

基于自适应双层无迹卡尔曼滤波神经网络的 铝电解电流效率预测模型

方小燕¹, 姚立忠^{2†}, 罗海军^{2,3}, 张玉泽¹, 易 军⁴

(1. 重庆科技大学 电子与电气工程学院, 重庆 401331; 2. 重庆师范大学 物理与电子工程学院, 重庆 401331;

3. 重庆国家应用数学中心, 重庆 401331; 4. 重庆科技大学 计算机科学与工程学院, 重庆 401331)

摘要: 针对铝电解过程强干扰和强时变导致模型精确度和稳定性不佳的难题, 本文提出一种基于自适应双层无迹卡尔曼滤波神经网络的建模方法. 该方法首先构建一种双层无迹卡尔曼滤波神经网络模型, 以提高模型对扰动系统的稳定性. 具体为: 使用双层无迹卡尔曼滤波在线更新神经网络的权值和阈值; 然后, 在双层无迹卡尔曼滤波神经网络的状态变量均方误差中引入约束调节参数; 同时, 采用梯度下降法自适应调整比例调节参数, 将其均方误差约束至较小的范围内, 以此来削弱滤波递归计算过程中误差累积对模型的影响; 最后, 通过铝电解电流效率预测, 验证了本文所提方法具有较高的精确度和稳定性.

关键词: 铝电解; 自适应建模; 双层无迹卡尔曼滤波; 人工神经网络; 电流效率

引用格式: 方小燕, 姚立忠, 罗海军, 等. 基于自适应双层无迹卡尔曼滤波神经网络的铝电解电流效率预测模型. 控制理论与应用, 2025, 42(3): 579 – 589

DOI: 10.7641/CTA.2024.30047

Prediction model for the current efficiency of aluminum electrolysis based on the adaptive double layer unscented Kalman filter neural network

FANG Xiao-yan¹, YAO Li-zhong^{2†}, LUO Hai-jun^{2,3}, ZHANG Yu-ze¹, YI Jun⁴

(1. School of Electronic and Electrical Engineering, Chongqing University of Science and Technology, Chongqing 401331, China;

2. College of Physics and Electronic Engineering, Chongqing Normal University, Chongqing 401331, China;

3. National Center for Applied Mathematics in Chongqing, Chongqing 401331, China)

4. School of Computer Science and Engineering, Chongqing University of Science and Technology, Chongqing 401331, China)

Abstract: This paper presents a novel modeling method based on an adaptive double layer unscented Kalman filter neural network (ADLUKFNN), which tackles the challenges of poor model accuracy and stability resulting from strong interference and time-varying disturbances in the aluminum electrolysis process. Firstly, this method constructs a double layer unscented Kalman filter neural network (DLUKFNN) model to enhance the stability of the model towards the disturbance system. Specifically, the weights and thresholds of the neural network are updated online using the double layer unscented Kalman filter. Then, a constraint adjustment parameter is introduced into the mean square error of the state variables in DLUKFNN. Meanwhile, by employing the gradient descent method to adaptively adjust the constraint adjustment parameter, the mean square error is constrained within a smaller range, thereby weakening the impact of error accumulation during the filtering recursive calculation on the model. Finally, the accuracy and stability of the proposed method are verified through aluminum electrolysis current efficiency prediction.

Key words: aluminum electrolysis; adaptive modeling; double layer unscented Kalman filtering; neural network; current efficiency

Citation: FANG Xiaoyan, YAO Lizhong, LUO Haijun, et al. Prediction model for the current efficiency of aluminum electrolysis based on the adaptive double layer unscented Kalman filter neural network. *Control Theory & Applications*, 2025, 42(3): 579 – 589

收稿日期: 2023-02-06; 录用日期: 2024-03-29.

†通信作者. E-mail: lizhong_yao@cqu.edu.cn.

本文责任编辑: 潘泉.

国家自然科学基金项目(62373069, 51805059), 重庆市教委科学技术研究项目(KJQN202200531, KJZD-M202101501), 重庆师范大学基金项目(22XLB014)资助.

Supported by the National Nature Science Foundation of China (62373069, 51805059), the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJQN202200531, KJZD-M202101501) and the Natural Science Foundation of Chongqing Normal University (22XLB014).

1 引言

我国原铝产量和消费量连续十几年保持世界第一, 但我国流程工业单位GDP能耗、物耗和污染排放较发达国家仍有较大差距^[1]. 提高电流效率是铝电解过程重要的生产目标之一, 同时, 也是降低能耗的关键因素. 目前, 研究者主要从改善硬件性能和优化槽控参数两个方面提高电流效率^[2-3]. 显然, 改善硬件设备性能需要投入大量人力财力资源, 对于中小铝电解企业并不现实. 相对改善硬件设备, 优化槽控参数可以在原有设备基础之上, 实现提高电流效率的目标. 但前提是需有准确且可靠的建模技术, 用于描述槽控参数和电流效率之间的潜在关系. 因此, 研究稳定性高、精确度高的铝电解电流效率建模技术, 对铝电解实际工业生产具有重要意义.

人工神经网络(neural network, NN)基于其卓越的非线性拟合能力, 已经在铝电解过程建模中具有丰富的研究. 例如: 卷积神经网络用于铝电解过热度检测^[4]; 长短时记忆神经网络用于铝电解槽阳极效应预测和锅温检测^[5-6]; 模糊神经网络制定氟化铝添加策略^[7]. 但是, 实际槽控参数优化过程需要在铝电解生产系统频繁波动的情况下, 实时求解最优槽控参数. 传统NN已被证实属于静态建模方法^[8]. 由于铝电解过程内外部干扰因素众多^[9], 因此, 若想建立精确的铝电解过程模型, 需要结合传统NN优点的同时, 具备对系统的自适应学习能力. 基于这种需求, 学者提出使用扩展卡尔曼滤波(extended Kalman filter, EKF)、无迹卡尔曼滤波(unscented Kalman filter, UKF)和粒子滤波(particle filter, PF)在线更新NN的权值和阈值. 文献[10]中介绍相较于传统反向传播学习算法, EKF实时训练NN参数具有更快的收敛性; 文献[11-12]设计了一种强跟踪UKF神经网络, 并将其用于铝电解过程建模, 相较于传统NN具有更高的鲁棒性; 文献[13]采用UKF训练径向基神经网络的权重、中心值和宽度, 并将该方法应用于抽油机故障诊断, 实验结果表明该方法可有效提高抽油机故障诊断地准确性; 文献[14-16]使用PF实时训练NN的权值和阈值, 其建立的粒子滤波神经网络(particle filter NN, PFNN)能较高精确度预测铝电解能耗. 但是, 上述滤波中, EKF已被证实处理强非线性系统时精确度和稳定性远不如UKF^[17-18]. PF中粒子退化和重要性函数的选取仍是研究的难点问题^[19-21]. 同时, 在复杂系统中, PF需要大量的粒子逼近系统会显著增加模型计算负担^[22-23]. 相较于PF, UKF采用一组确定粒子逼近系统的均值和方差, 不会产生粒子退化的同时可以达到泰勒三阶精确度^[24].

