

带有工况中心修正的多模型在线建模

丛秋梅^{1,2†}, 苑明哲¹, 柴天佑^{3,4}, 王宏^{1,5}

(1. 中国科学院沈阳自动化研究所 信息服务与智能控制技术研究室, 辽宁 沈阳 110016;

2. 辽宁石油化工大学 信息与控制工程学院, 辽宁 抚顺 113001; 3. 东北大学 自动化研究中心, 辽宁 沈阳 110819;

4. 东北大学 流程工业综合自动化国家重点实验室, 辽宁 沈阳 110819; 5. 沈阳中科博微自动化有限公司, 辽宁 沈阳 110179)

摘要: 针对运行工况频繁波动、单一模型难以描述过程特性的问题, 提出了带有工况中心修正的多模型在线建模方案, 包括工况识别机制、局部模型、多模型合成机制. 工况识别机制根据工况特征变量分析工况范围, 由相近度修正工况中心; 局部模型采用Hammerstein模型, 非线性增益由带有稳定学习算法的小波神经网络建立, 线性模型由带控制量的自回归模型(ARX)建立; 多模型合成机制采用加权求和方法. 在线修正工况中心可反映工况的时间变化特性, 参数稳定学习算法改善了模型精度和自适应能力. 采用此方法建立污水处理过程化学需氧量(COD)软测量模型, 结果表明, 模型在工况大范围变化时仍具有满意预测效果.

关键词: 工况; 多模型; 建模; 稳定学习; 污水处理过程

中图分类号: TP391.9 **文献标识码:** A

Online modeling for multi-model by adjusting the centers of operating ranges

CONG Qiu-mei^{1,2†}, YUAN Ming-zhe¹, CHAI Tian-you^{3,4}, WANG Hong^{1,5}

(1. Department of Information Service & Intelligent Control, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang Liaoning 110016, China;

2. School of Information and Control Engineering, Liaoning Shihua University, Fushun Liaoning 113001, China;

3. Research Center of Automation, Northeastern University, Shenyang Liaoning 110819, China;

4. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang Liaoning 110819, China;

5. Microcyber Inc, Shenyang Liaoning 110179, China)

Abstract: Because a single model cannot represent the characteristics of the complex industrial process in varying operating ranges, we propose an online modeling scheme for multiple models. This scheme includes the recognition mechanism of operating range, the local models and the combination mechanism for multiple models. The recognition mechanism analyzes the operating range according to its characteristic variables and adjusts the center of operating range according to similarity degrees. The local model is actually a Hammerstein model which is the serial connection of a wavelet neural network with a stable learning algorithm and an ARX model. The combination mechanism calculates the weighted sum of the outputs of local models, and online adjusts the centers of operating range to reflect the variation characteristics of the operating range. A stable learning algorithm of parameters improves the prediction accuracy and the adaptation ability. This method is implemented in a wastewater treatment process to measure the concentration of the chemical oxygen demand (COD). Experimental results show that this modeling scheme can obtain satisfactory effect in varying operating ranges.

Key words: operating range; multiple models; modeling; stable learning; wastewater treatment process

1 引言(Introduction)

大多数复杂工业过程受多种外部因素的影响, 运行工况经常处于大范围波动的情况下, 单一模型难以描述整个工况的过程特性. 由于软测量模型精度在很大程度上依赖于运行工况^[1], 一旦工况变化, 软

测量精度迅速降低, 基于分解-合成法则的多模型建模策略在一定程度上克服上述不足^[2-3]. 文献[4]提出了稀土串级萃取分离过程的多模型软测量模型, 在多个工作点附近建立局部线性模型, 根据性能指标选择最优局部模型作为元素组分含量软测量模

收稿日期: 2012-06-19; 收修改稿日期: 2013-01-30.

†通信作者. E-mail: cong_0828@163.com.

基金项目: 国家重点基础研究发展计划资助项目(2009CB320601); 创新引智计划(“111”计划)资助项目(B08015); 国家自然科学基金资助项目(61034008, 61004051); 中国博士后科学基金面上资助项目(2013M530953).

型,对于工况重叠区域的运行数据预测精度不高.文献[5]基于证据演化Gustafson-Kessel算法研究了非线性动态系统多模型建模的结构确定和参数辨识问题,采用信任函数理论表征了不同工况区域之间过渡的不确定性,估计了局部子模型的激活度,但仅修改最近邻工况中心的方法无法全面描述各工况区域的动态趋势.对于污水处理过程,文献[6]基于多模型建模方法,建立了污水处理厂输入输出变量数据的主元分析得分向量和目标水质COD之间关系的模型;文献[7]采用模糊C均值聚类算法识别操作工况,在不同工况下以神经网络偏最小二乘法分别建立非线性子模型,建立了某污水处理厂水质COD、悬浮固体浓度、氨态氮软测量模型.文献[6]和文献[7]均未考虑新样本使运行工况呈现实时波动的情況.

Hammerstein模型(简称H模型)是化工过程的常用模型之一,由无记忆非线性增益和有记忆线性系统串联构成,线性子系统描述对象动态特性,非线性增益校正线性系统模型.文献[8]以多项式描述H模型非线性增益,仅能代表弱非线性过程,对于具有中等或强非线性的过程,精度和适用性将有所下降.本文提出了带有工况中心修正的多模型在线建模方法,由工况识别机制分析工况,以新样本与

工况中心之间的相近度在线修正工况中心,有效地反映了工况的动态趋势;局部模型采用带有时变稳定学习算法的小波神经网络(wavelet neural network, WNN)和ARX模型串联的H模型,与固定学习速率相比,时变学习率具有更快的收敛性^[9],保证了误差的稳定性和自适应能力, WNN使模型对强非线性过程具有较好的拟合能力;多模型切换机制采用加权求和方法,实现了局部模型之间的平滑切换,较好地刻画工况重叠区域的过程特性.采用此方法建立水质COD软测量模型,结果表明模型在运行工况大范围波动时仍具有满意动态预测效果.

