分类和模式识别能力[14-19]。为进一步提高FFNN的预测效率与能力, 研究者陆续提出了Levenberg-Marquardt(L-M)、贝叶斯正则化(BR)、量化共轭梯度(SCG)、动量和Nesterov加速梯度等多种优化算法, 显著提高FFNN模型的预测能力和效率[20-22]。

本文以序批式活性污泥工艺为研究平台, 构建了一种基于可控参数的前馈神经网络出水总氮预测模型。与已有预测模型相比, 本模型具备以下两个特点: ①采用可控参数(表面气速与缺氧段时长)代替溶解氧作为模型主要输入参数, 明显提高模型可用可控性; ②采用算法优化的FFNN构建模型, 显著提高模型预测精准度。本文的研究目的主要在于评估优化后FFNN模型的SBR工艺出水TN预测准确性, 明确合理地运行参数控制策略以实现污水处理厂污染物高效去除与系统节能降耗。

2. 材料与方 法

2.1. SBR 工艺长期仿真实验

本研究设置两组平行运行的SBR反应器(R1、R2), 开展持续两个月的长期仿真实验。反应器活性污泥取自浙江省某城镇污水处理厂, 进水模拟该污水处理厂的 实际进水。反应器结构及模拟废水组成如附录A中的图S1、表S1所示。

参考该污水处理厂SBR工艺的实际运行模式, 每组反应器的运行周期设为4 h, 包括进水(5 min)、缺氧与曝气期(210 min)、沉降期(5 min)、出水(5 min)、闲置(15 min), 体积交换比设为50%。配置进水时,

随机控制主要水质指标浓度为预设值的75%~125%, 模拟实际进水水质波动情况。分别设表面气速与缺氧段时长作为控制变量, 其设计值如表1所示。每个周期分别对进出水进行取样, 采用标准方法分析进出水的总氮(TN)、氨氮($\text{NH}_4^+\text{-N}$)、化学需氧量(COD)和总磷(TP)浓度[23]。经两个月的仿真运行后, 从控制变量的16种组合中获取124组数据供后续建模使用。

为拟合更复杂的运行状况, 本研究引入缺氧段后进水占比并设置范围更广的控制变量后开展扩展实验, 共收集91组数据, 其设计值如表2所示。同时为仿真实际运行中的极端情况, 本研究分别设置表面气速为 $3.6 \text{ cm}\cdot\text{s}^{-1}$ 、 $4.8 \text{ cm}\cdot\text{s}^{-1}$, 缺氧段时长为0、150 min, 并从其组合中收集11组数据。

2.2. FFNN 建模

基础FFNN模型及其优化算法均在Matlab R2016a中搭建并运行。建模前, 将长期仿真得到的数据集归一化至0.001~0.999, 以消除数据间不同量纲的影响。FFNN模型在构建过程中根据模型预测值与实际值的误差不断优化权值网络[式(1)] [24], 其网络结构包括输入层、隐藏层及输出层三部分, 其中进水水质(COD、TN、 $\text{NH}_4^+\text{-N}$ 、TP)与运行参数(表面气速、缺氧段时长)6个变量作为模型的输入变量, 出水TN则作为模型的输出变量, 建模示意图如图1所示。由于实验数据维度较低, 设置FFNN模型隐藏层数为1层以缩短运算时长, 提高效率, 同时防止过拟合现象发生。而模型隐藏层最佳结点数则参考经验方程进行计算[式(2)]。

$$a = Wd + b \quad (1)$$

表1 SBR工艺控制变量的设计值

Level	Superficial gas velocity ($\text{cm}\cdot\text{s}^{-1}$)	Anoxic time (min)
1	0.5	0
2	1.0	20
3	1.5	40
4	2.0	60

表2 扩展实验中引入缺氧段后控制变量和进水比的设计值

Level	Superficial gas velocity ($\text{cm}\cdot\text{s}^{-1}$)	Anoxic time (min)	Influent ratio after anoxic period
1	0.4	30	0
2	1.2	60	0.25
3	2.0	90	0.5
4	2.8	120	0.75