需要指出的是, 滤波神经网络搭建状态空间方程

时, 状态变量一步预测过程不包含系统的时变信息. 当模型遭遇突变扰动时, 模型稳定性会降低. 文献[25]提出一种双层无迹卡尔曼滤波(double layer UKF, DLUKF)方法, 其主要思想之一是将现场量测值修正过的内层滤波结果作为外层的初始信息, 这则可以看作模型已经包含了系统的时变信息. 因此, 本文利用DLUKF实时训练NN的权值和阈值, 构建了一种双层无迹卡尔曼滤波神经网络(double layer UKFNN, DLUKFNN)方法, 用于解决状态变量一步预测过程中未包含系统时变信息的难题. 同时, DLUKFNN的估计状态变量经过两次现场量测值的修正, 能够进一步加强模型对扰动的自适应性. 另外, DLUKFNN并未改变UKF计算的核心过程, 本质仍是一种递归的估计方法. 由于递归计算过程中, 第 k 时刻估计时会考虑历史时刻的值, 随着滤波次数的增加, 存在误差累积的情况. 为此, 本文以DLUKFNN的状态变量均方误差最小为目标, 设计了一种自适应策略, 进一步提高模型精确度. 综上所述, 本文主要创新和贡献如下:

1) 利用UKF对动态系统的跟踪能力, 赋予传统NN动态性能, 具体为: 将NN的权值阈值作为UKF的状态变量, 输出作为其量测变量. 为提高模型对扰动系统的稳定性, 提出使用DLUKF实时训练NN的权值和阈值.

2) 为改善DLUKFNN中的误差累积, 提出一种状态变量的均方误差自适应调整策略. 在DLUKFNN中引入约束调节参数, 并结合状态变量均方误差构建代价函数. 然后, 采用梯度下降法自适应调整约束调节参数, 削弱因滤波次数增加而产生的误差累积.

3) 结合上述两种策略, 构建自适应双层无迹卡尔曼滤波神经网络(adaptive DLUKFNN, ADLUKFNN)的建模机制与完整架构, 并建立铝电解槽工艺电流效率动态预测模型. 同时, 以重庆某铝厂电解车间工艺系统为例, 采集9种主要铝电解过程槽控变量进行电流效率建模. 从拟合实验、随机实验和人为扰动3个角度评判所提模型对扰动系统的稳定性和精确度. 实验结果表明, ADLUKFNN具有较高的预测精度和稳定性.

2 ADLUKFNN动态预测模型设计

2.1 DLUKFNN算法设计及分析

将NN权值和阈值作为UKF的状态变量, 将NN预测输出作为UKF的预测量测变量, 详细无迹卡尔曼滤波神经网络(unscented Kalman filter neural network, UKFNN)算法可以参考文献[11]. UKFNN状态空间方程表示为

$$\begin{cases} x_k = x_{k-1} + q_{k-1}, \\ z_k = f(u_k, x_k) + r_k, \end{cases} \quad (1)$$

式中: $x_k \in \mathbb{R}^{n \times 1}$ 和 $z_k \in \mathbb{R}^{m \times 1}$ 分别表示第 k 时刻的状态变量和量测变量; u_k 表示第 k 时刻输入变量. q_{k-1} 和 r_k 分别表示第 k 时刻和 $k-1$ 时刻的系统噪声和量测噪声, 且两者互不相关. 本文假设 q_{k-1} 和 r_k 服从均值为 0, 协方差分别为 $E[q_{k-1}q_{k-1}^T] = Q_{k-1} \in \mathbb{R}^{n \times n}$ 和 $E[r_k r_k^T] = R_k \in \mathbb{R}^{m \times m}$ 的高斯白噪声. $f(\cdot)$ 为 NN 前馈传播网络结构.

基于式(1)状态空间方程, 将DLUKFNN动态预测模型的详细实施步骤描述如下:

步骤 1 初始化状态变量.

$$\begin{cases} \hat{x}_0 = E(x_0), \\ \Sigma_0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]. \end{cases} \quad (2)$$

步骤 2 sigma对称采样.

$$\begin{cases} \chi_{0,k-1}^* = \hat{x}_{k-1}^{**}, \\ \chi_{i,k-1}^* = \hat{x}_{k-1}^{**} + \sqrt{(n+a^2(n+\gamma)-n) \times (\sqrt{\hat{\Sigma}_{k-1}^{**}})_i}, \quad i = 1, \dots, n, \\ \chi_{i,k-1}^* = \hat{x}_{k-1}^{**} - \sqrt{(n+a^2(n+\gamma)-n) \times (\sqrt{\hat{\Sigma}_{k-1}^{**}})_{i-n}}, \quad i = n+1, \dots, 2n. \end{cases} \quad (3)$$

步骤 3 状态预测.

$$\begin{cases} \chi_{i,k|k-1}^* = \chi_{i,k-1}^*, \\ \hat{x}_k^{-*} = \sum_0^{2n} w_i^{(m)} \chi_{i,k|k-1}^*, \\ Z_{i,k}^* = f(\chi_{i,k|k-1}^*), \\ \hat{\Sigma}_k^{-*} = \sum_0^{2n} w_i^{(c)} (\chi_{i,k|k-1}^* - \hat{x}_k^{-*}) \times (\chi_{i,k|k-1}^* - \hat{x}_k^{-*})^T + Q_{k-1}, \\ \hat{z}_k^{-*} = \sum_{i=0}^{2n} w_i^{(m)} Z_{i,k}^*. \end{cases} \quad (4)$$

步骤 4 状态校正.

$$\begin{cases} \hat{\Sigma}_{z,k}^* = \sum_0^{2n} w_i^{(c)} (Z_{i,k}^* - \hat{z}_k^{-*}) \times (Z_{i,k}^* - \hat{z}_k^{-*})^T + R_k, \\ \hat{\Sigma}_{xz,k}^* = \sum_0^{2n} w_i^{(c)} (\chi_{i,k|k-1}^* - \hat{x}_{k|k-1}^{-*}) \times (Z_{i,k}^* - \hat{z}_k^{-*})^T, \\ K_k^* = \hat{\Sigma}_{xz,k}^* (\hat{\Sigma}_{z,k}^*)^{-1}, \\ \hat{x}_k^* = \hat{x}_k^{-*} + K_k^* (z_k - \hat{z}_k^{-*}), \\ \hat{\Sigma}_k^* = \hat{\Sigma}_{k|k-1}^{-*} - K_k^* \hat{\Sigma}_{z,k}^* (K_k^*)^T. \end{cases} \quad (5)$$

步骤 5 将 \hat{x}_k^* 和 $\hat{\Sigma}_k^*$ 作为下层滤波的初始信息, 定点再次进行sigma对称采样.

$$\begin{cases} \chi_{0,k}^{**} = \hat{x}_k^*, \\ \chi_{i,k}^{**} = \hat{x}_k^* + \sqrt{(n+a^2(n+\gamma)-n) \times (\sqrt{\hat{\Sigma}_k^{**}})_i}, \quad i = 1, \dots, n, \\ \chi_{i,k-1}^{**} = \hat{x}_{k-1}^* - \sqrt{(n+a^2(n+\gamma)-n) \times (\sqrt{\hat{\Sigma}_{k-1}^{**}})_{i-n}}, \quad i = n+1, \dots, 2n. \end{cases} \quad (6)$$

步骤 6 再次进行步骤2至步骤5过程, 得到 k 时刻下层估计状态变量 \hat{x}_k^{**} 和预测协方差 $\hat{\Sigma}_k^{**}$.