2 多模型在线建模策略(On-line modeling strategy based on multi-model)

本文多模型在线软测量建模方法由工况识别机制、局部模型、多模型合成机制组成,如图1所示.其中: $\Omega(k)$ 表示工况特征变量; $x(k)$ 表示其他输入变量; $O(k)$ 表示中间变量; $c_i(k)$ 表示第*i*工况中心; $\hat{y}_i(k)$ 表示第*i*局部模型 M_i 的输出; $\hat{y}(k)$ 表示输出估计值; $y(k)$ 表示输出实际值; $\lambda_i(k)$ 表示 $\Omega(k)$ 属于第*i*工况的隶属度, $i = 1, \dots, m$, m 为局部模型个数,即工况个数; $e(k)$ 表示建模误差, $e(k) = \hat{y}(k) - y(k)$, $k = 1, \dots, N$, N 表示样本数.

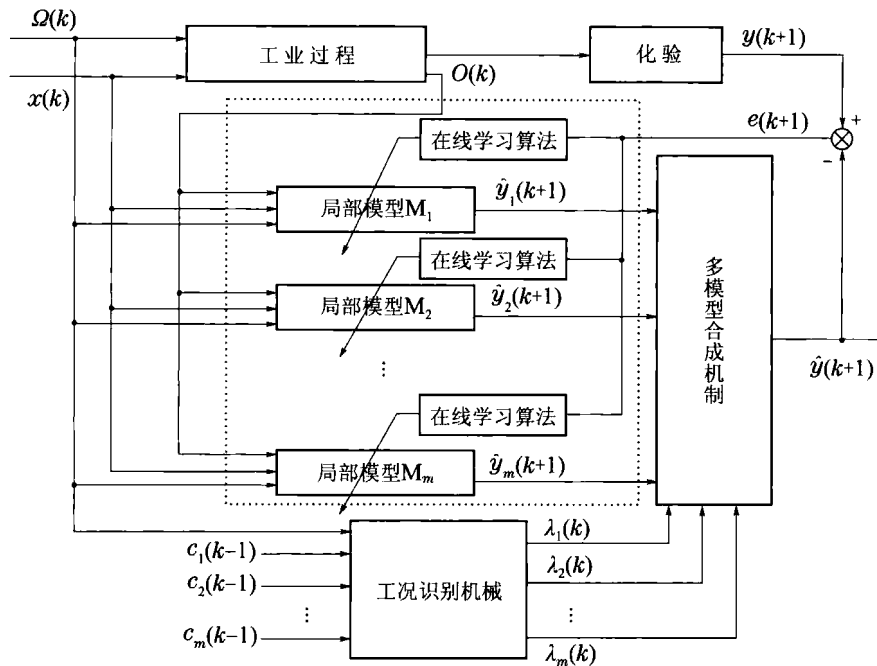


图1 多模型在线软测量模型结构

Fig. 1 The structure of on-line soft sensor based on multi-model

2.1 工况识别机制(The recognition mechanism of operating range)

工况识别机制包括工况中心在线修正算法和隶属度在线计算算法.工况中心在线修正算法计

算 $\Omega(k)$ (反映当前运行工况)与 $c_i(k-1)$ 之间的相近度,采用修正算法实时修正工况中心;隶属度在线计算算法计算当前 $\Omega(k)$ 属于各工况区域的隶属度 $\lambda_i(k)$.图2为工况识别机制结构图.

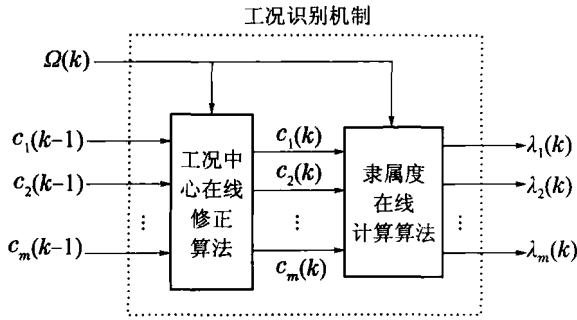


图2 工况识别机制结构

Fig. 2 The structure of operating regime recognition mechanism

1) 工况中心在线修正算法.

当增加样本 $\Omega(k)$ 时, 计算 $\Omega(k)$ 与 $c_i(k-1)$ 之间的相相似度函数 S_{ki} . S_{ki} 定义如下:

$$S_{ki} = e^{-\alpha \|\Omega(k) - c_i(k-1)\|}, \quad (1)$$

其中: α 表示相相似度修正系数, $\alpha = -\ln(0.5/\bar{D})$, \bar{D} 表示数据集平均距离^[10], 计算如下:

$$\bar{D} = \frac{\sum_{i=1}^{N_1} \sum_{j=1}^{N_1} \|\Omega(i) - \Omega(j)\|}{N_1(N_1 - 1)}, \quad (2)$$

其中 N_1 表示训练样本数. S_{ki} 取值范围介于(0,1]之间, S_{ki} 值越大, 表明 $\Omega(k)$ 与 $c_i(k-1)$ 之间距离越近, 样本属于此工况的程度越大. 基于相相似度函数的工况中心在线修正算法为

$$c_i(k) = c_i(k-1) + \sigma S_{ki}[\Omega(k) - c_i(k-1)], \quad (3)$$

其中 σ 表示工况中心修正率, $\sigma = 0.05$. σ 与工况个数有关, 当工况个数较多时, 工况并集接近于整个输入空间, σ 可取较小的值而不影响推理结果.

2) 隶属度在线计算算法.

修正各工况中心位置以后, 需要计算 $\Omega(k)$ 属于第 i 工况的隶属度 $\lambda_i(k)$, 作为局部模型 M_i 输出的权值, 表征当前工况与各工况匹配的程度. $\lambda_i(k)$ 由如下公式计算:

$$\lambda_i(k) = \frac{1}{\sum_{j=1}^m \left(\frac{L_{ki}^2}{L_{kj}^2} \right)}, \quad (4)$$

其中 $L_{ki}^2 = (\Omega(k) - c_i(k))^T (\Omega(k) - c_i(k))$.

