

基于动态模糊神经网络的出水含氮参数软测量方法

蒙西[†], 张寅, 乔俊飞

(北京工业大学 信息学部, 北京 100124; 智慧环保北京实验室, 北京 100124;

智能感知与自主控制教育部工程研究中心, 北京 100124)

摘要: 针对城市污水处理过程出水氨氮($\text{NH}_4^+\text{-N}$)和出水总氮(TN)难以实时准确检测的问题, 文中提出了一种基于动态模糊神经网络(DFNN)的出水含氮参数软测量方法. 首先, 采用自组织增删机制和快速二阶学习算法构建模糊神经网络(FNN), 以快速获得结构精简的软测量模型; 其次, 引入自适应激活强度阈值设计FNN分级更新策略, 确保软测量模型在非平稳环境下的预测精度; 最后, 通过基准仿真1号模型(BSM1)平台的数据验证了DFNN软测量方法的有效性, 实验结果表明, 所提出的方法能够实现出水 $\text{NH}_4^+\text{-N}$ 和出水TN的在线精准检测.

关键词: 城市污水处理过程; 模糊神经网络; 分级更新; 出水含氮量; 软测量

引用格式: 蒙西, 张寅, 乔俊飞. 基于动态模糊神经网络的出水含氮参数软测量方法. 控制理论与应用, 2024, 41(12): 2383–2392

DOI: 10.7641/CTA.2023.20667

Soft-sensing method for effluent nitrogen parameters based on a dynamic fuzzy neural network

MENG Xi[†], ZHANG Yin, QIAO Jun-fei

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China;

Beijing Laboratory of Smart Environmental Protection, Beijing 100124, China;

Engineering Research Center of Intelligent Perception and Autonomous Control, Ministry of Education, Beijing 100124, China)

Abstract: Aiming at the real-time and accurate measurements of the effluent ammonium nitrogen ($\text{NH}_4^+\text{-N}$) and the effluent total nitrogen (TN) in municipal wastewater treatment process, a soft-sensing method for effluent nitrogen parameters based on a dynamic fuzzy neural network (DFNN) is proposed in this paper. First, by utilizing a self-organizing growing-and-pruning mechanism and an improved second-order learning algorithm, a fuzzy neural network (FNN) is constructed in order to obtain a soft-sensing model with a simplified structure. Then, by introducing an adaptive firing strength threshold, a hierarchical updating strategy of FNN is designed, which can effectively ensure the prediction accuracy of the soft-sensing model under non-stationary environments. Finally, the effectiveness of the proposed DFNN soft-sensing method is verified based on the simulation data which were provided by the benchmark simulation model No. 1 (BSM1) platform. The simulation results show that the proposed soft-sensing method can achieve online and accurate measurements of the effluent $\text{NH}_4^+\text{-N}$ and the effluent TN.

Key words: municipal wastewater treatment process; fuzzy neural network; hierarchical updating; effluent nitrogen concentration; soft-sensing

Citation: MENG Xi, ZHANG Yin, QIAO Junfei. Soft-sensing method for effluent nitrogen parameters based on a dynamic fuzzy neural network. *Control Theory & Applications*, 2024, 41(12): 2383–2392

1 引言

城市污水处理是水环境保护和水资源持续再生利用的重要途径. 出水氨氮(ammonium nitrogen, $\text{NH}_4^+\text{-N}$)和出水总氮(total nitrogen, TN)是城市污水处理厂

的重要监测指标, 对其浓度的实时获取是实现脱氮过程稳定高效运行的前提. 受限于检测环境、仪器仪表成本等因素, $\text{NH}_4^+\text{-N}$, TN等含氮参数难以实时精准检测.

收稿日期: 2022-07-26; 录用日期: 2023-10-11.

[†]通信作者. E-mail: mengxi@bjut.edu.cn.

本文责任编辑: 吴敏.

国家自然科学基金项目(61903012, 622731013, 61890930-5, 62021003), 科技创新2030—“新一代人工智能”重大项目(2021ZD0112301, 2021ZD0112302), 国家重点研发计划项目(2019YFC1906004-2)资助.

Supported by the National Natural Science Foundation of China (61903012, 62273013, 61890930-5, 62021003), the National Key Research and Development Program of China (2021ZD0112301, 2021ZD0112302) and the National Key Research and Development Project (2019YFC1906004-2).

软测量方法具有良好的推理估计能力和快速的动态响应能力,能借助于易测辅助变量实现不可测或难测主导变量的在线实时连续估计,因此,成为城市污水处理过程(wastewater treatment process, WWTP)等复杂工业过程中广受青睐的关键参数在线检测方法^[1-2]。根据软测量模型的性质,可以分为基于机理模型的软测量方法与基于数据驱动模型的软测量方法两类^[3]。常用的机理模型为国际水协推出的活性污泥系列模型(activated sludge models, ASMs)^[4]。文献[5]对活性污泥3号模型进行了改进,并基于改进后的模型实现了对出水TN的软测量。Luo等^[6]基于改进后的ASM,即全耦合活性污泥4号模型实现了城市污水处理厂出水TN的软测量。由于城市WWTP包含复杂的生化反应,建立机理模型时通常会引入诸多假设条件,同时,模型参数众多且难以辨识,进而影响了软测量精度。随着数据挖掘、人工智能等技术的飞速发展,基于数据驱动模型的软测量方法已成为城市WWTP关键参数实时检测领域的研究热点^[7-9]。其中,人工神经网络(artificial neural network, ANN)凭借其强大的自学习能力和非线性逼近能力,成为建立数据驱动软测量模型的主流方法^[10-11]。文献[10]和文献[11]分别基于多层感知机和径向基函数(radial basis function, RBF)神经网络构建了出水水质参数软测量模型,实现了对出水参数的准确预测。模糊神经网络(fuzzy neural network, FNN)集成了神经网络的学习能力和模糊逻辑的不确定信息处理能力,因此被广泛用于描述复杂系统的非线性特性^[12]。Zaghloul等^[13]采用混合优化算法对自适应模糊推理系统的参数进行调整,进而建立了出水化学需氧量(chemical oxygen demand, COD)和出水 NH_4^+-N 的软测量模型。文献[14]引入了建模误差概率密度函数,设计了一种基于自适应梯度下降算法的FNN,并基于该FNN建立了出水 NH_4^+-N 软测量模型,提高了出水 NH_4^+-N 预测的准确性。尽管上述方法都取得了较为理想的预测精度,但城市WWTP被动接受进水流量、进水组分、污染物种类等不确定性因素的影响,所处环境是非平稳、动态时变的,当实际工业过程发生变化时,模型精度也将随之下降。因此,为了确保软测量模型在未知环境、非平稳状态下的稳定性及可靠性,软测量模型在应用过程中的自适应更新是不可或缺的环节。