步骤 7 \hat{x}_k^{**} 更新 NN 权值和阈值; 同时, 按式(7)进行 NN 前馈计算.

$$f(u_k, \hat{x}_k^{**}) = \sum_{j=1}^{12} \frac{w_{js,k}^{23}}{1 + \exp[-(\sum_{i=1}^9 w_{ij,k}^{12} u_{i,k} + b_{j,k}^3)]} + b_{s,k}^3. \quad (7)$$

步骤 8 重复步骤2至步骤7过程. 此外, 本文使用工业实况数据对模型进行验证. 当采集的数据运行完毕, 模型停止运行.

上述式子中: 上标*表示上层UKF计算结果, **表示下层UKF计算结果; x 表示状态变量; Σ 表示协方差矩阵; k 表示时刻; z_k 表示第 k 时刻的现场量测值; $(\sqrt{\Sigma})_i$ 采用 Cholesky 分解得到, i 表示 $\sqrt{\Sigma}$ 的第 i 列; $w_i^{(m)}$ 和 $w_i^{(c)}$ 表示采样点对应均值和协方差的权值, i 表示第 i 个采样点; n 表示状态变量维度; a 决定 sigma 点与均值距离的大小; γ 为自由度参数; $\hat{x}_k^{**} = [w_{i,j,k}^{12} \quad b_{j,k}^2 \quad w_{js,k}^{23} \quad b_{s,k}^3]^T$; $w_{ij,k}^{12}$ 表示第 i 个输入与第 j 个隐含层神经元在第 k 时刻的连接权值; $w_{js,k}^{23}$ 表示第 j 个隐含层神经元与第 s 个输出层神经元在第 k 时刻的连接权值; $b_{s,k}^3$ 表示输出层神经元在第 k 时刻的阈值; $u_{i,k}$ 表示 k 时刻的第 i 个输入参数; 9 表示输入层具有 9 个神经元; 12 表示隐含层具有 12 个神经元.

UKFNN 状态空间方程一步预测过程不包含系统的时变信息. UKFNN 的权值和阈值为其估计状态变量, 而估计状态变量通过一步预测值与在线量测值进行融合计算而得到, 若估计状态变量中缺乏系统的时变信息, NN 前馈计算时会存在较大的参数不确定性. 而从上述 DLUKFNN 搭建过程可知, DLUKFNN 将上层滤波结果作为下层滤波的初始信息, 则下层滤波结果中包含了系统的时变信息. 同时, DLUKFNN 估计状态变量经过两次现场量测值的修正, 对系统时变和环境扰动更具备自适应性. 综合上述策略, 图1给出了 DLUKFNN 模型框架.

2.2 ADLUKFN算法设计及分析

UKF 是一种基于最小均方误差原则的递推算法. 其中, DLUKFNN 状态变量预测误差斜方差矩阵表示为

$$E(\tilde{x}\tilde{x}^T)_k = \begin{pmatrix} \tilde{x}_1^2 & \cdots & \tilde{x}_1\tilde{x}_n \\ \vdots & & \vdots \\ \tilde{x}_n\tilde{x}_1 & \cdots & \tilde{x}_n^2 \end{pmatrix}_k, \quad (8)$$

式中: n 表示状态变量维度; \tilde{x}_n 表示状态变量最优估计值和真实值之差, 表示为

$$\tilde{x}_n = \hat{x}_n^{**} - x_n. \quad (9)$$

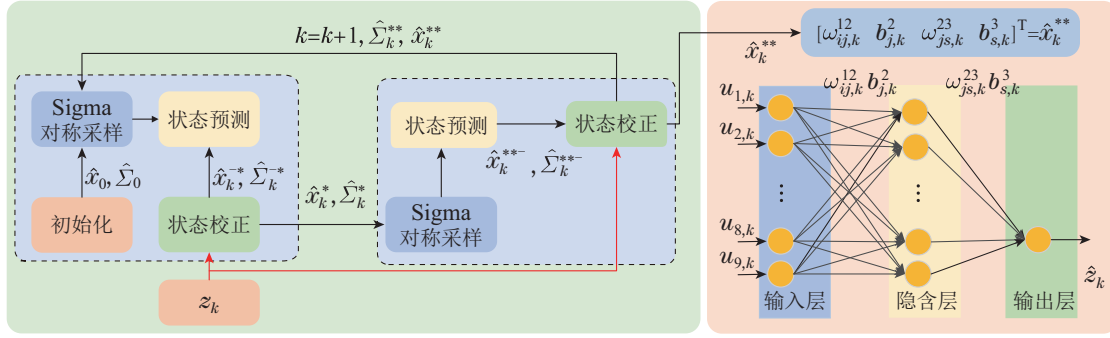


图1 DLUKFNN算法框架

Fig. 1 Algorithm framework of DLUKFNN

梯度下降法使用迭代计算的方式, 不断调整参数取值, 能够使目标函数的值不断减小. 受此启发, 本文在DLUKFNN状态变量的先验预测方差和量测预测方差中引入两个约束调节参数, 并构建代价函数. 然后, 使用梯度下降法自适应调整这两个约束调节参数. 期望能够将代价函数值约束至较小范围内, 以此来迫使DLUKFNN状态变量均方误差逼近较小的理想值.

将上述自适应策略优化DLUKFNN的具体实施步骤阐述如下:

步骤1 执行DLUKFNN得到 $\hat{\Sigma}_k^{**}$ 和 \hat{x}_k^{**} ;

步骤2 更新NN的权值和阈值. 同时, \hat{x}_k^{**} 按式(7)NN前馈计算;

步骤3 在DLUKFNN的状态变量估计方差中, 引入两个约束调节参数, 并构造式(10)代价函数, 即

$$J(\mu_k, \nu_k) = \mu_k^2 \text{tr}(\Sigma_k^{**}) + \nu_k^2 \text{tr}(-(K_k^{**} \hat{\Sigma}_{z,k} (K_k^{**})^T)), \quad (10)$$

式中: μ_k 和 ν_k 为第 k 时刻估计状态变量方差的约束调节参数, 上标2表示平方.

步骤4 基于式(10), 约束调节参数 μ_k 和 ν_k 使用梯度下降法更新;

$$\begin{vmatrix} \mu_k \\ \nu_k \end{vmatrix} = \begin{vmatrix} \mu_{k-1} \\ \nu_{k-1} \end{vmatrix} - \text{learn} \times \begin{vmatrix} \frac{\partial J(\mu_{k-1}, \nu_{k-1})}{\partial \mu_{k-1}} \\ \frac{\partial J(\mu_{k-1}, \nu_{k-1})}{\partial \nu_{k-1}} \end{vmatrix}, \quad (11)$$

即

$$\begin{vmatrix} \mu_k \\ \nu_k \end{vmatrix} = \begin{vmatrix} \mu_{k-1} \\ \nu_{k-1} \end{vmatrix} - \text{learn} \times$$

式(8)中, 对角线之和即为状态变量的均方误差. 根据最小均方误差原则, 状态变量的均方误差和即 $\text{tr}(E(\tilde{x}\tilde{x}^T))$ 应满足趋于极小值. 其中 tr 表示矩阵的迹. 但是, UKF在进行第 k 时刻估计时, 会考虑 k 时刻之前的所有数据, 随着滤波次数的增加会出现误差累积的情况. 而DLUKFNN并不改变UKF计算的核心过程, 因此, 误差累积依然存在.

$$\begin{vmatrix} 2 \times \mu_{k-1} \times \text{tr}(\hat{\Sigma}_k^{**}) \\ 2 \times \nu_{k-1} \times (-\text{tr}(K_k^{**} \hat{\Sigma}_{z,k} (K_k^{**})^T)) \end{vmatrix}, \quad (12)$$

式中: learn 表示学习率, k 表示时刻.