2.2 局部H模型(Local H model)

H模型采用WNN建立非线性增益模型, 与多项式相比, WNN非线性拟合能力更强, 可以更准确地描述非线性过程特性. 局部模型 M_1 结构如图3所示.

模型 M_1 中WNN模型可写为

$$u_1(k) = W_1(k)\phi_1(V_1(k)\mathbf{x}(k)), \quad (5)$$

M_1 可写为

$$\hat{y}_1(k) = -\sum_{i=1}^{n_a} a_{1i} \hat{y}_1(k-i) + \sum_{j=0}^{n_b} b_{1j} u_1(k-j). \quad (6)$$

将式(5)代入式(6):

$$\hat{y}_1(k) = -\sum_{i=1}^{n_a} a_{1i} \hat{y}_1(k-i) + \sum_{j=0}^{n_b} b_{1j} (W_1(k-j)\phi_1(V_1(k-j)\mathbf{x}(k-j))). \quad (7)$$

其中: $\mathbf{x}(k) \in \mathbb{R}^I$ 表示局部模型输入向量; $u_1(k)$ 表示 M_1 中WNN输出; $\hat{y}_1(k)$ 表示ARX模型输出, 即 M_1 的输出; $a_{1i}(k)$ ($i = 1, 2, \dots, n_a$)和 $b_{1j}(k)$ ($j = 1, 2, \dots, n_b$)表示ARX模型参数, n_a 和 n_b 表示模型阶次; $\phi_1 = [\phi_{1i}] \in \mathbb{R}^H$ 表示WNN隐含层节点的激励函数向量, $\phi_{1i} = \psi\left(\frac{\sum_{j=1}^I v_{1ij}x_j - d_{1i}}{s_{1i}}\right)$, ψ 表示小波母函数, 其中 $S_1 = [s_{1i}(k)] \in \mathbb{R}^H$, $s_{1i}(k)$ 表示第 i 个隐含层节点小波基函数的伸缩参数; $D_1 = [d_{1i}(k)] \in \mathbb{R}^H$, $d_{1i}(k)$ 表示平移参数; $V_1(k) = [v_{1ij}(k)] \in \mathbb{R}^{H \times I}$ 表示隐含层权值矩阵, $W_1(k) = [w_{1i}(k)] \in \mathbb{R}^{1 \times H}$ 表示输出层权值向量; H 表示隐含层节点数, I 表示输入层节点数.

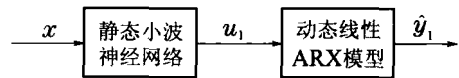


图3 局部H模型 M_1 结构

Fig. 3 The structure of local H model M_1

1) 非线性增益WNN参数辨识.

定理 1 为消除未建模动态和干扰影响, 若WNN参数采用如下稳定学习算法时:

$$W_1(k+1) = W_1(k) - \eta_1(k)\lambda_1(k)b_{10}(k)e(k)\phi_1(k), \quad (8)$$

$$V_1(k+1) = V_1(k) - \eta_1(k)\lambda_1(k)b_{10}(k)e(k) \cdot \phi_1'(k)W_1(k)\mathbf{x}^T(k), \quad (9)$$

$$s_{1i}(k+1) = s_{1i}(k) + \eta_1(k)\lambda_1(k)b_{10}(k)e(k) \cdot \frac{w_{1i}(k)}{s_{1i}^2(k)} \phi_1'(k) \left(\sum_{j=1}^I v_{1ij}(k)x_j(k) - d_{1i}(k) \right), \quad (10)$$

$$d_{1i}(k+1) = d_{1i}(k) + \eta_1(k)\lambda_1(k)b_{10}(k)e(k) \frac{w_{1i}(k)}{s_{1i}(k)} \phi_1'(k). \quad (11)$$

学习率:

$$\eta_1(k) = \frac{\eta_0}{1 + \Phi(k)}, \quad 0 < \eta_0 \leq 1,$$

$$\begin{aligned} \Phi(k) = & \|\lambda_1(k)b_{10}(k)\phi_1(k)\|^2 + \\ & \|\lambda_1(k)b_{10}(k)W_{V_1}(k)\mathbf{x}^T(k)\|^2 + \\ & \|\lambda_1(k)b_{10}(k)W_{S_1}(k)\|^2 + \\ & \|\lambda_1(k)b_{10}(k)W_{D_1}(k)\|^2, \end{aligned}$$

$$W_{V_1}(k) = \left[\frac{w_{11}}{s_{11}} \phi'_{11} \quad \cdots \quad \frac{w_{1H}}{s_{1H}} \phi'_{1H} \right]^T,$$

$$W_{S_1}(k) = \left[-\frac{w_{11}}{s_{11}^2} \phi'_{11} (\mathbf{E}_1 V_1 \mathbf{x} - d_{11}) \quad \cdots \right. \\ \left. -\frac{w_{1H}}{s_{1H}^2} \phi'_{1H} (\mathbf{E}_H V_1 \mathbf{x} - d_{1H}) \right],$$

$$\mathbf{E}_i = [0 \quad \cdots \quad 1 \quad \cdots \quad 0], \quad i = 1, 2, \dots, H,$$

$$W_{D_1}(k) = \left[-\frac{w_{11}}{s_{11}} \phi'_{11} \quad \cdots \quad -\frac{w_{1H}}{s_{1H}} \phi'_{1H} \right].$$

则可保证 M_1 的平均建模误差指标 \bar{J}_1 有界,并满足

$$\bar{J}_1 = \lim_{T \rightarrow \infty} \sup \frac{1}{T} \sum_{k=1}^T \|e(k)\|^2 \leq \frac{\eta_0}{\pi_1} \bar{\delta}_1, \quad (12)$$

其中: $\pi_1 = \frac{\eta(k)}{1 + \max_k \{\Phi(k)\}}$, M_1 的未建模动

态 $\bar{\delta}_1 = \max_k \{\|\delta_1(k)\|^2\}$.

因此多模型建模误差指标 \bar{J} 满足

$$\bar{J} \leq \max_i \{\bar{J}_i\}, \quad i = 1, 2, \dots, m. \quad (13)$$

证明详见附录.

2) 线性部分ARX模型参数辨识.