式中, a 为输出变量; W 为权值矩阵; d 为输入变量; b 为偏置。

$$h = \sqrt{i + o} + c \quad (2)$$

式中, h 为隐藏层结点数; i 为输入层结点数; o 为输出层结点数; c 为常数, 一般选在1~10之间。

2.3. FFNN 模型优化

基础FFNN模型参照传统的反向传播(BP)神经网络进行构建。在训练阶段, BP神经网络根据预测值与实际值误差的反向传播不断调整其权重网络, 使误差最小化。梯度下降法则是模型调整权重的最常用算法, 即沿梯度下降的方向调整权重并最小化误差[25]。但在实际训练过程中, 梯度下降法在迭代求解时容易最小化误差至局部极小值而不是全局最小值, 因而降低了模型学习效率与预测准确度[26]。为提高模型的学习效率与预测准确度, 本研究采用三种优化算法(L-M、BR以及SCG)对FFNN模型进行优化, 对比三种优化算法得到最适合SBR工艺的优化FFNN模型。同时, 为进一步提高模型的泛化能力, 采用更复杂的数据集(即包括扩展实验数据集)训练模型, 缺氧段后进水比将代替进水TP作为模型的输入参数。

2.3.1. L-M 算法

L-M算法结合高斯-牛顿算法以及梯度下降法的优点, 能有效避免局部最小值并提高全局最小值的收敛速

度, 其算法如下。

若将BP神经网络中各层间的权值用向量 W 表示, 其误差平方和(E)为

$$E = \frac{1}{2} \sum \sum (t_{nj} - O_{nj})^2 = \frac{1}{2} \sum \varepsilon_n^2 = \frac{1}{2} \|\varepsilon\|^2 \quad (3)$$

式中, n 为样本编号; t_{nj} 为样本 n 在输出层 j 结点的期望输出; O_{nj} 为该结点实际输出; ε 是以 ε_n 为元素的向量。

在式(4)中, k 代表迭代次数(即权重更新的次数)。在计算 W^{k+1} 的过程中, 如果移动量 $W^{k+1} - W^k$ 很小, 则可将 ε 展开成一阶泰勒级数:

$$\varepsilon W^{k+1} = \varepsilon W^{k+1} + Z(W^{k+1} - W^k) \quad (4)$$

式中, Z 是 ε 的雅可比矩阵, Z 的元素为

$$Z_{nj} = \frac{\partial \varepsilon_n}{\partial W_j} \quad (5)$$

由此, 误差函数可改为

$$E = \frac{1}{2} \left\| \varepsilon W^k + Z(W^{k+1} - W^k) \right\|^2 \quad (6)$$

要使误差函数最小, 可对 W^{k+1} 进行求导, 得到如下高斯-牛顿迭代公式:

$$W^{k+1} = W^k - (Z^T Z)^{-1} Z^T \varepsilon W^k \quad (7)$$

式中, T代表转置。

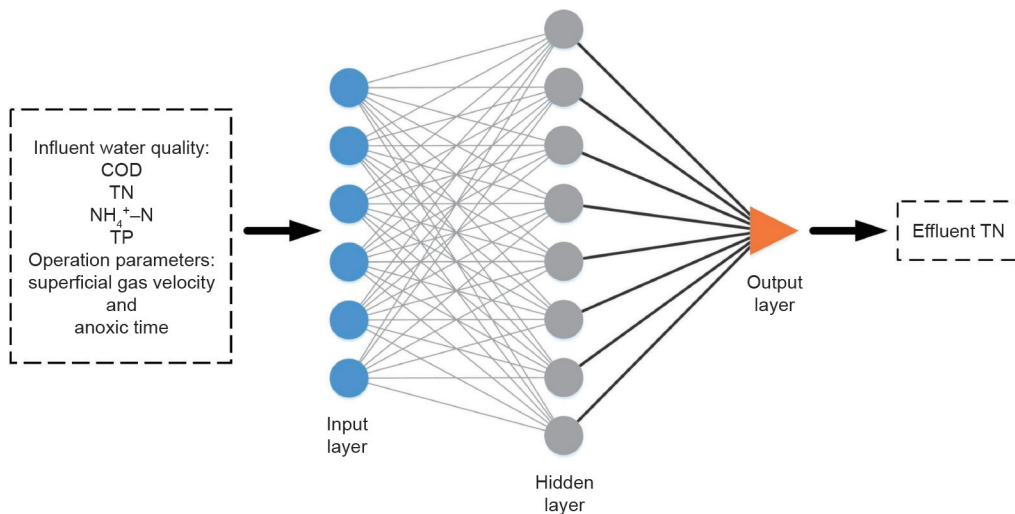


图1. 出水总氮预测FFNN模型的网络结构。

为克服高斯-牛顿法中经常出现的雅可比矩阵奇异现象, 把误差函数改写为

$$E = \frac{1}{2} \left\| \boldsymbol{\varepsilon} \mathbf{W}^k + \mathbf{Z} (\mathbf{W}^{k+1} - \mathbf{W}^k) \right\|^2 + \lambda \left\| \mathbf{W}^{k+1} - \mathbf{W}^k \right\|^2 \quad (8)$$