步骤5 使用 μ_k 和 ν_k 更新DLUKFNN估计状态变量的方差. 将 $\hat{\Sigma}_k^{**}$ 和 $-(K_k^{**} \hat{\Sigma}_{z,k} (K_k^{**})^T)$ 的对角线值分别写作向量 A_k 和 B_k . 使用步骤4得到的 μ_k 和 ν_k , 按照式(13)更新 A_k 和 B_k , 即

$$\begin{cases} A'_k = \mu_k^2 \times A_k, \\ B'_k = \nu_k^2 \times B_k, \end{cases} \quad (13)$$

式中 A'_k 和 B'_k 表示更新后的 A_k 和 B_k , 最终状态变量方差的更新应为

$$V_k = A'_k + B'_k. \quad (14)$$

为了和原DLUKFNN预测状态变量协方差区分, 将更新后的估计状态变量协方差写为 $\hat{\Sigma}_k^{**'}$.

注1 将向量 A_k 和 B_k 写为

$$A_k = \begin{vmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{vmatrix}_k, \quad B_k = \begin{vmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{vmatrix}_k, \quad (15)$$

取向量 A_k 和 B_k 中任意一个元素, 比如 a_1 和 b_1 , 按照式(12), 其更新方式写为

$$\begin{cases} a'_{1,k} = [\mu_{k-1} - \text{learn} \times 2 \times \mu_{k-1} \times \text{tr}(\hat{\Sigma}_k^{**})]^2 \times a_{1,k}, \\ b'_{1,k} = [\nu_{k-1} - \text{learn} \times 2 \times \nu_{k-1} \times \text{tr}(-(K_k^{**} \hat{\Sigma}_{z,k} (K_k^{**})^T))]^2 \times b_{1,k}, \end{cases} \quad (16)$$

式中 $a'_{1,k}$ 和 $b'_{1,k}$ 分别表示更新后的 $a_{1,k}$ 和 $b_{1,k}$, 按照式(14)将 $(\tilde{x}_1)^2$ 写为

$$\begin{aligned}
 (\tilde{x}_1)^2 &= a'_{1,k} + b'_{1,k} = \\
 &[\mu_{k-1} - \text{learn} \times 2 \times \mu_{k-1} \times \text{tr}(\hat{\Sigma}_k^{**})]^2 \times a_{1,k} + \\
 &[\nu_{k-1} - \text{learn} \times 2 \times \nu_{k-1} \times \text{tr}(-(K_k^{**} \hat{\Sigma}_{z,k} (K_k^{**})^T))]^2 \times b_{1,k},
 \end{aligned}
 \tag{17}$$

其中, 根据式(10)可知, b_1 为负值. 因此, 并不能保证更新后的 $(\tilde{x}_1)^2$ 始终为正. 为此, 需要对每个状态变量方差更新时加入大于0的约束条件, 保证每个状态变量方差更新过程中均大于0. 由于状态变量方差为预测协方差矩阵的对角线值, 也即该矩阵的特征值. 对于实对称矩阵来说, 如果矩阵特征值大于0, 也能保证矩阵的正定性.

注2 根据式(10)可知, 构造的代价函数为标准凸函数形式, 梯度将沿着梯度减小的方向变化. 因此, 能达到状态变量约束方差中“误差累积”的情况.

步骤6 重复步骤1至步骤5过程.

步骤7 当 $J(\mu_k, \nu_k)$ 达到预设极小值时, 停止步骤3至步骤5过程. 此外, 本文使用工业实况数据验证模型的有效性. 当采集的数据运行完毕, 模型停止运行.

综合上述策略, 图2给出了ADLUKFNN模型框架.

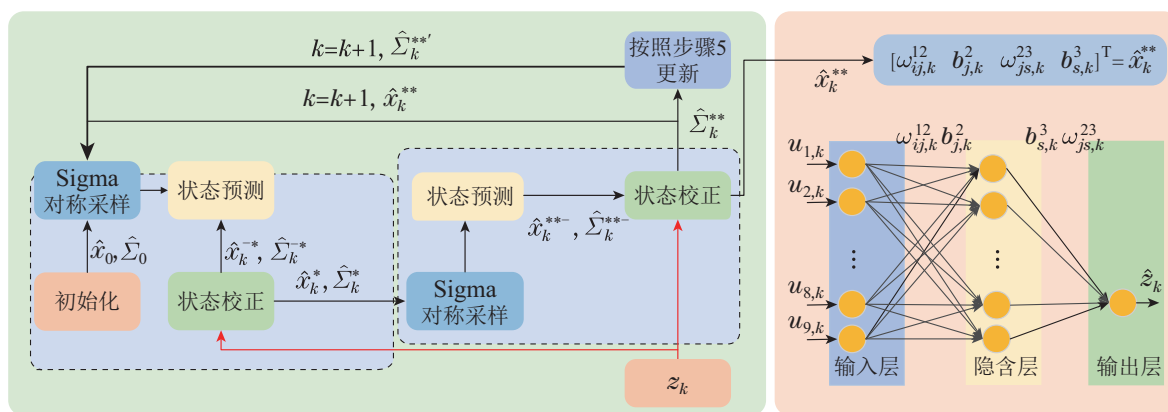


图2 ADLUKFNN算法框架

Fig. 2 Algorithm framework of ADLUKFNN

3 铝电解电流效率建模应用实验

3.1 实验对象

本文采用重庆某铝制造车间铝电解装备作为实验对象, 如图3所示. 电流效率定义为单位时间电解产出铝质量和按照法拉第定律计算的理论产出之比^[26], 实际生产过程中, 希望电流效率能够尽可能接近100. 使用数学模型优化槽控参数的前提是需要建立的数学模型实时指导求解最优槽控参数, 使得槽内始终保持物料和能量的平衡, 只有在这种情况下电流效率才会尽可能地高. 分析铝电解生产过程, 得到影响电流效率的主要槽控参数有系列电流、工作电压、打壳下料次数(number of break-Offs, NB)、分子比、铝水平、电解质水平、槽温度、实际出铝和氟化盐日用量^[14], 表1为截取的部分样本.

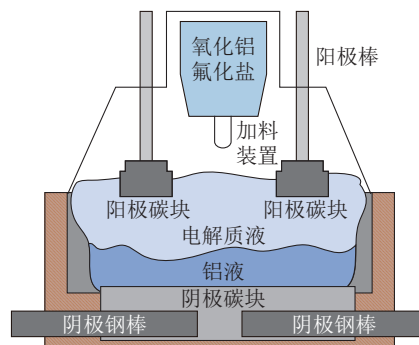


图3 电解槽结构示意图

Fig. 3 Schematic diagram of the structure of the aluminum electrolytic cell

3.2 动态建模

原始数据共1757组, 将数据采集过程中存在的两类明显无效异常值: 1) 电流效率小于10或大于100; 2) 大面积空白缺失, 作删除处理. 对氟化盐日用量缺失值, 采用插值法补齐^[27]. 按照该两种方式处理后数据总1539组. 为还原铝电解过程的复杂性和扰动性, 本实验将保留输入数据波动较大的值如表格中加粗部分. 同时, 也保留原始数据的时间顺序, 以此来验证模型对强时变系统的稳定性. 由于数据采集量纲不同, 对该1539组数据进行0-1归一化处理^[11, 28].