定义ARX模型参数向量 θ_1 和数据向量 φ_1 为

$$\theta_1 = [a_{11} \quad \cdots \quad a_{1n_a} \quad b_{10} \quad \cdots \quad b_{1n_b}]_{(n_a+n_b+1) \times 1}^T, \quad (14)$$

$$\varphi_1(k) =$$

$$\begin{bmatrix} -\hat{y}_1(k-1) & \cdots & -\hat{y}_1(k-n_a) \\ u_1(k) & \cdots & u_1(k-n_b) \end{bmatrix}_{(n_a+n_b+1) \times 1}^T. \quad (15)$$

θ_1 采用带有遗忘因子的递推最小二乘法辨识:

$$\theta_1(k+1) = \theta_1(k) + \mathbf{K}_1(k+1)e_1^T(k), \quad (16)$$

$$\mathbf{K}_1(k+1) = \frac{\mathbf{P}_1(k)\varphi_1(k+1)}{\varpi_1 + \varphi_1^T(k+1)\mathbf{K}_1(k)\varphi_1(k+1)}, \quad (17)$$

$$\mathbf{P}_1(k+1) = \frac{[\mathbf{I} - \mathbf{K}_1(k+1)\varphi_1^T(k+1)]\mathbf{P}_1(k)}{\varpi_1}, \quad (18)$$

其中遗忘因子 $\varpi_1 = 0.9$.

M_1 的建模误差 e_1 定义为

$$e_1(k) = \hat{y}_1(k) - y_1(k), \quad (19)$$

其中: $\hat{y}_1(k)$ 表示 M_1 输出; $y_1(k)$ 表示理想输出,实质上是理想虚拟输出,无法获知实际精确值.采用下式间接计算 $e_1(k)$:

$$\frac{\partial E(k)}{\partial \hat{y}_1(k)} = -e(k)\lambda_1(k-1) = -e_1(k), \quad (20)$$

其中误差性能指标 E 定义为

$$E(k) = \frac{1}{2} \|e(k)\|^2 = \frac{1}{2} (\hat{y}(k) - y(k))^2, \quad (21)$$

可得 $e_1(k) = e(k)\lambda_1(k-1)$.

采用类似方法推导 M_i 的参数辨识算法.

2.3 多模型合成机制(Multi-model combination mechanism)

由于大多数工业过程运行工况相互重叠,若仅选中与当前工况最接近的局部模型,无法刻画工况重叠区域的过程特性,存在较大软测量误差.采用基于加权求和的多模型合成机制不会造成输出的较大跳跃,更接近实际过程的运行情况.软测量模型输出由下式计算:

$$\hat{y}(k+1) = \sum_{i=1}^m \lambda_i(k) \hat{y}_i(k+1). \quad (22)$$

局部模型输出对应的权值表示当前工况与各工况匹配的程度,数值上采用 $\Omega(k)$ 属于各工况的隶属度 $\lambda_i(k)$.

3 多模型建模算法步骤(The procedure of on-line modeling algorithm based on multi-model)

1) 采用减法聚类算法对 Ω 历史数据进行聚类,得到初始工况中心位置 $\mathbf{c}_i(0)$ 及聚类数据集 Ξ_i ;

2) 利用数据集 Ξ_i ,离线学习 M_1 中WNN权值 W_1, V_1 和小波尺度参数 S_1, D_1 ,以及ARX模型参数 θ_1 ,作为 M_1 的初始参数;同时学习 M_i 的参数;

3) 采集新样本 $\Omega(k)$,由式(3)修正 $\mathbf{c}_i(k)$,由式(4)计算 $\lambda_i(k)$;

4) 由式(8)-(11)在线学习 $W_1(k+1), V_1(k+1)$ 和 $S_1(k+1), D_1(k+1)$,由式(16)-(18)在线学习 $\theta_1(k+1)$;以类似算法学习 M_i 的参数;

5) 由式(7)计算 M_1 的输出,采用类似方法计算 M_i 的输出,由式(22)计算模型输出,返回步骤3)计算下一时刻模型输出.

4 仿真实验(Simulation experiments)

4.1 工艺流程描述(Description of technological process)

由于受季节、天气、温度、居民习惯等因素影响,污水处理过程工况频繁波动,经常工作在负荷大范围波动条件下,生化反应速率和参数随工况波动而非线性变化,单一模型难以准确估计水质.由于COD难以连续在线检测,采用所提方法建立缺氧/好氧(A/O, Anoxic/Oxic)过程出水COD软测

量模型. 图4为A/O过程工艺流程图.

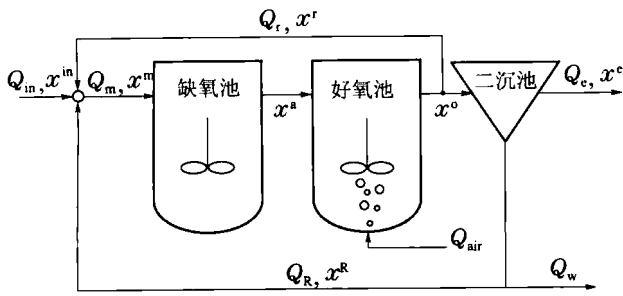


图4 A/O过程工艺流程图

Fig. 4 Flow chart of A/O wastewater treatment process

图中: Q_{in} 表示进水流量, Q_m 表示混合进水流量, Q_r 表示内回流混合液流量, Q_R 表示外回流污泥流量, Q_w 表示过剩污泥流量, Q_e 表示二沉池出水溢流流量, x^{in} 表示进水组分浓度, x^m 表示混合进水组分浓度, x^a 表示缺氧池出水组分浓度, x^o 和 x^r 表示好氧池出水组分浓度, x^R 表示二沉池浓缩区组分浓度, x^e 表示二沉池澄清区出水组分浓度.

4.2 数据描述(Description of data samples)

采用某污水处理厂250组输入/输出数据对, 前150组进行工况划分和离线建模, 后100组进行COD在线软测量实验. 模型输入变量选为进水流量 x_1 、COD x_2 、悬浮固体浓度 x_3 、氨态氮 x_4 、缺氧池氧化还原电位 x_5 、好氧池溶解氧浓度 x_6 .