式中, λ 为阻尼系数。

此时对 E 求导, 便可得到基于高斯-牛顿法的L-M迭代公式:

$$\mathbf{W}^{k+1} = \mathbf{W}^k - \left(\mathbf{Z}^T \mathbf{Z} + \lambda \mathbf{I} \right)^{-1} \mathbf{Z}^T \boldsymbol{\varepsilon} \mathbf{W}^k \quad (9)$$

式中, \mathbf{I} 为单位矩阵, 为迭代变量。在迭代过程中搜索方向与训练步长受 λ 变化影响。在计算初始阶段 λ 可取较大值, 这时 $\mathbf{Z}^T \mathbf{Z}$ 与 $\lambda \mathbf{I}$ 相比是可以忽略的, 于是上式可写为

$$\mathbf{W}^{k+1} = \mathbf{W}^k - \frac{1}{\lambda} \mathbf{g} \quad (10)$$

式中, \mathbf{g} 为梯度。若令 λ 趋于零, 则该表达式可变为高斯-牛顿迭代公式。

2.3.2. BR 算法

BR算法指贝叶斯方法正则化神经网络。所谓正则化是指训练阶段为神经网络添加惩罚项来限制网络的复杂度。经正则化后, 神经网络可有效避免过拟合现象的发生, 提高模型的泛化能力。通常来讲, 神经网络的性能函数可表示为:

$$F = E \quad (11)$$

经过添加惩罚项 E_w 后, 其性能函数转变为:

$$F = \alpha E_w + \beta E \quad (12)$$

$$E_w = \frac{1}{2} \|\mathbf{W}\|^2 \quad (13)$$

α 和 β 的相对大小决定了惩罚项所占比例。如果 $\alpha \ll \beta$, 则近似于无正则化条件, 侧重于最小化训练误差, 但可能过拟合。若 $\alpha \gg \beta$, 则侧重于限制网络权值规模, 但可能降低模型预测性能。因此, 如何确定 α 和 β 值是非常重要的。在贝叶斯分析的框架下, MK推导出[27]:

$$\alpha = \gamma / (2E_w) \quad (14)$$

$$\beta = \frac{N - \gamma}{2E} \quad (15)$$

式中, $\gamma = N - 2\alpha \text{tr}(\mathbf{H})^{-1}$ 表示有效权值数, N 为样本总数; \mathbf{H} 是 F 的海森矩阵:

$$\mathbf{H} = \nabla^2 \alpha E_w + \nabla^2 E \quad (16)$$

但海森矩阵计算量很大, Foresee和Hagan用高斯-牛顿法近似计算海森矩阵, 大大降低了计算量[28]:

$$\mathbf{H} = \nabla^2 F \approx 2\beta \mathbf{J}^T \mathbf{J} + 2\alpha \mathbf{I}_N \quad (17)$$

式中, \mathbf{J} 是训练误差的雅可比矩阵。

2.3.3. SCG 算法

SCG算法是一种标准BP神经网络的改进算法。在神经网络训练中, 传统的梯度下降法只能沿负梯度的方向进行收敛, 并不一定会正确收敛到全局极小点。共轭梯度法则可将负梯度方向与上一次搜索方向结合起来, 计算出新的搜索方向。但该算法每次迭代时都需要重新确定搜索方向, 进而导致其计算量巨大。Moller [29]提出的SCG算法将模型置信区间法与共轭梯度法结合起来, 成功解决了上述问题。SCG算法的权值调整方法转变为:

$$w_{k+1} = w_k + \theta_k \mathbf{p}_k \quad (18)$$

式中, w_k 为 \mathbf{W}_k 中一点; \mathbf{p}_k 为第 k 次迭代时的搜索方向; θ_k 为第 k 次迭代时的搜索步长。

$$\theta_k = -\frac{\mathbf{g}_k^T \mathbf{p}_k}{\mathbf{p}_k^T \mathbf{H}_k \mathbf{p}_k} \quad (19)$$

式中, \mathbf{g}_k 为第 k 次迭代时的梯度; \mathbf{H}_k 为第 k 次迭代时的海森矩阵。

令 $\mathbf{s}_k = \mathbf{H}_k \mathbf{p}_k$, $\delta_k = \mathbf{p}_k^T \mathbf{s}_k$, $u_k = -\mathbf{g}_k^T \mathbf{p}_k$, 则有 $\theta_k = u_k / \delta_k$ 。同时为了保证 $\delta_k > 0$, 令 $\delta_k = \mathbf{p}_k^T \mathbf{s}_k + \lambda_k |\mathbf{p}_k|^2$ 。