为了解决以上问题,部分学者开始将动态更新方法引入城市WWTP及其他流程工业过程中。针对出水生化需氧量(biochemical oxygen demand, BOD)软测量模型, Xiong等^[15]基于模型预测误差设计了一种选择性更新策略,通过调节误差阈值,在模型的预测精度与更新效率之间取得平衡,进而有效提高了BOD的软测量精度。然而,以单个样本的预测误差为依据控制模型更新,使得模型容易受到异常值的影响,并

且可能会导致模型更新不及时,进一步影响模型的预测精度。Wu等^[16]在建立出水BOD软测量模型时引入了时间差分策略,增强了模型在动态环境中的鲁棒性。为了提高软测量模型的自适应能力,文献[17]在采用长短期记忆网络建立软测量模型时,引入了滑动窗口学习法来实现对网络权值的动态调整。Fan等^[18]采用一种在线学习机制,通过将在线样本加入训练集来对模型参数进行实时调整,以实现出水 NH_4^+-N 软测量模型在复杂、时变环境下的动态自适应。文献[16-18]中的更新方法需要在每次在线样本到达时对模型全部参数进行调整,虽然提升了预测精度,但同时也影响了模型的更新效率。为此,陆荣秀等^[19]提出了一种局部模型更新策略,根据萃取过程工况的实际变化情况决定每个局部模型更新与否。文献[20]设计了一种局部更新方法,采用递归最小二乘(recursive least square, RLS)算法对具有较大激活强度的规则后件参数进行更新,取得了优于全局更新的结果。Gu等^[21]同样采取了局部更新的思想,根据在线数据对FNN模型规则后件参数进行调整,以保证模型在非平稳环境中的良好性能。虽然上述研究均从模型更新的精度以及效率角度出发进行了一系列探索,但如何同时确保自适应更新策略的精度和效率仍是一个开放问题。

针对上述问题,文中提出了基于动态模糊神经网络(dynamic FNN, DFNN)的出水含氮参数软测量方法,该方法结合网络学习精度和规则激活强度进行软测量模型结构自组织设计,进而提高建模效率和模型精度,同时引入自适应阈值设计模型分级更新策略,提高模型的更新效率与精度,实现了对出水 NH_4^+-N 和出水TN的在线精准检测。

2 基础知识

2.1 城市污水处理脱氮过程描述

目前我国城市污水处理厂广泛采用活性污泥法去除污水中的污染物,其中,氮是城市WWTP中主要的污染物指标,出水 NH_4^+-N , TN等含氮参数指标不达标将引起水体富营养化问题,严重危害水体生态环境。因此,生物脱氮一直是城市污水处理的重点和难点。以A/O工艺为例(如图1所示),首先,污水通过格栅、初沉池,一些较大体积的悬浮物可被去除;然后,污水进入缺氧池进行有机物的初步降解,在缺氧池中同时还进行着内循环回流的亚硝态氮(nitrite nitrogen, NO_2^--N)和硝态氮(nitrate nitrogen, NO_3^--N)的反硝化脱氮,此过程中 NO_2^--N 和 NO_3^--N 被反硝化细菌还原成氮气释放,从而实现污水脱氮;随后,污水流入好氧池发生硝化反应, NH_4^+-N 被硝化细菌氧化为 NO_2^--N 和 NO_3^--N ;同时,好氧池中的硝化混合液和沉淀后的部分污泥回流到缺氧池,缺氧池从回流液中得到大量的 NO_2^--N 和 NO_3^--N ,继续进行反硝化反应^[22]。在上

述过程中,脱氮效果主要取决于硝化与反硝化反应. 为了对脱氮过程进行有效的控制,提高脱氮效率,出

水含氮参数,即出水NH₄⁺-N、出水TN的实时准确检测至关重要.

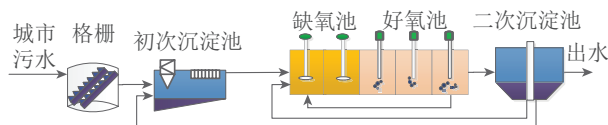


图 1 城市污水处理生物脱氮过程简图

Fig. 1 Schematics of biological denitrification process in municipal wastewater treatment

2.2 FNN

文中所采用的FNN结构如图2所示,包括输入层、RBF层、规则化层、输出层4层.该FNN中第j条模糊规则具有如下形式:

Rule *j*: IF x_1 is A_{1j} ... and x_i is A_{ij} and ...
and x_M is A_{Mj} ,

THEN y_j is w_j , $j = 1, 2, \dots, K$, (1)

其中: x_i 为第*i*维输入变量, A_{ij} 是第*j*个规则对第*i*个输入变量的隶属函数, y_j 是第*j*个规则的输出, w_j 为第*j*个规则的后件参数, M 为输入变量维数, K 为模糊规则个数. FNN每一层输出的数学描述如下.

输入层: 该层共有*M*个神经元,代表了*M*维输入向量.对于输入变量 $\mathbf{x} = [x_1 \dots x_M]^T$,该层的输出可以表示为

$$u_i = x_i, i = 1, 2, \dots, M, \quad (2)$$

其中 u_i 为输入层第*i*个神经元的输出.

RBF层: 该层共有*K*个神经元,每个神经元代表了模糊规则的前件.该层的输出可以计算为

$$\varphi_j = \prod_{i=1}^M A_{ij}(x_i), j = 1, 2, \dots, K, \quad (3)$$

其中: φ_j 为RBF层第*j*个神经元的输出; $A_{ij}(\cdot)$ 表示该层第*j*个规则对第*i*个输入的隶属函数,文中采用高斯型隶属函数,即

$$A_{ij}(x_i) = \exp\left(-\frac{(x_i - c_{ij})^2}{2\sigma_{ij}^2}\right), \quad (4)$$

其中 c_{ij} 和 σ_{ij} 分别为隶属函数的中心与宽度.

规则化层: 该层共有*K*个神经元,其输出可以表示为

$$h_j = \frac{\varphi_j}{\sum_{j=1}^K \varphi_j}, \quad (5)$$

其中 h_j 为该层第*j*个神经元的输出.

输出层: 该层的作用是解模糊,对于单输出FNN来说,网络输出可计算为

$$y = \sum_{j=1}^K w_j h_j = \sum_{j=1}^K w_j \frac{\exp\left(-\sum_{i=1}^M \frac{(x_i - c_{ij})^2}{2\sigma_{ij}^2}\right)}{\sum_{j=1}^K \exp\left(-\sum_{i=1}^M \frac{(x_i - c_{ij})^2}{2\sigma_{ij}^2}\right)}, \quad (6)$$

其中 w_j 为第*j*个规则化层神经元到输出层的连接权值,代表了规则的后件参数.