定义 $\Omega = [x_1 \ x_2 \ x_3]^T$, 工况识别机制采用150组样本划分工况, 聚类半径选为 $\delta_a = 0.5$ 时, 将样本分为3个聚类子集, 聚类中心分别作为初始工况中心. 以进水COD为横轴, 流量为纵轴, 工况识别结果如图5所示, “•”表示工况中心位置, 1-3分别对应进水高负荷、正常负荷、低负荷. 可看出, 污水处理厂大多数情况下运行于高负荷和正常负荷之间.

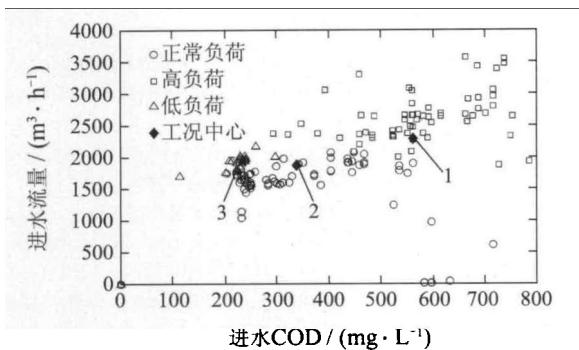


图5 工况识别结果

Fig. 5 The results of operating condition recognition

按经验选择WNN隐含层节点数 $H = 10$, 故WNN结构为6-10-1; 初始学习率取 $\eta_0 = 1$; 小波

母函数 ψ 选为Morlet小波, $\psi(x) = e^{-\frac{x^2}{2}} \cos(1.75x)$. 根据AIC准则选择ARX模型阶次为 $n_a = 2, n_b = 1$.

4.3 实验结果(Results of experiments)

1) 在线修正和不修正工况中心的水质多模型软测量结果与分析.

将离线得到的模型参数值作为初始值进行在线实验. 新数据会对污水处理过程运行工况产生影响, 表现为工况中心的修正. 图6为在线修正工况中心2的动态结果, 图中数字表示数据序号, 箭头表示工况中心随数据实时修正的动态趋势.

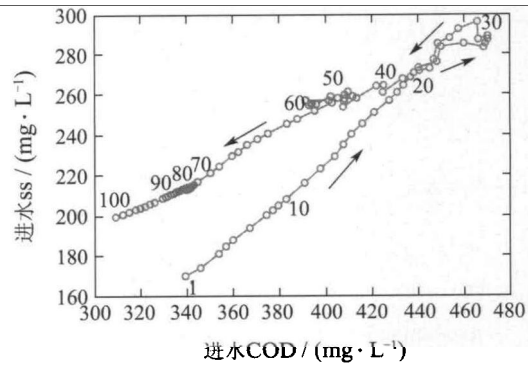


图6 工况中心2的在线修正结果

Fig. 6 The online adjusting result of the second operating center

图7为修正和不修正工况中心两种情况下, COD软测量结果与实际COD的比较.

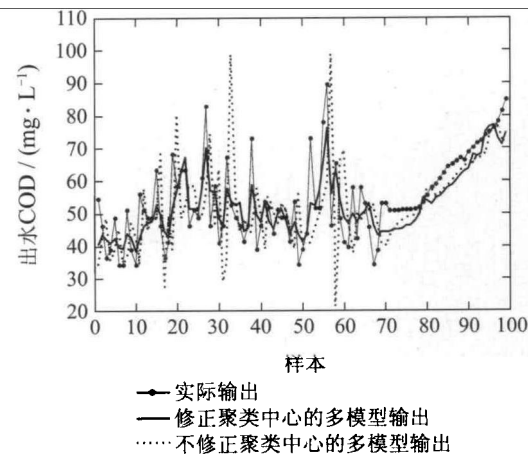


图7 水质COD多模型在线软测量结果

Fig. 7 The estimated COD of soft sensor based on multi-model

在线修正工况中心时COD软测量结果与实际COD较为接近; 不修正时在工况波动下软测量偏差较大. 通过修正工况中心实时跟踪工况变化趋势, 较准确地反映了污水处理过程实际运行状况. 图8为基于多模型的软测量结果、各局部模型结果与实际COD的比较. 多模型结果与实际值拟合较

好,但各局部模型输出与实际值之间均存在较大偏差,以任意局部模型输出或当前时刻隶属度最大的局部模型输出作为软测量输出均不能获得较好的精度,因此多模型合成机制可以刻画工况重叠区域的过程特性,具有满意的软测量性能。

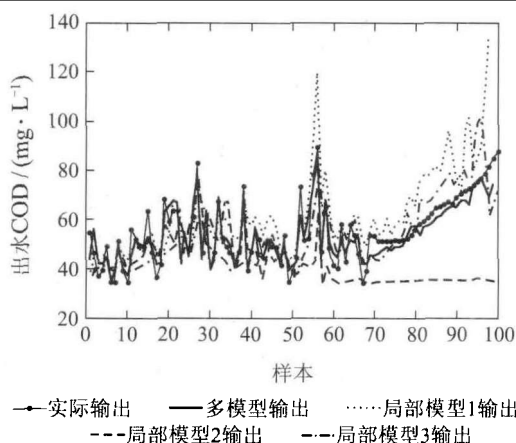


图8 多模型和局部模型在线软测量结果

Fig. 8 The results of soft sensor based on multi-model and local models

2) 局部模型采用WNN和多层感知器(multi-layer perceptron, MLP)时软测量结果与分析。

文献[11]采用MLP描述局部H模型的非线性增益。图9为采用本文方法的结果、采用文献[11]方法的结果与实际COD值的比较。

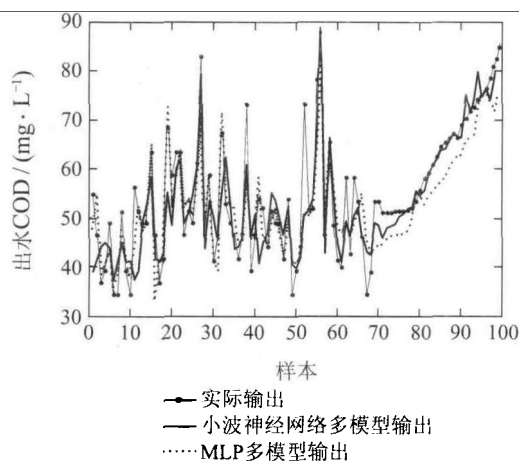


图9 采用WNN与MLP的多模型软测量结果

Fig. 9 The estimated COD of WNN-based and MLP-based multi-model

比较两类模型计算的COD值和实际COD之间的均方根误差RMSE(root-mean-square error),采用MLP时, RMSE为18.45 mg/L; 采用WNN时, RMSE为15.54 mg/L。图10为RMSE趋势图,平均建模误差指标逐渐降低并趋近恒值,证明了所提方法建模误差的有界性。