最后可得到步长 θ_k :

$$\theta_k = \frac{u_k}{\delta_k} = \frac{u_k}{\mathbf{p}_k^T \mathbf{s}_k + \lambda_k |\mathbf{p}_k|^2} \quad (20)$$

2.4. 模型性能评价指标

本研究采用相关系数 (R) [式 (21)] 评估各个模型的预测能力, 采用均方误差 (MSE) [式 (22)] 评估训练过程中各模型的拟合能力。 R 与 MSE 均在 Matlab R2016a 中计算完成。

$$R = \frac{\text{Cov}(\text{TN}_P, \text{TN}_A)}{\sigma_P \sigma_A} \quad (21)$$

$$\text{MSE} = \frac{1}{m} \sum_{n=1}^m (\text{TN}_{Pn} - \text{TN}_{An})^2 \quad (22)$$

式中, $\text{Cov}(\text{TN}_P, \text{TN}_A)$ 代表出水 TN 预测值与实际值的协方差; σ_P 、 σ_A 分别为出水 TN 预测值与实际值的标准差; m 为数据集样本量; P_i 为第 i 个样本的出水 TN 预测值; A_i 为第 i 个样本的出水 TN 实际值。

3. 结果与讨论

3.1. 采用运行参数作为主要输入变量构建 FFNN 模型的可行性

建模时首先确定神经网络隐藏层结点数。在其他参数不变的情况下, 根据式 (2) 计算可得, 隐藏层结点数 q 应取在 3~12 之间。训练完成后, 不同结点数模型的均方误差如图 2 所示。对比发现, 神经网络隐藏层结点数为 8 时, MSE 最小, 因而确定基础 FFNN 模型的网络结构为 6-8-1。

为训练和评估基础 FFNN 模型, 本研究将 SBR 工艺长期仿真实验中所收集的数据随机分为训练集 (104 组) 和测试集 (20 组)。图 3 为模型对整体数据集的拟合程

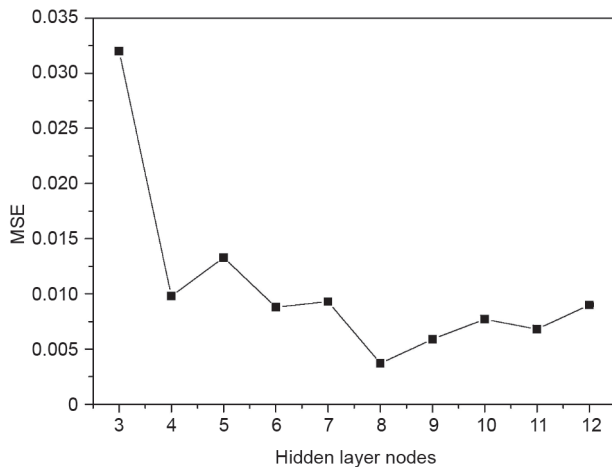


图2. 具有不同隐藏层结点数的FFNN模型的均方误差。

度, 可见其左侧 (训练集) 的预测值与实际值较为接近 ($R = 0.91973$), 表明采用可控参数 (如表面风速、缺氧段时长) 代替 DO 作为输入构建模型是可行的。但右侧测试集中预测值与实际值相差甚远 ($R = 0.5057$), 说明该模型仍存在过拟合或陷于局部极小值等降低模型预测性能的问题。针对这些问题, 本研究采用 L-M、BR 和 SCG 三种算法对神经网络模型进行优化。同时采用更复杂的数据集对模型进行训练, 并针对不同的算法建立更合理的训练模式, 找出最适合模型的算法。

3.2. 采用 L-M 算法优化 FFNN 模型

将数据集随机分为训练集 (80%) 和验证集 (20%)。通过尝试法优选后, 设置隐藏层结点数为 30, 最小失败次数为 100, 隐藏层与输出层间传递函数为 `tansig`。训练 113 步后, 因验证集失败次数超过设定值, 停止训练, 共用时 1.0 s, 其拟合结果如图 4 所示。由图可知, 该模型的 MSE 为 0.00200, 在训练集、验证集与全集中的 R 值分别为: 0.93603、0.92316、0.93392, 表明该模型对出水 TN 的预测性能良好。L-M 算法优化后的模型比之前的基础 FFNN 模型具有更高的 R 值和更快的收敛速度, 不仅有效避免了局部最小值, 而且提高了收敛到全局最优的速度。

3.3. 采用 BR 算法优化 FFNN 模型

BR 算法不需要验证集, 因此将数据随机分为训练集 (80%) 与测试集 (20%)。模型调试过程发现, 当隐藏层结点数小于 20 时, 预测值与实际测量值之间的相关系数低于 0.6, 表明模型拟合程度较差; 随着隐藏层