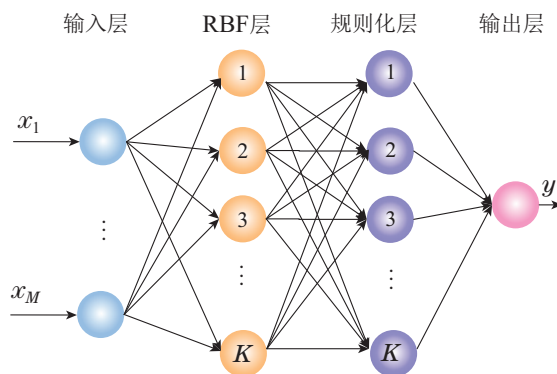


图 2 FNN结构

Fig. 2 The structure of FNN

3 基于动态模糊神经网络的出水含氮参数软测量方法

为了实现不确定、非平稳环境下出水含氮参数的准确测量,文中提出了一种基于DFNN的出水含氮参数软测量方法,包括基于互信息的辅助变量选择和DFNN设计两部分.

3.1 基于互信息的辅助变量选择

城市WWTP反应复杂,影响出水NH₄⁺-N和出水TN的变量众多且相互耦合,相互关联.为了提高软测量模型的计算效率,文中基于互信息进行辅助变量选择.互信息能够表示两个随机变量之间的关联程度,其定义来源于信息论中熵的概念.对于随机变量 $X = \{x_i | i = 1, 2, \dots, N\}$,其熵计算公式为

$$H(X) = -\sum_{i=1}^N p_X(x_i) \log p_X(x_i), \quad (7)$$

其中 $p_X(\cdot)$ 为*X*的概率密度函数.对于两个随机变量和 $X = \{x_i | i = 1, 2, \dots, N\}$ 和 $Y = \{y_j | j = 1, 2, \dots, M\}$,其联合熵可计算为

$$H(X, Y) = - \sum_{i=1}^N \sum_{j=1}^M p_{X,Y}(x_i, y_j) \log p_{X,Y}(x_i, y_j), \quad (8)$$

其中 $p_{X,Y}(\cdot)$ 为 X, Y 的联合概率密度函数. X, Y 之间的互信息计算公式为

$$I(X, Y) = H(X) + H(Y) - H(X, Y). \quad (9)$$

当变量 X 和 Y 完全不相关时, 互信息值为0; 反之, 两个变量的相关性越高, 计算得到的互信息值也就越大.

首先, 根据上述互信息计算方法, 计算出所有变量与出水 $\text{NH}_4^+\text{-N}$ 、 TN 之间的相关性; 然后, 结合专家知识剔除了实际应用中难测得或测量周期较长的变量; 最后, 根据人工经验选取了6个与出水含氮参数相关性较大的变量作为辅助变量. 表1、表2分别给出了基于基准仿真1号模型(benchmark simulation model no.1, BSM1)阴雨天气数据所得到的出水 $\text{NH}_4^+\text{-N}$ 、出水 TN 软测量模型输入变量.

表1 出水 $\text{NH}_4^+\text{-N}$ 软测量模型输入变量相关性分析

Table 1 Correlation analysis of the input variables selected for effluent $\text{NH}_4^+\text{-N}$ soft-sensing model

序号	变量名称	相关性分析结果
1	混合液悬浮固体	0.8927
2	出水固体悬浮物	0.8831
3	出水COD	0.8571
4	进水TN	0.8505
5	进水固体悬浮物	0.8383
6	进水 $\text{NH}_4^+\text{-N}$	0.8125

表2 出水 TN 软测量模型输入变量相关性分析

Table 2 Correlation analysis of the input variables selected for effluent TN soft-sensing model

序号	变量名称	相关性分析结果
1	混合液悬浮固体	0.8552
2	出水固体悬浮物	0.8415
3	出水COD	0.8134
4	进水TN	0.8084
5	进水COD	0.7948
6	进水流量	0.6456

3.2 DFNN设计

DFNN的设计主要可分为自组织增删机制与分级更新策略两部分. 其中, DFNN自组织增删机制又包括规则增长阶段、规则修剪阶段以及网络参数调整.

3.2.1 规则增长阶段

初始时刻, 网络中不含有规则. 根据一种误差补偿算法^[23]的思想, 逐个增长网络中的规则, 即增加相应的RBF神经元与规则化层神经元, 以补偿每次迭代开始时误差曲面中的最大误差. 考虑到初始时刻没有规则, 即RBF神经元与规则化层神经元个数均为0, 首先找到训练样本中具有最大绝对输出值的样本 p_1 , 即

$$p_1 = \operatorname{argmax}_{p \in P} \|y_{dp}\|, \quad (10)$$

其中 y_{dp} 为第 p 个样本的期望输出. 然后, 根据样本 p_1 新增第1个规则, 设置相应RBF神经元与规则化层神经元的初始参数分别为

$$\mathbf{c}_1 = \mathbf{x}_{p_1}, \quad (11)$$

$$\boldsymbol{\sigma}_1 = \mathbf{d}_1, \quad (12)$$

$$w_1 = y_{dp_1}, \quad (13)$$

其中: \mathbf{c}_1 和 $\boldsymbol{\sigma}_1$ 分别是第1个RBF神经元的中心与宽度, w_1 是第1个规则化层神经元到输出层的连接权值, \mathbf{x}_{p_1} 和 y_{dp_1} 分别为第 p_1 个样本的输入与输出, $\mathbf{d}_1 = [d_1 \cdots d_1] \in \mathbb{R}^M$ 是元素为 $d_1 = 1$ 的向量.

初始化第1个规则后, 采用第3.2.3节介绍的网络参数调整算法对网络参数进行调整. 网络参数调整完成后, 规则增长阶段将继续进行. 在 k 时刻, 对所有训练样本, 网络误差向量为

$$\mathbf{e}(k) = [e_1(k) \ e_2(k) \ \cdots \ e_p(k) \ \cdots \ e_P(k)]^T, \quad (14)$$

其中 $e_p(k)$ 为 k 时刻第 p 个样本的期望输出与实际输出之间的误差

$$e_p(k) = y_{dp} - y_p(k). \quad (15)$$

然后, 找到当前最大绝对误差对应的样本 p_k , 即

$$p_k = \operatorname{argmax}_{p \in P} \|e_p(k)\|. \quad (16)$$

根据样本 p_k 设置 k 时刻新增RBF神经元与规则化层神经元的初始参数分别为

$$\mathbf{c}_k = \mathbf{x}_{p_k}, \quad (17)$$

$$\boldsymbol{\sigma}_k = \mathbf{d}_k, \quad (18)$$

$$w_k = y_{dp_k}, \quad (19)$$

其中: \mathbf{c}_k 和 $\boldsymbol{\sigma}_k$ 分别是第 k 个RBF神经元的中心与宽度, w_k 为第 k 个规则化层神经元到输出层的连接权值, $\mathbf{d}_k = [d_k \cdots d_k] \in \mathbb{R}^M$ 是元素为 $d_k = 0.7 \min_{j \neq k} \|\mathbf{c}_k - \mathbf{c}_j\|$ 的向量, 其中 $\min_{j \neq k} \|\mathbf{c}_k - \mathbf{c}_j\|$ 表示当前新增RBF神经元与现有神经元的最近欧氏距离. 然后, 采用第3.2.3节中的参数调整算法对当前网络参数进行调整.