为验证本文所提ADLUKFNN方法的有效性, 依托上述实验对象, 在相同数据样本和仿真平台(Python 3.9)条件下, 同经典的反向传播人工神经网络(back propagation NN, BPNN)^[29]算法、PFNN和UKFNN开展对比实验. 实验所涉及的网络结构均构建为9-12-1形式. 因此, 可以计算得出NN的权值和阈值总参数为133个, 即状态变量维数为133. 其中, 9为网络输入, 在本文中则为筛选出的9种关键槽控参数, 12为隐含层神经元个数, 1为输出层神经元个数. 隐含层激活函数设置为sigmoid, 输出层激活函数设置为relu. 其中,

PFNN分别使用100, 300和500个粒子观察粒子数量同模型精确度的联系, 由PFNN(100), PFNN(300)和PFNN(500)表示. 此外, 将上述数据集前1489组数据作为训练集, 后50组数据作为测试集.

表1 电解槽数据样本
Table 1 Sample of electrolytic cell data

槽控参数	范围	样本				
		1	2	3	...	1539
系列电流/A	1214~1754	1649	1720	1720	...	1720
工作电压/mv	3395~4809	3785	3719	4211	...	4108
NB次数	307~1082	667	707	639	...	697
分子比	2.34~2.95	3	3	2	...	3
铝水平/cm	11~26	14	12	18	...	22
电解质水平/cm	6~29	18	19	19	...	17
槽温/°C	927~1006	958	946	964	...	955
实际出铝/kg	402~1339	1270	1230	1187	...	1266
氟化盐日用量/kg	1.8~307	11	33	50	...	25

3.3 结果和讨论

本文中, 构建的NN结构输出层节点为1, 即模型的预测电流效率值. 因此, 使用拟合评价指标衡量各模型预测性能. 文中, 平均绝对误差(mean absolute error, MAE)、误差平方和(sum of squared errors, SSE)和相对误差(relative error, RE)评价模型预测精度, 确定系数(correlation of coefficient, R)衡量模型拟合能力.

图4给出了BPNN, PFNN, UKFNN在测试集中的拟合效果. BPNN基于反向传播算法, 使用训练集获得一组最优权值和阈值. 从图4(a)看出, BPNN在测试集中的拟合效果并不理想. 究其原因当系统时变或遭受扰动后, 该组使用历史数据(训练集)获得的最优权值和阈值便不再满足本时刻预测要求. 此外, 图4(b)–(d)显示, PFNN随着粒子数量的增加其拟合效果逐步提升, UKFNN采用固定粒子数近似系统分布, 根据上述状态变量维度可以计算得出其粒子数为 $133 \times 2 + 1 = 267$, 即当PFNN粒子数为300时, 应和UKFNN拟合效果相当. 然而, 图4和表2显示, 当PFNN粒子数为300时, UKFNN的拟合效果优于PFNN. 此外, PFNN粒子数为500时, 和UKFNN拟合效果相当, 但表2显示, 其计算耗时明显高于UKFNN. 由于PFNN的重要性函数选取仍是目前研究的难点问题, 本文使用最优重要性函数^[30]构建PFNN, 该方法缺乏现场值对重要性函数的修正, 其实验结果并不稳定. 综合考虑铝电解过程建模时效和精度要求, 本文研究基于UKFNN铝电解过程电流效率预测模型构建方法.

表1截取实验数据样本显示, 铝电解过程蕴含较大的异常扰动. 因此, 需要模型即使面临异常扰动时仍能保持较高的预测性能, 则要求状态变量(NN的权值

和阈值)跟随工况进行自适应调整. 而DLUKFNN的主要特点是采用两次现场量测值修正状态变量, 故较UKFNN而言, 其状态变量更符合系统当前时刻特点. 与此同时, DLUKFNN将上层滤波结果作为下层滤波初始信息, 也能解决状态空间方程一步预测过程未包含系统时变信息的弊端. 图5给出了DLUKFNN和ADLUKFNN在测试集上的拟合效果, 图5(a)显示, DLUKFNN在测试集上拟合效果较好. 此外, ADLUKFNN对DLUKFNN状态变量的预测均方误差构建了一种代价函数, 使用梯度下降法迫使其值逼近于较小的理想值. 图5(c)给出了两者的MAE, 图5(c)显示, ADLUKFNN相较于DLUKFNN其MAE得到了进一步降低.

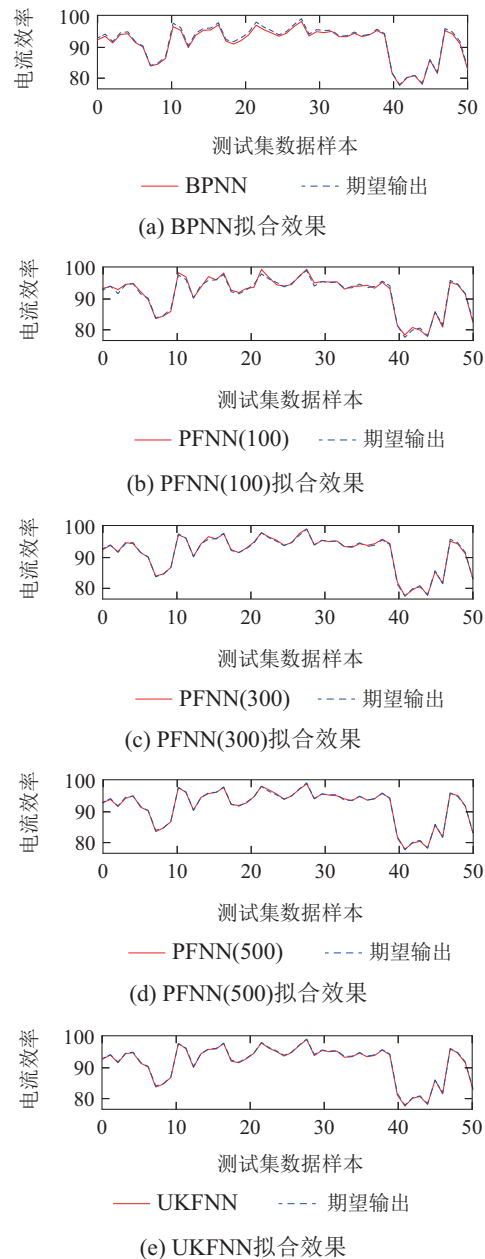


图4 不同模型拟合效果

Fig. 4 The fitting effect of different models

表 2 不同模型相关性能指标比较

Table 2 The comparison of related performance indicators from different models

模型	误差统计指标						
	Average	Max	Min	MAE	SSE	R	运行时间/s
BPNN	0.0919	3.1930	-3.9549	0.5976	1287.8678	0.9732	0.5474
PFNN(100)	0.0188	6.1252	-2.4015	0.4474	594.7745	0.9876	9.4547
PFNN(300)	0.0134	3.6343	-3.2487	0.2534	207.9112	0.9957	23.5237
PFNN(500)	-0.0097	0.7003	-2.9632	0.1887	97.2059	0.9980	40.5895
UKFNN	-0.1838	1.2565	-1.1487	0.2000	83.0534	0.9983	31.1837
DLUKFNN	-0.0572	0.0000	-0.0975	0.0572	5.0780	0.9999	61.2604
ADLUKFNN	-0.0489	0.0000	-0.0673	0.0489	3.7001	0.9999	67.9314