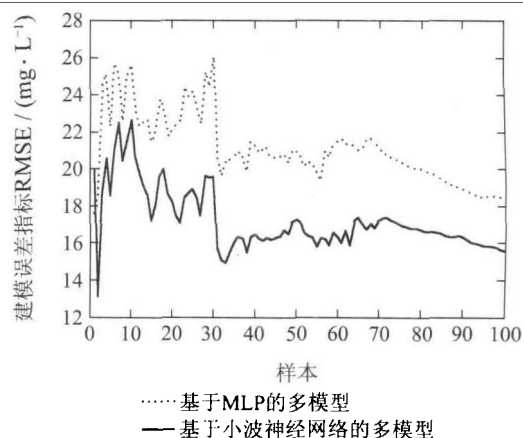


图10 建模误差指标趋势图

Fig. 10 the curves of the modeling error

当局部H模型非线性增益采用MLP时,多模型具有较好的软测量性能;而MLP表征非线性的能力明显弱于WNN,当采用WNN时在不同尺度下描述特定工况的非线性特性,故基于WNN的多模型软测量结果相对理想,优于基于MLP的多模型结果。

5 结论(Conclusions)

本文针对运行工况频繁波动的工业过程建模问题,提出了带有工况中心修正的多模型在线软测量方法,其中在线修正工况中心有效地反映了工况的动态变化趋势;WNN和H模型结合的局部模型拟合了各工况的非线性特性,稳定学习算法保证了建模误差有界;基于加权求和的多模型切换机制实现了局部模型之间的平滑切换,多模型具有满意的软测量性能。污水处理过程的仿真实验表明了所提方法的有效性。

参考文献(References):

- [1] JOHANSEN T A. *Operating regime based process modeling and identification* [D]. Trondheim: The Norwegian Institute of Technology, University of Trondheim, 1994.
- [2] RAGOT J, GRAPIN G, CHATELLIER P, et al. Modelling of a water treatment plant: a multi-model representation [J]. *Environmetrics*, 2001, 12(7): 599 – 611.
- [3] 刘德馨, 李晓理, 丁大伟, 等. 基于雷达观测数据的高炉料面多模型控制 [J]. *控制理论与应用*, 2012, 29(10): 1277 – 1283. (LIU Dexin, LI Xiaoli, DING Dawei, et al. Multi-model control of blast furnace burden surface based on observed data of radars [J]. *Control Theory & Applications*, 2012, 29(10): 1277 – 1283.)
- [4] 贾文君, 柴天佑. 稀土串级萃取分离过程元素组分含量的多模型软测量 [J]. *控制理论与应用*, 2007, 24(4): 568 – 580. (JIA Wenjun, CHAI Tianyou. Soft-sensor of element component content based on multiple models for the rare earth cascade extraction process [J]. *Control Theory & Applications*, 2007, 24(4): 568 – 580.)
- [5] SERIR L, RAMASSO E, NECTOUX P, et al. E2GKpro: an evidential evolving multi-modeling approach for system behavior prediction with applications [J]. *Mechanical Systems and Signal Processing*, 2012, (7): 1 – 13.

- [6] YOO C K, VANROLLEGHEM P A, LEE I B. Nonlinear modeling and adaptive monitoring with fuzzy and multivariate statistical methods in biological wastewater treatment plants [J]. *Journal of Biotechnology*, 2003, 105(1): 135 - 163.
- [7] 赵立杰, 张金玲, 陶成虎. 污水处理出水水质指标的非线性动态软测量模型 [J]. 沈阳化工学院学报, 2009, 23(3): 258 - 265. (ZHAO Lijie, ZHANG Jinling, TAO Chenghu. Nonlinear dynamic soft sensing model of wastewater treatment effluent water quality indicators [J]. *Journal of Shenyang University of Chemical Technology*, 2009, 23(3): 258 - 265.)
- [8] LAKSHMINARAYANAN S, SHAH S L, NANDAKUMAR K. Modeling and control of multivariable processes: dynamic PLS approach [J]. *AIChE Journal*, 1997, 43(9): 2307 - 2322.
- [9] YU W. Nonlinear system identification using discrete-time recurrent neural networks with stable learning algorithms [J]. *Information Sciences*, 2004, 158(1): 131 - 147.
- [10] YAO J, DASH M, TAN S T, et al. Entropy-based fuzzy clustering and modeling [J]. *Fuzzy Sets and Systems*, 2000, 113(3): 381 - 388.
- [11] 丛秋梅, 赵立杰, 柴天佑. 一类污水处理过程水质多模型在线软测量方法 [J]. 东北大学学报(自然科学版), 2010, 31(9): 1221 - 1225. (CONG Qiumei, ZHAO Lijie, CHAI Tianyou. A multi-model soft-sensing method of water quality in wastewater treatment process [J]. *Journal of Northeastern University (Nature Science)*, 2010, 31(9): 1221 - 1225.)