规则的增长将按照式(14)–(19)进行, 在此期间采用均方根误差(root mean square error, RMSE)度量网

络学习精度

$$\text{RMSE} = \sqrt{\frac{1}{P} \sum_{p=1}^P (y_p - y_{dp})^2}, \quad (20)$$

其中: P 为训练样本个数, y_p 为第 p 个样本的网络输出。当网络学习精度达到预设期望值或网络中规则个数达到预设最大值时, 规则增长阶段完成, 计算当前网络 RMSE 值 (RMSE1) 并进入下一阶段。

3.2.2 规则修剪阶段

在规则增长阶段结束后, 可获得一个含有 J 个规则的 FNN。为了保证网络结构的精简性, 可删除一些不重要的规则。文中采用规则激活强度衡量规则的重要程度^[24], 激活强度较低的规则被认为在当前网络中发挥的作用较小, 因此, 可以在不影响网络学习精度的前提下对其进行修剪。第 j 个模糊规则的激活强度计算如下:

$$\text{FS}_j = \sum_{p=1}^P \text{FS}_j(\mathbf{x}_p), \quad (21)$$

$$\text{FS}_j(\mathbf{x}_p) = \exp\left(-\sum_{i=1}^M \frac{(x_p^i - c_{ij})^2}{2\sigma_{ij}^2}\right), \quad (22)$$

其中 x_p^i 表示第 p 个样本的第 i 维输入。

首先, 计算当前所有规则的激活强度, 并找到具有最小激活强度的第 j_{\min} 个规则, 即

$$j_{\min} = \underset{j \in J}{\operatorname{argmin}}(\text{FS}_j). \quad (23)$$

然后, 删除第 j_{\min} 个规则并采用第 3.2.3 节所述的参数调整算法调整当前网络参数。参数调整完成后, 计算当前网络 RMSE 值 (RMSE2)。若 $\text{RMSE1} \geq \text{RMSE2}$, 则表明删除第 j_{\min} 个规则不会降低网络学习精度, 此时将按照式 (21)–(23) 继续进行修剪直到 $\text{RMSE1} < \text{RMSE2}$ 。

规则修剪阶段结束后, 网络的结构设计完成。

3.2.3 网络参数调整

每次增长或修剪规则后, 均采用一种改进二阶 (improved second-order, ISO) 算法^[25] 对所有网络参数进行调整, 公式如下:

$$\boldsymbol{\psi}_{t+1} = \boldsymbol{\psi}_t - (\mathbf{Q}_t + \mu \mathbf{I})^{-1} \mathbf{g}_t, \quad (24)$$

其中: $\boldsymbol{\psi}$ 为参数向量, 包含所有需要调整的网络参数 (即中心、宽度、权值), \mathbf{Q} 为类海森矩阵, μ 为学习率, \mathbf{I} 为单位矩阵, \mathbf{g} 为梯度向量。

为了减小雅可比矩阵计算过程中的复杂度, 将类海森矩阵 \mathbf{Q} 转化为多个类海森子矩阵的和, 即

$$\mathbf{Q} = \sum_{p=1}^P \mathbf{q}_p, \quad (25)$$

$$\mathbf{q}_p = \mathbf{j}_p^T \mathbf{j}_p, \quad (26)$$

其中: \mathbf{q}_p 为类海森子矩阵, \mathbf{j}_p 为雅可比向量。

类似地, 将梯度向量 \mathbf{g} 转化为多个梯度子向量的和, 即

$$\mathbf{g} = \sum_{p=1}^P \boldsymbol{\eta}_p, \quad (27)$$

$$\boldsymbol{\eta}_p = \mathbf{j}_p^T e_p, \quad (28)$$

其中 $\boldsymbol{\eta}_p$ 为梯度子向量。因此, 类海森矩阵和梯度向量的计算可转化为雅可比向量 \mathbf{j}_p 的计算, 即

$$\mathbf{j}_p = \begin{bmatrix} \frac{\partial e_p}{\partial w_1} & \cdots & \frac{\partial e_p}{\partial w_j} & \cdots & \frac{\partial e_p}{\partial w_K} & \frac{\partial e_p}{\partial c_{11}} & \cdots & \frac{\partial e_p}{\partial c_{ij}} \\ \cdots & \frac{\partial e_p}{\partial c_{MK}} & \frac{\partial e_p}{\partial \sigma_{11}} & \cdots & \frac{\partial e_p}{\partial \sigma_{ij}} & \cdots & \frac{\partial e_p}{\partial \sigma_{MK}} \end{bmatrix}, \quad (29)$$

其中: M 为输入层神经元个数, K 为 RBF 层神经元个数。

根据链式求导法则, 雅可比向量中每个分量计算如下:

$$\frac{\partial e_p}{\partial w_j} = -h_j(\mathbf{x}_p), \quad (30)$$

$$\frac{\partial e_p}{\partial c_{ij}} = -h_j(\mathbf{x}_p)(1 - h_j(\mathbf{x}_p))\varphi_j(\mathbf{x}_p) \frac{\|x_p^i - c_{ij}\|}{\sigma_{ij}^2}, \quad (31)$$

$$\frac{\partial e_p}{\partial \sigma_{ij}} = -h_j(\mathbf{x}_p)(1 - h_j(\mathbf{x}_p))\varphi_j(\mathbf{x}_p) \frac{\|x_p^i - c_{ij}\|^2}{\sigma_{ij}^3}. \quad (32)$$

3.2.4 分级更新策略

为了确保软测量模型在非平稳环境中具有良好的性能, 文中设计了一种分级更新策略。该更新策略从模型的泛化能力出发, 建立模型预测精度与更新方式之间的联系, 实现模型自适应分级更新。当模型对当前在线数据能够保持良好的预测精度时, 只需对部分激活强度较大的规则进行局部更新, 以提高更新效率; 当模型预测精度较差时, 则需要对所有的规则参数进行全局更新, 其中, 采用了模型更新的差异性来挑选用于全局更新的样本, 以进一步提高模型的精准度。

首先, 为了衡量模型对已到达在线样本的预测精度, 设计累积误差指标如下:

$$E_t = \sum_{i=1}^t \lambda^{t-i} |e_i(t)|, \quad (33)$$

其中: E_t 为 t 时刻计算得到的累积误差指标, λ 为遗忘因子, $e_i(t)$ 为 t 时刻第 i 个在线样本的误差。在累积误差的计算中, 遗忘因子 $\lambda \in (0, 1)$, 其作用是对于不同时刻的在线样本赋予不同的权重, 越接近当前时刻的样本权重越大, 从而使得在计算累积误差时能够更侧重于最新样本。为了平衡新样本与历史样本在累积误

差计算中所占的比重,文中取 $\lambda = 0.5$.