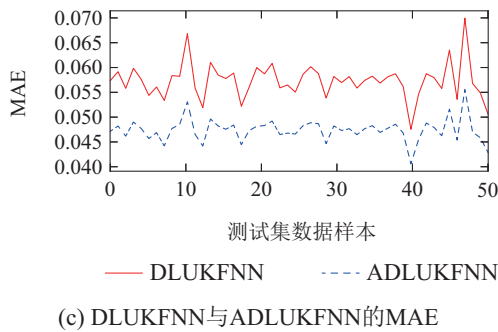
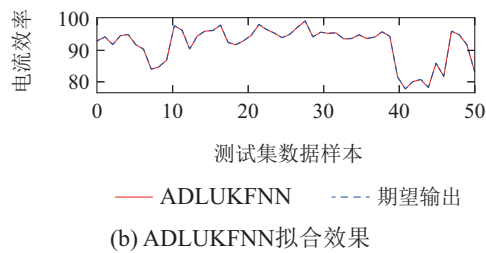
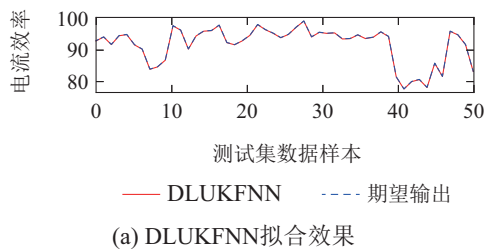


图 5 DLUKFNN和ADLUKFNN的拟合效果

Fig. 5 The fitting effect of DLUKFNN and ADLUKFNN

PFNN和UKFNN的现场量测值会对状态变量不断修正,原则上两者在测试集和训练集上均能获得良好的预测性能.因此,表2和图6分析了各模型所有样本的预测性能.表2显示,BPNN运行时间最短.但是其MAE和SSE最大,R最小.因此,预测误差较大,BPNN不能满足铝电解过程高精度建模需求.PFNN随着粒子数量的增加,其拟合效果逐步提高,当粒子数量为500时,其MAE稍小UKFNN.但是,其SSE大于UKFNN,R稍小于UKFNN,即PFNN存在部分预测值更大偏离实际值得情况.同时,在时效性上,UKFNN表现出更大的优势.

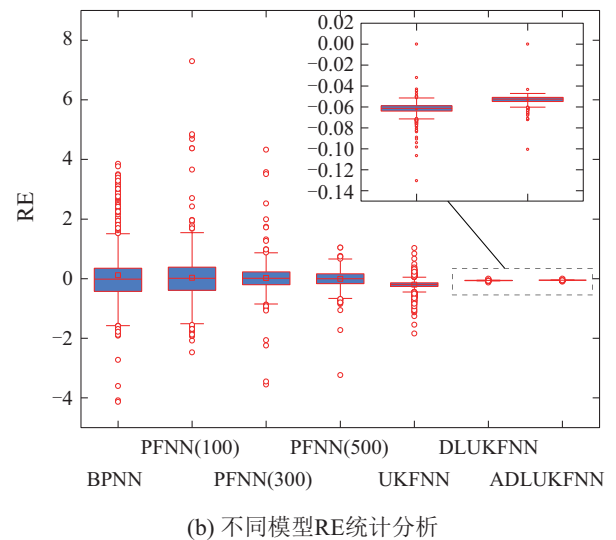
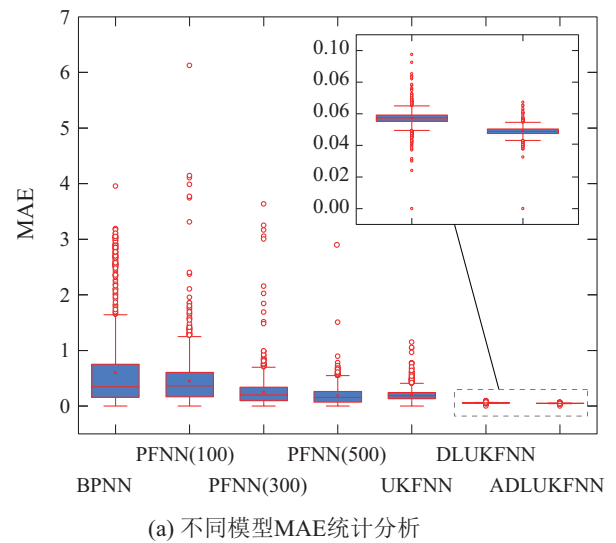


图 6 不同模型的误差统计分析

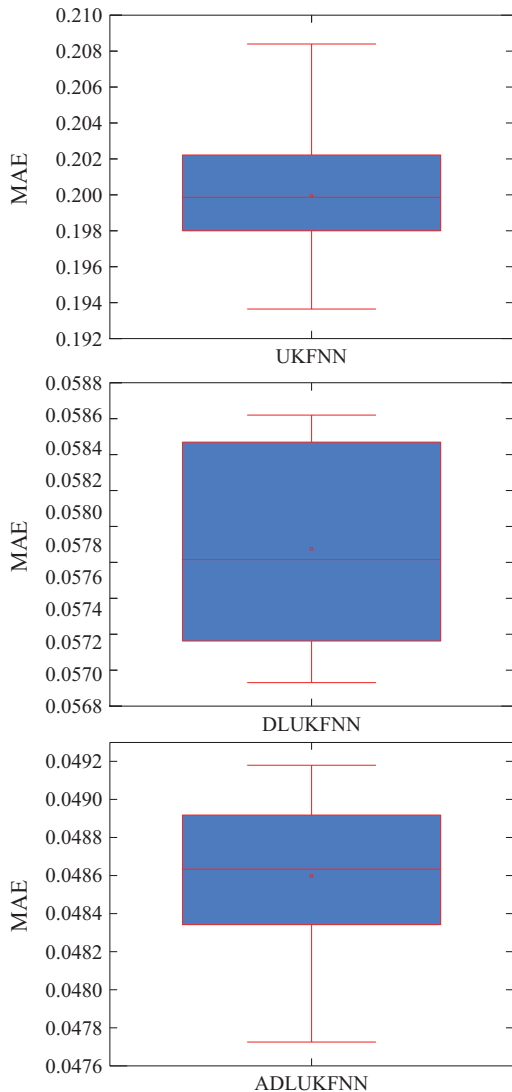
Fig. 6 Statistical analysis of errors in different models

表2显示,DLUKFNN相较于UKFNN其精确度得到了较大的提升,其MAE降低了约3.5倍,SSE则降低了约16.4倍.此外,对DLUKFNN采用误差自适应策略后,运行时间在可接受的范围内有一定增加,但是,其MAE降低了约1.2倍,SSE降低了约1.4倍.同时,ADLUKFNN在各模型中的误差最大值(Max)和误差最