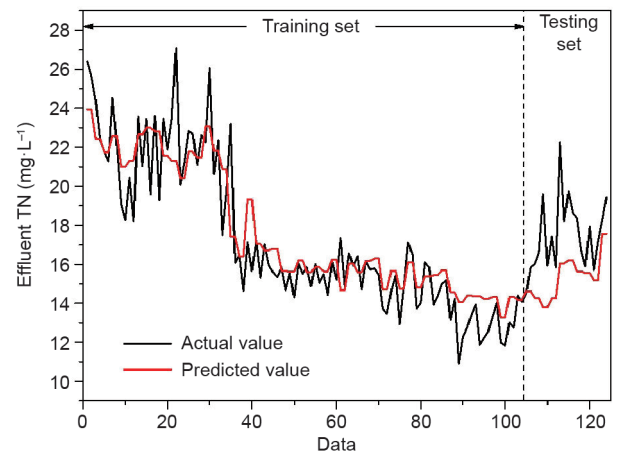


图3. FFNN 模型预测值与实际值的比较。

结点数逐渐增加至30, R 值逐渐升高, 结点数超过30后 R 值升高不明显, 但训练时间大幅增长, 故选取隐藏层结点数为30。选取Tansig作为传递函数, 设置训练步数上限为1000后, 该模型训练用时19 s。如图5所示, 训练集的 R 值为0.90232, 而测试集的 R 值为0.83000, 说明该模型的预测性能优于基础模型, 但弱于L-M算法优化的模型。BR算法通过限制网络的复杂性提高了模型性能, 使得 R 值比基础模型更高。尽管如此, 正规化可能会使得模型丢失一些输入输出之间的关系, 特别是在模型结构复杂性中等的情况下。

3.4. 采用 SCG 算法优化 FFNN 模型

与L-M算法相似, 数据集被划分为训练集 (80%) 和验证集 (20%)。隐层结点数选为30, 以保证与三种算法模型结构一致。选择Tansig作为隐藏层与输出层之间的传递函数, 最小失败次数设为100。因失败次数超

过预设值, 该模型在训练156步后停止, 其训练时间接近于零。如图6所示, 其训练集的 R 值为0.916, 验证集的 R 值为0.962。验证集 R 值高于其训练集 R 值表明该模型不存在过拟合现象且具有良好的预测性能。

3.5. 三种优化算法的比较

基于不同优化算法的FFNN模型训练参数如表3所示。由表可知, 训练期间三种算法优化后的FFNN模型的MSE均较低 (< 0.001), 表明其拟合能力均优于基础模型 ($MSE = 0.0037$)。其中, SCG算法的训练时间最短。

基于各优化算法所构建的FFNN模型 R 值如表4所示。训练期间三种模型的 R 值分别为0.936、0.902和0.917, 说明模型训练集的输入 (进水水质和运行参数) 和输出 (出水TN) 相关性较强。在预测阶段, 所有FFNN模型的 R 值均高于基础模型, 表明优化算法大大提高了FFNN神经网络模型对出水总氮的预测性能。其

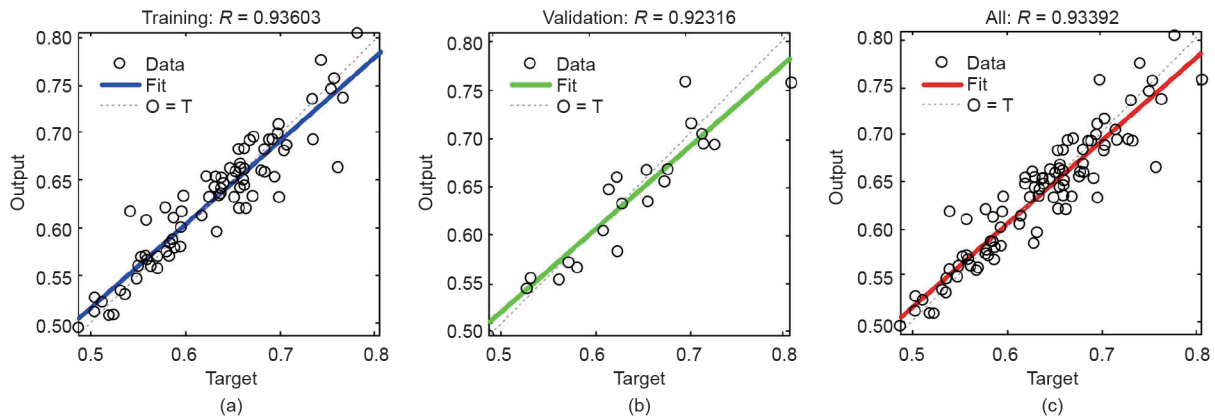


图4. 基于L-M算法的FFNN模型仿真结果。(a) 训练集 (输出 $\approx 0.88 \times$ 目标+0.075); (b) 验证集 (输出 $\approx 0.88 \times$ 目标+0.074); (c) 全集 (输出 $\approx 0.88 \times$ 目标+0.077)。“Fit”一行表示目标和输出之间的关系;“O=T”一行表示目标等于输出。