然后,根据模型对在线样本的预测精度设置自适应激活强度阈值,用于实现分级更新.该激活强度阈值与模型预测精度应具有相反变化趋势,当模型预测精度降低时,即模型对在线样本累积误差增大时,激活强度阈值也应增大,使得模型倾向于全局更新.因此,文中选择了一个常用的单增函数 $f(x) = \frac{1}{1 + \exp(-x)}$ 来刻画激活强度阈值 η_t 与累积误差指标 E_t 的关系.同时,为了使得该函数变化范围与激活强度实际值相符,需要为该函数设置一个上限值,即激活强度阈值最大值.因此,最终所构造的激活强度阈值变化函数为

$$\eta_t = \eta_{\max} \left(\frac{2}{1 + \exp(-E_t)} - 1 \right), \quad (34)$$

其中 η_{\max} 为激活强度阈值最大值.

对于 t 时刻在线样本 $\{\mathbf{x}(t), y_d(t)\}$,假设经过规则增长与修剪后网络中的规则个数为 K ,若 $\exists j \in \{1, 2, \dots, K\}$, $FS_j(\mathbf{x}(t)) \geq \eta_t$,则对于所有满足条件的 K^* 个规则,采用如下RLS算法更新其后件参数:

$$\mathbf{w}^*(t) = \mathbf{w}^*(t-1) + \mathbf{P}(t-1) \mathbf{h}^*(t)^T (y_d(t) - y(t)), \quad (35)$$

$$\mathbf{P}(t) = \mathbf{P}(t-1) - \frac{\mathbf{P}(t-1) \mathbf{h}^*(t)^T \mathbf{h}^*(t) \mathbf{P}(t-1)}{1 + \mathbf{h}^*(t) \mathbf{P}(t-1) \mathbf{h}^*(t)^T}, \quad (36)$$

其中: $\mathbf{w}^*(t) = [w_1(t) \dots w_{j^*}(t) \dots w_{K^*}(t)]^T$ 为所选择出的 K^* 个规则对应的权值, $\mathbf{h}^*(t) = [h_1(t) \dots h_{j^*}(t) \dots h_{K^*}(t)]$ 为规则化层输出, $y(t)$ 为 t 时刻在线样本的网络输出, $\mathbf{P}(t)$ 为协方差矩阵,且 $\mathbf{P}(0) = \alpha \mathbf{I}$, α 为一个取值较大的正数,文中取 $\alpha = 10^4$.

对于 t 时刻在线样本 $\{\mathbf{x}(t), y_d(t)\}$,若 $\forall j \in \{1, 2, \dots, K\}$, $FS_j(\mathbf{x}(t)) < \eta_t$,则采用ISO算法对所有规则进行全局更新.为了能够确定一组合适的样本用于全局更新,引入了如下指标量化模型更新前后的差异性^[26]:

$$D(\mathcal{M}(t-1), \mathcal{M}(t)) = \sqrt{\sum_{j=1}^K (w_j(t-1) - w_j(t))^2}, \quad (37)$$

其中: $\mathcal{M}(t-1)$ 代表 $t-1$ 时刻更新后的模型, $\mathcal{M}(t)$ 代表 t 时刻更新后的模型. $D(\mathcal{M}(t-1), \mathcal{M}(t))$ 越大,意味着 $t-1$ 时刻模型难以适用于 t 时刻样本的可能性越大.在整个更新过程中,采用一个滑动窗口记录每次更新后模型的差异性,在需要进行全局更新时,找到滑动窗口内最大值所对应的样本 $\{\mathbf{x}(t_{D_{\max}}), y_d(t_{D_{\max}})\}$,从该样本开始直至当前最新的样本 $\{\mathbf{x}(t), y_d(t)\}$ 被用于全局更新,更新公式如式(24)–(32)所示.

因此,分级更新策略的具体实施步骤如下:

步骤1 初始时刻 $t = 1$,根据式(33)计算累积误差指标 E_t ,根据式(34)确定激活强度阈值 η_t .

步骤2 计算网络中所有规则的激活强度,若存在规则激活强度大于或等于 η_t ,则进入步骤3;否则,转向步骤4.

步骤3 根据式(35)–(36)对激活强度大于或等于阈值的规则后件参数进行更新,转向步骤5.

步骤4 按照式(37)计算滑动窗口内的差异性,并找到差异性最大的样本 $\{\mathbf{x}(t_{D_{\max}}), y_d(t_{D_{\max}})\}$,基于 $\{\mathbf{x}(i), y_d(i) | i = t_{D_{\max}}, \dots, t\}$ 根据式(24)–(32)对所有规则前后件参数进行更新.

步骤5 根据式(37)计算差异性并更新滑动窗口内的数据.

步骤6 若基于所有在线数据更新完毕,则在线更新完成;否则, $t = t + 1$ 并返回步骤1.

根据上述所提出的辅助变量选择方法以及DFNN的设计思路,给出基于DFNN的软测量方法流程图如图3所示.