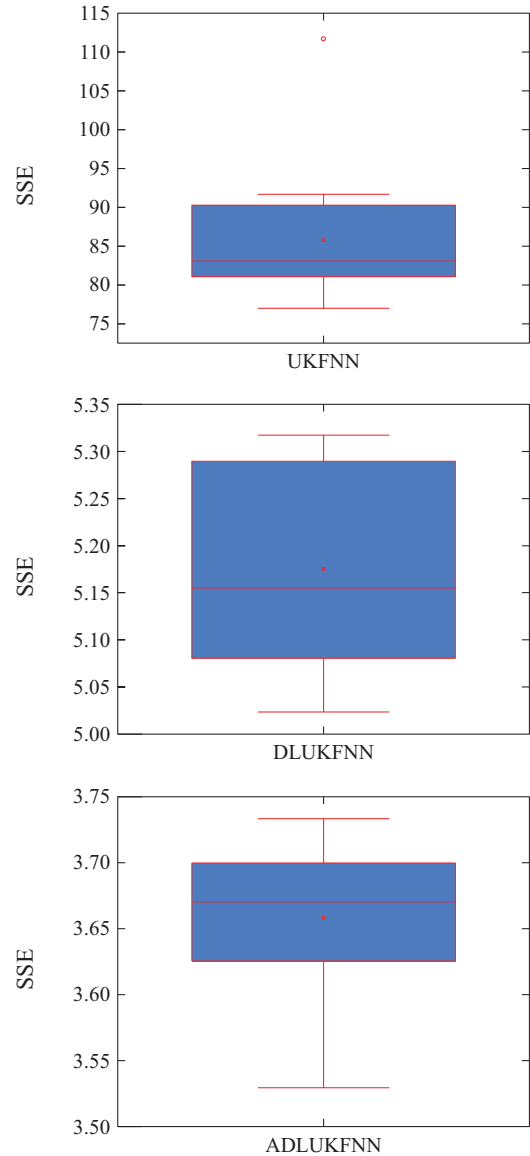
小值(Min)之间距离最短,即误差集中于较小的范围内.图6显示,DLUKFNN和ADLUKFNN将所有样本的MAE分布于0.0至0.1之间,集中于0.05至0.06左右.RE分布于-0.14至0.02之间,集中于-0.06至-0.04左右.同时,图6显示,ADLUKFNN相较于DLUKFNN其误差均有一定程度的降低.综合上述,ADLUKFNN预测性能最佳.

此外,由于网络的初始值是随机产生的,因此,每次实验结果具有一定的随机性.为了避免上述实验结果的偶然性,对各模型开展15次独立随机实验.本文采用最优重要性函数构建PFNN,该方法由于缺乏现场值对重要性函数的修正.因此,其实验结果并不稳定.BPNN基于训练集获得最优权值和阈值后不再改变.因此,不再对PFNN和BPNN开展独立随机实验.图7统计了各模型的MAE和SSE.

图7(a)显示,15次随机实验中,DLUKFNN的MAE最低可以达到0.0568左右,最高不超过0.0588.而ADLUKFNN的MAE最低能达到0.0476左右,最高不超过0.0492.



(a) 各模型随机实验MAE分布



(b) 各模型随机实验SSE分布

图7 随机实验分析

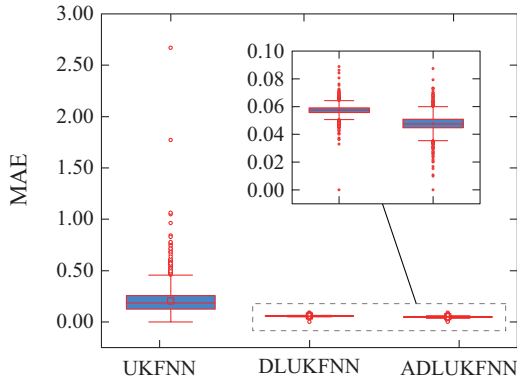
Fig. 7 Analysis of random experiments

图7(b)显示,DLUKFNN的SSE最低能达到5.00左右,最高不超过5.35,而ADLUKFNN的SSE最低能达到3.50,最高不超过3.75.由此可见,即使在随机实验下,ADLUKFNN的预测误差并没有产生过大的波动.

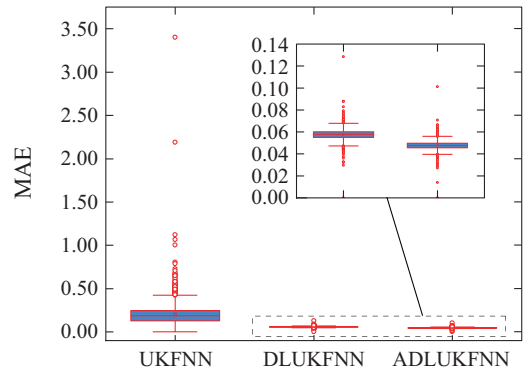
铝电解生产过程具有较多的人为操作,如阳极更换和出铝等^[12].因此,槽内具有较强的时变性和扰动性.为了进一步验证所提模型对铝电解生产系统时变和扰动的自适应性,对UKFNN,DLUKFNN和ADLUKFNN的输入数据按照 $X = X + d \times \text{rand}(m, n) \times X$ 的方式施加人为扰动.其中: X 为输入数据,rand为 $(-1, 1)$ 均匀分布的随机数, m 为1539, n 为9, d 是扰动因子.本文中, d 分别取5%,10%,15%,20%,25%,30%,35%,40%,45%,50%.图8显示,当面临不同程度的人为干扰时,DLUKFNN相较于UKFNN,能够将MAE压缩至更小范围之内,另外,ADLUKFNN相

较于DLUKFNN也将MAE压缩至更小的范围之内. 因此, 即使ADLUKFNN面临不同程度的人为干扰时, 依

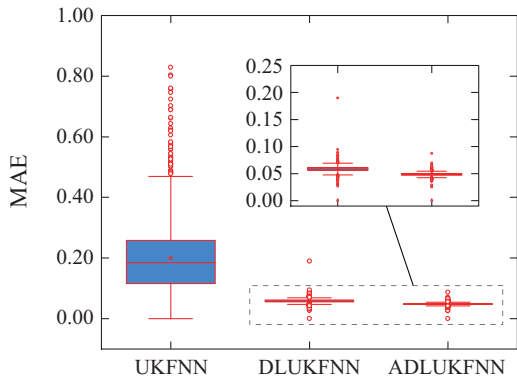
然能保持最高的预测性能, 再次证明了本方法的有效性.