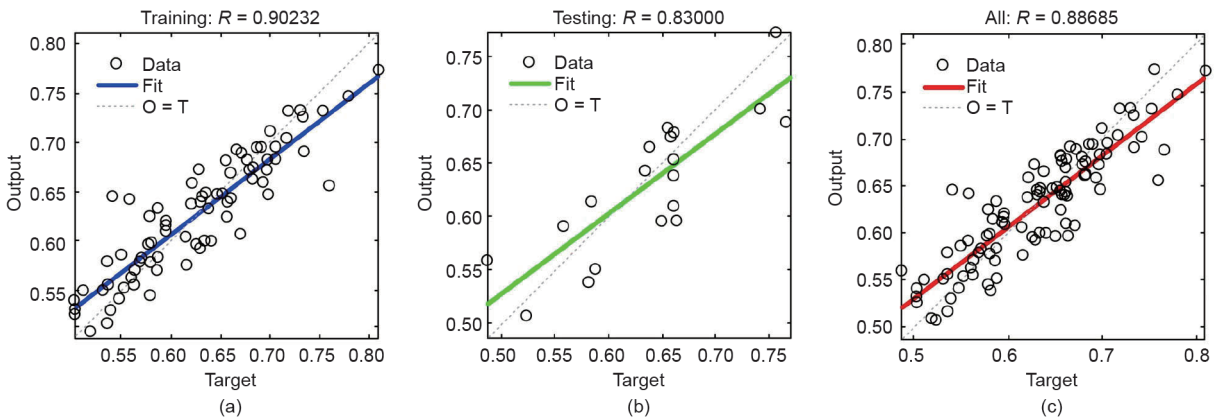


图5. 基于BR算法的FFNN模型仿真结果。(a) 训练集 (输出 $\approx 0.76 \times$ 目标+0.15); (b) 测试集 (输出 $\approx 0.75 \times$ 目标+0.15); (c) 全集 (输出 $\approx 0.76 \times$ 目标+0.15)。

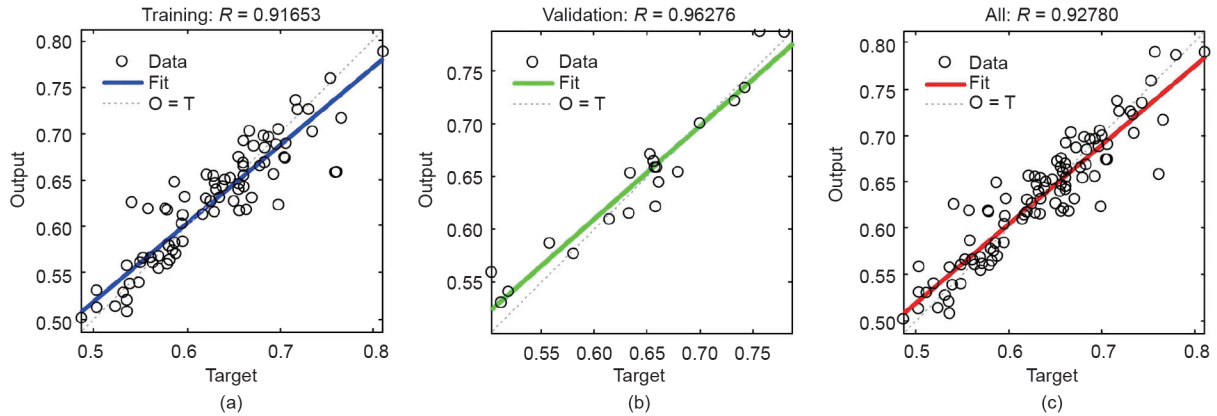


图6. 基于SCG算法的FFNN模型训练结果。(a) 训练集 (输出 $\approx 0.84 \times$ 目标 $+0.1$); (b) 验证集 (输出 $\approx 0.88 \times$ 目标 $+0.079$); (c) 全集 (输出 $\approx 0.85 \times$ 目标 $+0.092$)。

表3 基于不同优化算法的FFNN模型训练参数

Optimization algorithm	Training step	Training time (s)	MSE
L-M	113	1	0.000200
BR	1000	19	0.000935
SCG	156	$\ll 1$	0.000512