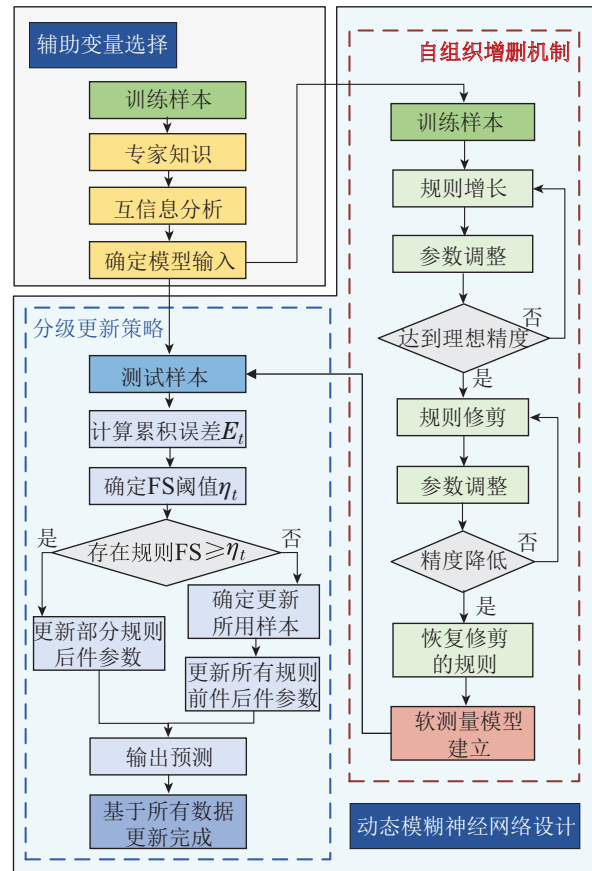


图3 基于DFNN的软测量方法流程图

Fig. 3 Flowchart for soft-sensing method based on DFNN

4 仿真实验与结果分析

为了验证所提出软测量方法的有效性,文中基于

BSM1 阴雨天气下的数据进行实验验证. 该数据集涵盖了 14 天内的数据, 共 1344 组, 其中从第 9 天至第 11 天模拟了降雨的发生. 文中选取第 1 至 7 天的数据作为训练集, 第 8 至 14 天的数据作为测试集, 所有数据在实验前均进行了归一化. 由于降雨的发生导致污水处理过程处于非平稳状态, 在此期间数据分布也将相应发生变化, 进而使得测试集中将有一段数据分布明显区别于训练集, 相应地将会影响模型在测试集上的预测效果. 结合第 3.1 节中的辅助变量选择结果, 采用 DFNN 分别建立出水 $\text{NH}_4^+\text{-N}$ 和出水 TN 软测量模型, 测试效果如图 4-5 所示. 为了验证所提出软测量方法的有效性, 将所提出方法与基于其它经典算法的软测量方法进行了比较, 如极限学习机(extreme learning machine, ELM)^[27]、支持向量回归(support vector regression, SVR)^[28]. 此外, 还比较了两种固定结构 FNN, 分别为基于梯度下降算法调整参数的 FNN¹ 和基于 ISO 算法调整参数的 FNN². 采用 RMSE、平均绝对百分比误差(mean absolute percentage error, MAPE)、决定系数(coefficient of determination, R^2) 对不同方法的性能进行评价, RMSE 的计算公式如式(20)所示, MAPE 与 R^2 计算公式为

$$\text{MAPE} = \frac{1}{P} \sum_{p=1}^P \left| \frac{y_p - y_{dp}}{y_{dp}} \right| \times 100\%, \quad (38)$$

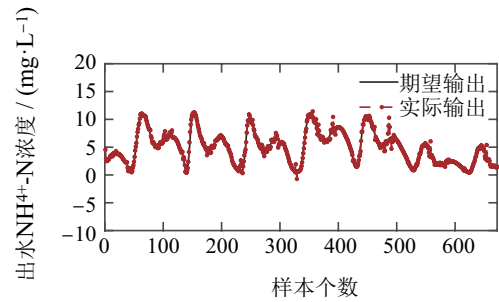
$$R^2 = 1 - \frac{\sum_{p=1}^P (y_p - y_{dp})^2}{\sum_{p=1}^P (y_p - \bar{y}_d)^2}, \quad (39)$$

其中: P 为样本个数, y_p 是第 p 个样本的模型输出值, y_{dp} 是第 p 个样本的期望输出值, \bar{y}_d 是所有样本期望输出值的平均.

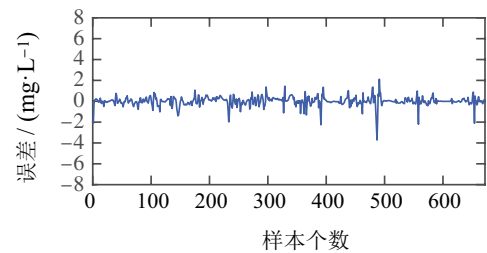
由图 4-5 中的结果可以看出, 所提出方法的预测曲线能够较好地跟踪期望输出值. 基于 DFNN 所建立的出水 $\text{NH}_4^+\text{-N}$ 软测量模型测试 RMSE, MAPE, R^2 分别为 0.4515, 8.4314%, 0.9747; 基于 DFNN 所建立的出水 TN 软测量模型测试 RMSE, MAPE, R^2 分别为 0.3513, 1.5499%, 0.9861, 表明所提出的软测量方法对出水含氮参数具有理想的测量精度.

表 3-4 给出了不同软测量模型对出水含氮参数的预测结果对比. 对于出水 $\text{NH}_4^+\text{-N}$ 来说, 虽然 ELM, SVR 和 FNNs 在训练集上的各个性能指标均在可接受范围内, 但基于上述几种方法建立的模型在测试集上精度较差, 这是由于在降雨发生的情况下, 上述几种模型无法准确跟踪期望输出值. 而文中所提方法在测试集上的各个性能指标明显优于其他 4 种算法, 这是由于 DFNN 引入了分级更新策略, 因此, 基于 DFNN 的软测量模型在复杂环境中仍然具有较高的准确性,

有效提高了对出水 $\text{NH}_4^+\text{-N}$ 的预测精度. 在出水 TN 的软测量实验中, 尽管 ELM 在训练集上的预测效果具有一定优势, 但是该方法难以在测试集上取得理想的精度. 文中所提出的 DFNN 在预测精度方面相较于其他几种方法具有明显优势. 此外, 从出水 $\text{NH}_4^+\text{-N}$ 、出水 TN 的预测结果中均可以看出 DFNN 相较于固定结构 FNN 具有更好的性能, 表明 DFNN 的自组织增删机制有助于进一步提高网络学习精度.



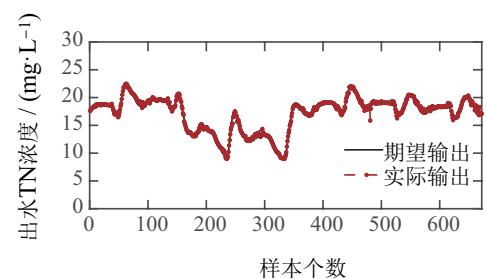
(a) 测试结果



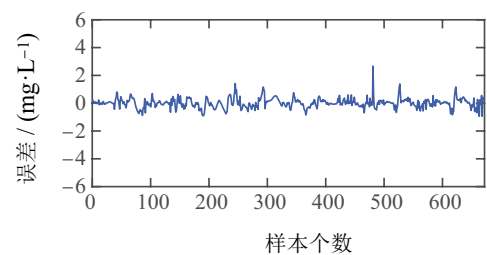
(b) 测试误差

图 4 出水 $\text{NH}_4^+\text{-N}$ 软测量模型测试结果

Fig. 4 The testing results of effluent $\text{NH}_4^+\text{-N}$ soft-sensing model



(a) 测试结果



(b) 测试误差

图 5 出水 TN 软测量模型测试结果

Fig. 5 The testing results of effluent TN soft-sensing model

表3 不同软测量模型出水NH₄⁺-N预测结果对比Table 3 Prediction results of various effluent NH₄⁺-N soft-sensing models