附录(Appendix)

采用梯度下降法学习式(7)中参数. 为了最小化 E , 根据误差反传法的链式规则, 各参数在 E 的负梯度下降方向进行训练, 修正量分别为

$$\Delta w_{1i}(k) = -\eta_1(k) \frac{\partial E(k)}{\partial w_{1i}(k)} = -\eta_1(k) e_{n1}(k) \phi_{1i}(k), \tag{A1}$$

$$\Delta v_{1ij}(k) = -\eta_1(k) \frac{\partial E(k)}{\partial v_{1ij}(k)} = -\eta_1(k) e_{n1}(k) w_{1i}(k) \phi'_{1i}(k) x_j(k), \tag{A2}$$

$$\Delta s_{1i}(k) = -\eta_1(k) \frac{\partial E(k)}{\partial s_{1i}(k)} = \eta_1(k) e_{n1}(k) \frac{w_{1i}(k)}{s_{1i}^2(k)} \phi'_{1i}(k) \left(\sum_{j=1}^I v_{1ij}(k) x_j(k) - d_{1i}(k) \right), \tag{A3}$$

$$\Delta d_{1i}(k) = -\eta_1(k) \frac{\partial E}{\partial d_{1i}} = \eta_1(k) e_{n1}(k) \frac{w_{1i}(k)}{s_{1i}(k)} \phi'_{1i}(k), \tag{A4}$$

其中: $\eta_1(k)$ 表示WNN的学习率; $e_{n1}(k)$ 表示WNN的建模误差, 定义为 $e_{n1}(k) = \hat{u}_1(k) - u_1(k)$, $\hat{u}_1(k)$ 表示WNN输出, $u_1(k)$ 表示WNN假定理想输出, 无法精确测量, 因此 $e_{n1}(k)$ 是虚拟误差指标, $i = 1, 2, \dots, H, j = 1, 2, \dots, I$.

$e_{n1}(k)$ 与 $e(k)$ 之间的关系由链式规则推导:

$$\frac{\partial E}{\partial w_{1i}} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial \hat{u}_1} \frac{\partial \hat{u}_1}{\partial w_{1i}} = -\lambda_1(k) b_{10}(k) e(k) \phi_{1i}(k). \tag{A5}$$

当以 $e_{n1}(k)$ 校正模型参数时, 下式成立:

$$\frac{\partial E}{\partial w_{1i}} = -e_{n1}(k) \phi_{1i}(k). \tag{A6}$$

根据式(A5)和式(A6), $e(k)$ 和 $e_{n1}(k)$ 具有如下关系:

$$e_{n1}(k) = \lambda_1(k) b_{10}(k) e(k). \tag{A7}$$

将式(A7)带入式(A1)-(A4), 在线更新算法写为

$$W_1(k+1) = W_1(k) - \eta_1(k) \lambda_1(k) b_{10}(k) e(k) \phi_1(k), \tag{A8}$$

$$V_1(k+1) =$$

$$V_1(k) - \eta_1(k) \lambda_1(k) b_{10}(k) e(k) \phi'_1(k) W_1(k) \mathbf{x}^T(k), \tag{A9}$$

$$s_{1i}(k+1) =$$

$$s_{1i}(k) + \eta_1(k) \lambda_1(k) b_{10}(k) e(k) \frac{w_{1i}(k)}{s_{1i}^2(k)} \cdot$$

$$\phi'_{1i}(k) \left(\sum_{j=1}^I v_{1ij}(k) x_j(k) - d_{1i}(k) \right), \tag{A10}$$

$$d_{1i}(k+1) =$$

$$d_{1i}(k) + \eta_1(k) \lambda_1(k) b_{10}(k) e(k) \frac{w_{1i}(k)}{s_{1i}(k)} \phi'_{1i}(k). \tag{A11}$$

根据Stone-Weierstrass定理, COD软测量多模型写为

$$\hat{y}(k) = \sum_{l=1}^m \lambda_l \left(- \sum_{i=1}^{n_a} a_{li} \hat{y}_l(k-i) + \sum_{j=0}^{n_b} b_{lj} W_l^*(k-j) \phi_l^*(V_l^*(k-j) \mathbf{x}(k-j)) \right) - \mu(k), \tag{A12}$$

其中 W_l^* 和 V_l^* , S_l^* 和 D_l^* 分别表示使辨识误差 $\mu(k)$ 最小时 M_l 的权值矩阵和小波参数向量: $W_l^* = [w_{li}^*]$, $V_l^* = [v_{lij}^*]$, $D_l^* = [d_{li}^*]$, $\phi_l^*(x) = \psi\left(\frac{x - d_{li}^*}{s_{li}^*}\right)$.

$e(k)$ 可写为

$$e(k) = - \sum_{l=1}^m \lambda_l(k) \left(\sum_{i=1}^H \phi_{li}(w_{li} - w_{li}^*) + \sum_{i=1}^H \sum_{j=1}^I \frac{\partial y_l}{\partial v_{lij}} (v_{lij} - v_{lij}^*) + \sum_{i=1}^H \frac{\partial y_l}{\partial s_{li}} (s_{li} - s_{li}^*) + \sum_{i=1}^H \frac{\partial y_l}{\partial d_{li}} (d_{li} - d_{li}^*) \right) + \delta(k), \tag{A13}$$

其中: $\delta(k)$ 表示过程未建模动态, $\delta(k) = \mu(k) + \varepsilon(k)$, $\varepsilon(k)$ 为Taylor级数高阶项. 因此 $e_{n1}(k)$ 可写为

$$e_{n1}(k) = \hat{u}_1(k) - u_1(k) = \sum_{i=1}^H \phi_{1i}(w_{1i} - w_{1i}^*) + \sum_{i=1}^H \sum_{j=1}^I \frac{\partial u_1}{\partial v_{1ij}} (v_{1ij} - v_{1ij}^*) + \sum_{i=1}^H \frac{\partial u_1}{\partial s_{1i}} (s_{1i} - s_{1i}^*) + \sum_{i=1}^H \frac{\partial u_1}{\partial d_{1i}} (d_{1i} - d_{1i}^*) + \mu_1(k).$$