中, 基于SCG算法的模型 R 值最高并高于其训练集的 R 值且训练时间最短, 表明SCG算法能够使模型获得最快的计算速度和出色的预测能力, 是最佳的优化算法。

3.6. 与其他模型比较

针对厌氧-缺氧-好氧工艺(A²/O)、SBR和氧化沟等不同的污水处理过程, 研究者陆续提出了各种适合于特定环境仿真模型用来预测出水水质。如为预测屠宰场废水中NH₄⁺-N的去除率, Kundu等[30]构建了基于进水水质和DO的BP神经网络模型。基于进水参数构建的ANN模型可预测好氧颗粒污泥的TN去除效率[31]。Ebrahimi等[32]提出了多变量回归模型预测氧化沟的出水生化需氧量(BOD)、TP浓度。然而, 上述研究选择的输入参数主要为DO与进水水质, 具有不可控性。与以往研究相比, 本研究从仿真模型的可控性与预测精度两个方面出发, 首次证实FFNN模型可采用实际可控参数(如表面风速和缺氧时间)代替不可控的DO实现出水总氮预测。为提高预测精度, 采用多种算法对FFNN模型进行优化, SCG优化后的模型预测精度最高。在此基础上, 将构建的模型与多种机器学习模型对比(表5), 结果表明经SCG算法优化的FFNN模型预测性能最佳, 具有最小的RMSE和最高的 R 值。总的来看, 研究构建的模型可根据进水水质变化调节可控运行参数, 以保证出水TN处于稳定达标状态, 同时有效避免过度曝

表4 基于各优化算法所构建的FFNN模型 R 值

Dataset	L-M	BR	SCG
Training set	0.93603	0.90232	0.91653
Validation set	0.92316	—	0.96276
Testing set	—	0.83000	—
The whole set	0.93392	0.88685	0.92780

表5 不同预测方法的性能比较 (R 值四舍五入为小数点后两位数)

Method	MSE	R
FFNN	0.000512	0.93
LR	0.009622	0.79
Lasso	0.009696	0.79
Ridge	0.009299	0.80
SVR	0.007687	0.84
RF	0.006941	0.85

气[33], 降低系统运行能耗。

4. 结论

本研究基于主流SBR工艺构建了一种基于可控参数的前馈神经网络出水总氮预测模型, 选择表面风速与缺氧时长代替DO作为FFNN模型的主要输入参数, 运用长期仿真数据提高了模型可用可控性; 选择SCG作为FFNN模型优化算法, 实现出水TN及最佳运行参数的精准预测, 可为污水处理系统的稳定运行及节能降耗提供有价值的解决方案。

致谢

本文得到了国家水体污染控制与治理科技重大专项(2017ZX07201003)、国家自然科学基金(51961125101)、

浙江省重点研发计划（2018C03003）的大力支持。

Compliance with ethics guidelines

Zihao Zhao, Zihao Wang, Jialuo Yuan, Jun Ma, Zheling He, Yilan Xu, Xiaojia Shen, and Liang Zhu declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2020.07.027>.