算法	训练集			测试集		
	RMSE	MAPE/%	R ²	RMSE	MAPE/%	R ²
ELM	1.2558	23.0898	0.8844	6.1211	91.9195	0.2513
SVR	1.4378	13.2438	0.8809	8.4265	132.6207	0.1148
FNN ¹	1.3804	19.7745	0.8765	2.896	87.6671	0.1933
FNN ²	1.2874	20.9566	0.8814	2.5652	59.7651	0.4785
DFNN	1.1900	18.9821	0.8905	0.4515	8.4314	0.9747

表4 不同软测量模型出水TN预测结果对比

Table 4 Prediction results of various effluent TN soft-sensing models

算法	训练集			测试集		
	RMSE	MAPE/%	R ²	RMSE	MAPE/%	R ²
ELM	0.9877	3.4476	0.8909	12.7241	76.8425	0.2060
SVR	1.5523	4.1757	0.7610	16.8877	51.0711	0.0590
FNN ¹	1.3841	5.6507	0.6582	2.7341	12.9924	0.2235
FNN ²	1.2937	5.5419	0.6803	3.0151	18.4441	0.3431
DFNN	1.2177	5.2540	0.6839	0.3513	1.5499	0.9861

为了进一步验证分级更新策略的有效性,在不考虑更新以及考虑3种不同更新方法的情况下,分别建立出水含氮参数软测量模型进行实验对比,3种更新方法具体描述如下:方法1:对所有规则后件参数采用RLS算法进行更新;方法2:对激活强度大于固定阈值的规则后件参数采用RLS算法进行更新;方法3:采用文中所提出的分级更新策略进行更新.在出水NH₄⁺-N软测量实验中,方法2中固定阈值设置为0.02,方法3中阈值最大值设置为2,滑动窗口长度设置为50;在出水TN软测量实验中,方法2中固定阈值设置为0.02,方法3中阈值最大值设置为3,滑动窗口长度设置为50.结合不同更新方法的出水NH₄⁺-N、出水TN软测量模型预测结果分别如图6-7所示.

由图6-7可以看出,与不考虑更新相比较,加入了3种更新方法的出水含氮参数软测量模型输出均能够更好地拟合期望输出值的变化趋势.因此,在进行出水含氮参数软测量时,引入更新方法能够有效提高预测精度.其中,方法3所代表的分级更新策略对于期望输出值拟合效果优于其它两种更新方法.表5-6给出了不同更新方法下出水NH₄⁺-N、出水TN软测量模型的预测结果对比,可以看出相比较于不考虑更新,3种更新方法的各个性能指标均有不同程度的提升.其中,方法3在各个指标上均能够取得优于其他两种方法的结果,在预测精度方面表现更佳,进一步验证了所提出软测量方法的有效性.

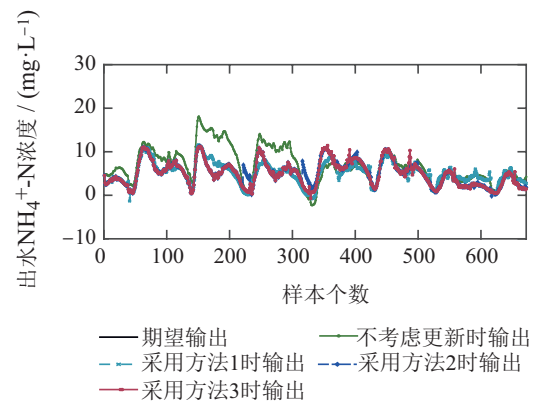


图6 不同更新方法下出水NH₄⁺-N软测量模型预测结果
Fig. 6 Prediction results of the effluent NH₄⁺-N soft-sensing models with different updating methods

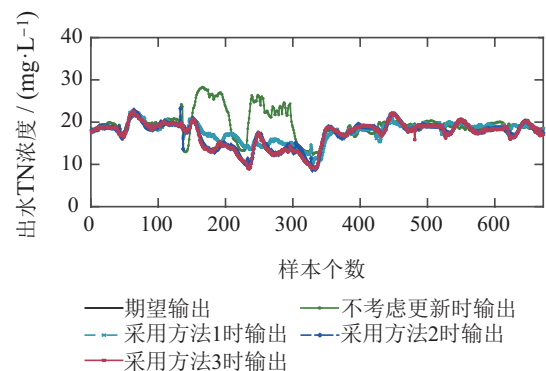


图7 不同更新方法下出水TN软测量模型预测结果
Fig. 7 Prediction results of the effluent TN soft-sensing models with different updating methods

表 5 不同更新方法下出水 $\text{NH}_4^+\text{-N}$ 软测量模型预测结果对比Table 5 Prediction results of the effluent $\text{NH}_4^+\text{-N}$ soft-sensing models with different updating methods

更新方法	RMSE	MAPE/%	R^2
—	3.2542	66.3360	0.4246
方法1	1.3910	36.4632	0.6888
方法2	0.7714	18.5118	0.9186
方法3	0.4515	8.4314	0.9747

表 6 不同更新方法下出水TN软测量模型预测结果对比

Table 6 Prediction results of the effluent TN soft-sensing models with different updating methods

更新方法	RMSE	MAPE/%	R^2
—	4.7892	20.5339	0.3172
方法1	1.6512	8.1877	0.4905
方法2	0.4328	1.6776	0.9774
方法3	0.3513	1.5499	0.9861

5 结论

针对城市WWTP出水 $\text{NH}_4^+\text{-N}$ 、出水TN难以实时准确检测的问题,文中提出了一种基于DFNN的含氮参数软测量方法.该方法具有以下特点:

1) 采用一种自组织增删机制与快速二阶学习算法构建FNN,保证了软测量模型具有精简的结构与良好的精度;

2) 引入自适应激活强度阈值,设计了FNN分级更新策略,提高了模型在非平稳动态环境中的自适应性;

3) 基于所提出的软测量方法,实现了对于出水含氮参数 $\text{NH}_4^+\text{-N}$ 、TN的在线准确测量.

为了能够建立更加准确、可靠的模型以进一步提高对出水含氮参数的预测精度,未来工作将致力于挖掘数据中蕴含的时延信息.