写成矩阵形式为

$$e_{n1}(k) = \tilde{W}_1(k) \phi_1 + E^T \tilde{V}_1(k) W_{V1}(k) \mathbf{x}^T(k) E + W_{S1}(k) \tilde{S}_1(k) + W_{D1}(k) \tilde{D}_1(k) + \delta_1(k), \tag{A14}$$

$$\tilde{W}_1(k) = W_1(k) - W_1^* = [w_{11} - w_{11}^* \quad \dots \quad w_{1H} - w_{1H}^*],$$

$$\tilde{V}_1(k) = V_1(k) - V_1^* = \begin{pmatrix} v_{111} - v_{111}^* & \dots & v_{1H1} - v_{1H1}^* \\ \vdots & & \vdots \\ v_{11I} - v_{11I}^* & \dots & v_{1HI} - v_{1HI}^* \end{pmatrix},$$

$$\tilde{S}_1(k) = S_1(k) - S_1^* = [s_{11} - s_{11}^* \quad \dots \quad s_{1H} - s_{1H}^*]^T,$$

$$\bar{D}_1(k) = D_1(k) - D_1^* = [d_{11} - d_{11}^* \cdots d_{1H} - d_{1H}^*]^T,$$

$$\delta_1(k) = \varepsilon_1(k) + \mu_1(k).$$

其中: $\bar{W}_1(k)$ 和 $\bar{V}_1(k)$, $\bar{S}_1(k)$ 和 $\bar{D}_1(k)$ 分别表示WNN权值矩阵和尺度参数的估计误差; $\varepsilon_1(k)$ 表示Taylor级数高阶项, $\delta_1(k)$ 表示 M_1 的未建模动态, $\mu_1(k)$ 表示WNN的辨识误差.

证 定义正定函数 L_1 为

$$L_1(k) = \|\bar{W}_1(k)\|^2 + \|\bar{V}_1(k)\|^2 + \|\bar{S}_1(k)\|^2 + \|\bar{D}_1(k)\|^2.$$

由式(A8)–(A11)和式(A14)可得

$$\begin{aligned} \Delta L_1(k) &= L_1(k+1) - L_1(k) = \\ &\eta_1^2(k)\|\lambda_1 b_{10}\|^2\|e(k)\|^2(\|\phi_1\|^2 + \|W_{V1}(k)\mathbf{x}^T(k)\|^2 + \\ &\|W_{S1}(k)\|^2 + \|W_{D1}(k)\|^2) - 2\eta_1(k)\|\lambda_1 b_{10}\|^2\|e(k)\| \cdot \\ &\|\bar{W}_1(k)\phi_1 + \mathbf{E}^T \bar{V}_1(k)W_{V1}(k)\mathbf{x}^T(k)\mathbf{E} + \\ &W_{S1}(k)\bar{S}_1(k) + W_{D1}(k)\bar{D}_1(k)\| \leq \\ &-\eta_1(k)\|e(k)\|^2[1 - \eta_1(k)(\|\lambda_1 b_{10}\phi_1\|^2 \cdot \\ &\|\lambda_1 b_{10}W_{V1}(k)\mathbf{x}^T(k)\|^2 + \|\lambda_1 b_{10}W_{S1}(k)\|^2 + \\ &\|\lambda_1 b_{10}W_{D1}(k)\|^2)] + \eta_1(k)\|\delta_1(k)\|^2 \leq \\ &-\pi_1\|e(k)\|^2 + \eta_1(k)\|\delta_1(k)\|^2. \end{aligned}$$

因为

$$\begin{aligned} I[\min(\tilde{w}_{1i}^2) + \min(\tilde{v}_{1ij}^2) + \min(\tilde{s}_{1i}^2) + \min(\tilde{d}_{1i}^2)] &\leq \\ L_1(k) &\leq I[\max(\tilde{w}_{1i}^2) + \max(\tilde{v}_{1ij}^2) + \max(\tilde{s}_{1i}^2) + \max(\tilde{d}_{1i}^2)], \end{aligned}$$

并且 $I[\min(\tilde{w}_{1i}^2) + \min(\tilde{v}_{1ij}^2) + \min(\tilde{s}_{1i}^2) + \min(\tilde{d}_{1i}^2)]$ 和 $I[\max(\tilde{w}_{1i}^2) + \max(\tilde{v}_{1ij}^2) + \max(\tilde{s}_{1i}^2) + \max(\tilde{d}_{1i}^2)]$ 是 κ_∞ 函数, $\pi_1\|e(k)\|^2$ 是 κ_∞ 函数, $\eta_1(k)\|\delta_1(k)\|^2$ 是 κ 函数; $L_1(k)$ 是 $e(k)$ 和 $\delta_1(k)$ 的函数, 因此 $L_1(k)$ 满足一个ISS-Lyapunov函数. 将 $\delta_1(k)$ 看做“输入”, $e(k)$ 看做“状态”, 由于 $\delta_1(k)$ 是有界的, 因此 $e(k)$ 也是有界的, 并且平均建模误差指标满足

$$\bar{J}_1 = \lim_{T \rightarrow \infty} \sup \frac{1}{T} \sum_{k=1}^T \|e(k)\|^2 \leq \frac{\eta_0}{\pi_1} \bar{\delta}_1, \quad (\text{A15})$$

其中 $\bar{\delta}_1 = \max_k \{\|\delta_1(k)\|^2\}$. 因此多模型建模误差指标满足

$$\bar{J} \leq \max_i \{\bar{J}_i\}, \quad i = 1, 2, \dots, m. \quad (\text{A16})$$

作者简介:

丛秋梅 (1978-), 女, 博士, 目前研究方向为复杂工业过程的建模、控制与能效优化, E-mail: cong.0828@163.com;

苑明哲 (1971-), 男, 研究员, 博士, 目前研究方向为复杂工业工程建模与优化控制、工业通讯技术, E-mail: mzyuan@sia.cn;

柴天佑 (1947-), 男, 中国工程院院士, 教授, 博士生导师, IEEE Fellow, IFAC Fellow, 欧亚科学院院士, 目前研究方向为自适应控制、多变量智能解耦控制、流程工业过程综合自动化理论、方法与技术, E-mail: tychai@mail.neu.edu.cn;

王宏 (1963-), 男, 研究员, 博士生导师, 目前研究方向为工业控制系统、现场总线和智能仪器仪表, E-mail: wang.hong@micro-cyber.cn.