References

- [1] National Bureau of Statistics of China. China statistical yearbook 2019. Beijing: China Statistics Press; 2019. Chinese.
- [2] Lu J, Wang X, Liu H, Yu H, Li W. Optimizing operation of municipal wastewater treatment plants in China: the remaining barriers and future implications. *Environ Int* 2019;129:273–8.
- [3] Irvine RL, Ketchum LH Jr, Asano T. Sequencing batch reactors for biological wastewater treatment. *Crit Rev Environ Control* 1989;18(4):255–94.
- [4] Vieira M, Brito AG, Nogueira R. Nitrogen removal in a sequencing batch biofilm reactor: effect of carbon availability and intermittent aeration. *World Rev Sci Technol Sustainable Dev* 2009;6(2–4):173.
- [5] Ren Y, Ngo HH, Guo W, Wang D, Peng L, Ni BJ, et al. New perspectives on microbial communities and biological nitrogen removal processes in wastewater treatment systems. *Bioresour Technol* 2020;297:122491.
- [6] Hu L, Wang J, Wen X, Qian Y. Study on performance characteristics of SBR under limited dissolved oxygen. *Process Biochem* 2005;40(1):293–6.
- [7] Boon AG. Sequencing batch reactors: a review. *Water Environ J* 2003;17(2):68–73.
- [8] Singh M, Srivastava RK. Sequencing batch reactor technology for biological wastewater treatment: a review. *Asia-Pac J Chem Eng* 2011;6(1):3–13.
- [9] Ingimundarson A, Ocampo-Martinez C, Bemporad A. Suboptimal model predictive control of hybrid systems based on mode-switching constraints. In: *Proceedings of the 2007 46th IEEE Conference on Decision and Control*; 2007 Dec 12–14; New Orleans, LA, USA. New York: IEEE; 2008.
- [10] Olsson G, Newell B. Wastewater treatment systems: modelling, diagnosis and control. London: IWA Publishing; 2015.
- [11] Baxter CW, Zhang Q, Stanley SJ, Shariff R, Tupas R-RT, Stark HL. Drinking water quality and treatment: the use of artificial neural networks. *Can J Civ Eng* 2001;28(S1):26–35.
- [12] Miron M, Frangu L, Ifrim G, Garaman S. Modeling of a wastewater treatment process using neural networks. In: *Proceedings of the 2016 20th International Conference on System Theory, Control and Computing*; 2016 Oct 13–15; Sinaia, Romania. New York: IEEE; 2016.
- [13] Fan M, Li T, Hu J, Cao R, Wu Q, Wei X, et al. Synthesis and characterization of reduced graphene oxide-supported nanoscale zero-valent iron (nZVI/rGO) composites used for Pb(II) removal. *Materials* 2016;9(8):687.
- [14] Dreyfus G. Neural networks: an overview. In: Dreyfus G, editor. *Neural networks—methodology and applications*. Heidelberg: Springer; 2005. p. 1–83.
- [15] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Netw* 1989;2(5):359–66.
- [16] Ata R. Artificial neural networks applications in wind energy systems: a review. *Renew Sustain Energy Rev* 2018;84:173. Retraction of: Ata R. *Renew Sustain Energy Rev* 2015;49:534–62.
- [17] Almási A, Woźniak S, Cristea V, Leblebici Y, Engbersen T. Review of advances in neural networks: neural design technology stack. *Neurocomputing* 2016;174 (Pt A):31–41.
- [18] Lin D, Vasilakos AV, Tang Y, Yao Y. Neural networks for computer-aided diagnosis in medicine: a review. *Neurocomputing* 2016;216:700–8.
- [19] Lesnik KL, Liu H. Predicting microbial fuel cell biofilm communities and bioreactor performance using artificial neural networks. *Environ Sci Technol* 2017;51(18):10881–92.
- [20] Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, et al. Mastering the game of Go without human knowledge. *Nature* 2017;550(7676):354–9.
- [21] Qian N. On the momentum term in gradient descent learning algorithms. *Neural Netw* 1999;12(1):145–51.
- [22] Sadrzadeh M, Mohammadi T, Ivakpour J, Kasiri N. Neural network modeling of Pb²⁺ removal from wastewater using electro dialysis. *Chem Eng Process* 2009;48(8):1371–81.
- [23] Association APH. Standard methods for the examination of water and wastewater. 21st ed. Washington, DC: American Public Health Association; 2005.
- [24] Ojha VK, Abraham A, Snášel V. Metaheuristic design of feedforward neural networks: a review of two decades of research. *Eng Appl Artif Intel* 2017;60:97–116.
- [25] Anzai Y. Learning by neural networks. In: Anzai Y, editor. *Pattern recognition & machine learning*. San Francisco: Morgan Kaufmann; 1992. p. 297–335.
- [26] Zhang S, Choromanska A, LeCun Y. Deep learning with elastic averaging SGD. 2014. arXiv:1412.6651.
- [27] MacKay DJC. A practical bayesian framework for backpropagation networks. *Neural Comput* 1992;4(3):448–72.
- [28] Foresee F, Hagan M. Gauss–Newton approximation to Bayesian regularization. In: *Proceedings of International-joint Conference on Neural Networks*; 1997 Jun 9–12; Houston, TX, USA; New York: IEEE; 1997.
- [29] Moller MF. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks* 1993;6(4):525–33.
- [30] Kundu P, Debsarkar A, Mukherjee S. Artificial neural network modeling for biological removal of organic carbon and nitrogen from slaughterhouse wastewater in a sequencing batch reactor. *Adv Artif Neural Syst* 2013;2013:268064.
- [31] Gong H, Pishgar R, Tay JH. Artificial neural network modelling for organic and total nitrogen removal of aerobic granulation under steady-state condition. *Environ Technol* 2019;40(24):3124–39.
- [32] Ebrahimi M, Gerber EL, Rockaway TD. Temporal performance assessment of wastewater treatment plants by using multivariate statistical analysis. *J Environ Manage* 2017;193:234–46.
- [33] Brđjanovic D, Slamet A, van Loosdrecht MCM, Hooijmans CM, Alaerts GJ, Heijnen JJ. Impact of excessive aeration on biological phosphorus removal from wastewater. *Water Res* 1998;32(1):200–8.