参考文献:

- [1] ZHANG Shuai, ZHOU Ping. Recursive bilinear subspace modeling and model-free adaptive control of wastewater treatment. *Acta Automatica Sinica*, 2022, 48(7): 1747 – 1759.
(张帅, 周平. 污水处理过程递推双线性子空间建模及无模型自适应控制. *自动化学报*, 2022, 48(7): 1747 – 1759.)
- [2] WANG Wei, YANG Chunhua, HAN Jie, et al. Forward variable selection method based on k-nearest neighbor mutual information and its application in soft sensor modeling of water quality parameters. *Systems Engineering – Theory & Practices*, 2022, 42(1): 253 – 261.
(王威, 阳春华, 韩洁, 等. 基于k-近邻互信息的前向式变量选择方法及在水质参数软测量中的应用. *系统工程理论与实践*, 2022, 42(1): 253 – 261.)
- [3] ZHOU P, WANG X, CHAI T. Multiobjective operation optimization of wastewater treatment process based on reinforcement self-learning and knowledge guidance. *IEEE Transactions on Cybernetics*, 2022, 53(11): 6896 – 6909.
- [4] GERNAEY K, VAN L M, HENZE M, et al. Activated sludge wastewater treatment plant modelling and simulation: state of the art. *Environmental Modelling & Software*, 2004, 19(9): 763 – 783.
- [5] ZHOU M, GONG J, YANG C, et al. Simulation of the performance of aerobic granular sludge SBR using modified ASM3 model. *Bioresource Technology*, 2012, 127: 473 – 481.
- [6] LUO T, YANG M, HAN J, et al. A novel model-based adaptive control strategy for step-feed SBRs dealing with influent fluctuation. *Bioresource Technology*, 2014, 167: 476 – 483.
- [7] GUO R, LIU H, XIE G, et al. A self-interpretable soft sensor based on deep learning and multiple attention mechanism: From data selection to sensor modeling. *IEEE Transactions on Industrial Informatics*, 2022, 19(5): 6859 – 6871.
- [8] TAOUFIK N, BOUMYA W, ACHAK M, et al. The state of art on the prediction of efficiency and modeling of the processes of pollutants removal based on machine learning. *Science of the Total Environment*, 2022, 807: 150554.
- [9] WEN Chaoyao, ZHOU Ping. Sparse robust modeling of effluent quality indices in wastewater treatment process. *Acta Automatica Sinica*, 2022, 48(6): 1469 – 1481.
(闻超垚, 周平. 污水处理过程出水水质稀疏鲁棒建模. *自动化学报*, 2022, 48(6): 1469 – 1481.)
- [10] DE CANETE J F, DEL SAZ-OROZCO P, GÓMEZ-DE-GABRIEL J, et al. Control and soft sensing strategies for a wastewater treatment plant using a neuro-genetic approach. *Computers & Chemical Engineering*, 2021, 144: 107146.
- [11] JIA Lijie, LI Wenjing, QIAO Junfei. Self-organizing design of radial basis function neural network based on neuron characteristics. *Control Theory & Applications*, 2020, 37(12): 2618 – 2626.
(贾丽杰, 李文静, 乔俊飞. 基于神经元特性的径向基函数神经网络自组织设计方法. *控制理论与应用*, 2020, 37(12): 2618 – 2626.)
- [12] DING Haixu, TANG Jian, XIA Heng, et al. Modeling of MIMO controlled object in municipal solid waste incineration process based on TS-FNN. *Control Theory & Applications*, 2022, 39(8): 1529 – 1540.
(丁海旭, 汤健, 夏恒, 等. 基于TS-FNN的城市固废焚烧过程MIMO被控对象建模. *控制理论与应用*, 2022, 39(8): 1529 – 1540.)
- [13] ZAGHLOUL M S, HAMZA R A, IORHEMEN O T, et al. Comparison of adaptive neuro-fuzzy inference systems (ANFIS) and support vector regression (SVR) for data-driven modelling of aerobic granular sludge reactors. *Journal of Environmental Chemical Engineering*, 2020, 8(3): 103742.
- [14] QIAO J, QUAN L, YANG C. Design of modeling error PDF based fuzzy neural network for effluent ammonia nitrogen prediction. *Applied Soft Computing*, 2020, 91: 106239.
- [15] XIONG W, SHI X. Soft sensor modeling with a selective updating strategy for Gaussian process regression based on probabilistic principle component analysis. *Journal of the Franklin Institute*, 2018, 355(12): 5336 – 5349.
- [16] WU J, CHENG H, LIU Y, et al. Learning soft sensors using time difference-based multi-kernel relevance vector machine with applications for quality-relevant monitoring in wastewater treatment. *Environmental Science and Pollution Research*, 2020, 27(23): 28986 – 28999.

- [17] ALHARBI M, HONG P, LALEG-KIRATI T. Sliding window neural network based sensing of bacteria in wastewater treatment plants. *Journal of Process Control*, 2022, 110: 35 – 44.
- [18] FAN Y, GUO Z, WANG J, et al. Soft-sensor design via task transferred just-in-time-learning coupled transductive moving window learner. *Environmental Research*, 2022, 210: 113015.
- [19] LU Rongxiu, RAO Yunchun, YANG Hui, et al. Prediction of Pr/Nd component content based on improved just-in-time learning algorithm. *Control Theory & Applications*, 2020, 37(8): 1846 – 1854. (陆荣秀, 饶运春, 杨辉, 等. 基于改进即时学习算法的镨/钕元素组分含量预测. *控制理论与应用*, 2020, 37(8): 1846 – 1854.)
- [20] YEH J, SU S. Efficient approach for RLS type learning in TSK neural fuzzy systems. *IEEE Transactions on Cybernetics*, 2016, 47(9): 2343 – 2352.
- [21] GU X, SHEN Q. A self-adaptive fuzzy learning system for streaming data prediction. *Information Sciences*, 2021, 579: 623 – 647.
- [22] GUJER W, HENZE M, MINO T, et al. Activated sludge model No. 3. *Water Science & Technology*, 1999, 39(1): 183 – 193.
- [23] YU H, REINER P D, XIE T, et al. An incremental design of radial basis function networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 25(10): 1793 – 1803.
- [24] AHWIADI M, WANG W. An adaptive evolving fuzzy technique for prognosis of dynamic systems. *IEEE Transactions on Fuzzy Systems*, 2021, 30(3): 841 – 849.
- [25] WILAMOWSKI B M, YU H. Improved computation for Levenberg-Marquardt training. *IEEE Transactions on Neural Networks*, 2010, 21(6): 930 – 937.
- [26] YANG Z, AL-DAHIDI S, BARALDI P, et al. A novel concept drift detection method for incremental learning in nonstationary environments. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 31(1): 309 – 320.
- [27] HUANG G, ZHU Q, SIEW C. Extreme learning machine: Theory and applications. *Neurocomputing*, 2006, 70(1/3): 489 – 501.
- [28] VAPNIK V N. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 1999, 10(5): 988 – 999.

作者简介:

蒙西 副教授, 硕士生导师, 目前研究方向为神经网络分析与设计、复杂工业过程智能建模, E-mail: mengxi@bjut.edu.cn;

张寅 硕士研究生, 目前研究方向为污水处理过程智能建模, E-mail: zhangyin0028@emails.bjut.edu.cn;

乔俊飞 教授, 博士生导师, 目前研究方向为计算智能与智能优化控制、复杂工业过程智能建模与优化控制, E-mail: adqiao@bjut.edu.cn.