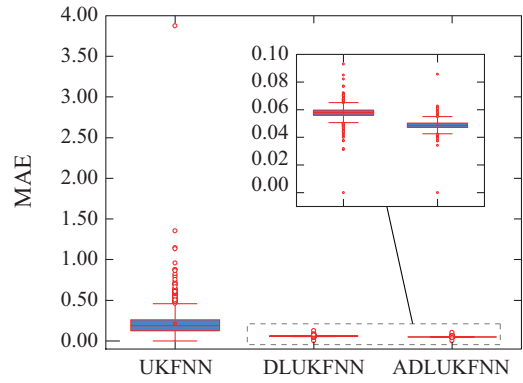
(a) 施加扰动±5%



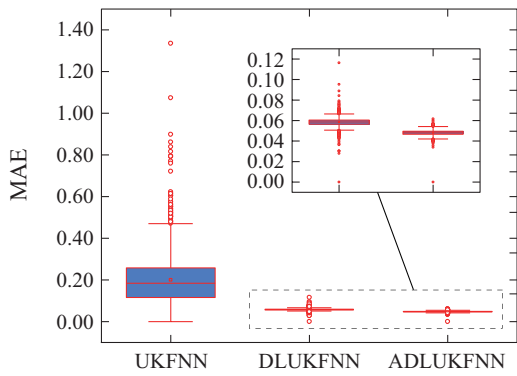
(b) 施加扰动±10%



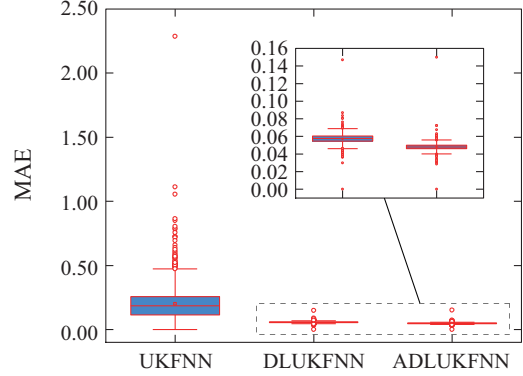
(c) 施加扰动±15%



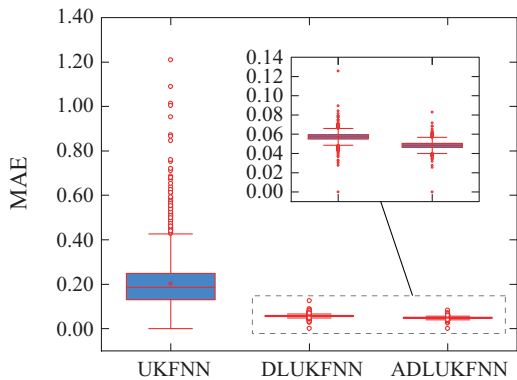
(d) 施加扰动±20%



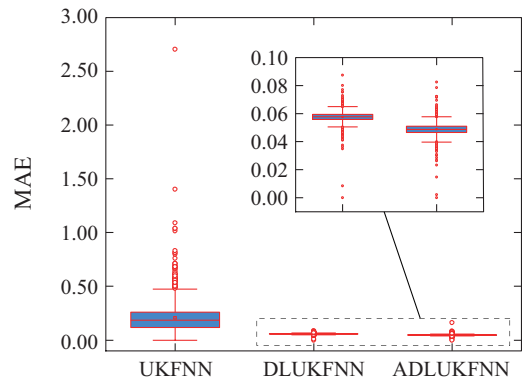
(e) 施加扰动±25%



(f) 施加扰动±30%



(g) 施加扰动±35%



(h) 施加扰动±40%

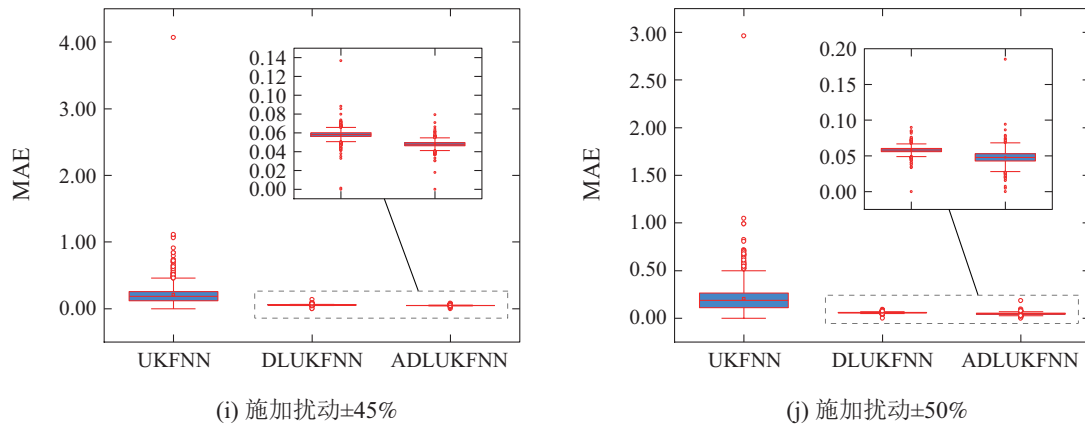


图8 扰动实验MAE统计分析

Fig. 8 MAE statistical analysis of perturbation experiment

4 结论

本文旨在建立一种稳定且可靠的铝电解过程电流效率预测模型. 首先, 提出采取双层结构的无迹卡尔曼滤波更新人工神经网络权值和阈值, 以此来提高模型对铝电解这种强扰动和强时变系统的稳定性. 然后, 以DLUKFNN状态变量的均方误差最小为目标, 构建了一种自适应策略. 以此来削弱DLUKFNN中误差累积对模型的影响, 进一步提高模型的精确度. 最后, 本文使用上述策略搭建了铝电解过程电流效率动态预测模型, 使用拟合实验、随机实验和人为扰动实验证实了DLUKFNN和ADLUKFNN具有对扰动系统的稳定性和预测准确性.

在后续研究工作中, 将进一步探索精度和计算复杂度平衡的难题. 同时, 将进一步考虑融合能耗、电流效率、碳排放等生产与环保指标, 建立一个平衡多指标的系统优化模型, 并探索多指标预测问题中精度和计算复杂度平衡的难题.

参考文献:

- [1] GUI Weihua, YUE Weichao, CHEN Xiaofang, et al. Process industry knowledge automation and its applications in aluminum electrolysis production. *Control Theory & Applications*, 2018, 35(7): 887 – 899. (桂卫华, 岳伟超, 陈晓方, 等. 流程工业知识自动化及其在铝电解生产中的应用. *控制理论与应用*, 2018, 35(7): 887 – 899.)
- [2] PENG J P, SONG Y, DI Y Z, et al. Towards improved current efficiency of hall-héroult cells by using a novel cathode and process parameters. *JOM*, 2019, 71(2): 508 – 513.
- [3] YI J, BAI J J, ZHOU W, et al. Operating parameters optimization for the aluminum electrolysis process using an improved quantum-behaved particle swarm algorithm. *IEEE Transactions on Industrial Informatics*, 2018, 14(8): 3405 – 3415.
- [4] LIN Qingyang, CHEN Xiaofang, XIE Yongfang. An superheat identification method in aluminum electrolysis based on residual convolutional self-attention neural network. *Journal of Northeastern University (Natural Science)*, 2023, 44(1): 8 – 17. (林清扬, 陈晓方, 谢永芳. 基于残差卷积自注意力神经网络的铝电解过热度识别方法. *东北大学学报(自然科学版)*, 2023, 44(1): 8 – 17.)

- [5] YIN Gang, CHEN Gen, HE Wen, et al. Anode effect prediction of 300kA aluminium electrolysis cell based on deep learning. *The Chinese Journal of Nonferrous Metals*, 2021, 31(1): 161 – 170. (尹刚, 陈根, 何文, 等. 基于深度学习的300kA铝电解槽阳极效应预测. *中国有色金属学报*, 2021, 31(1): 161 – 170.)
- [6] LEI Y X, KARIMI H R, CHEN X F. A novel self-supervised deep LSTM network for industrial temperature prediction in aluminum processes application. *Neurocomputing*, 2022, 502: 177 – 185.
- [7] WANG J, XIE Y F, XIE S W, et al. Optimization of aluminum fluoride addition in aluminum electrolysis process based on pruned sparse fuzzy neural network. *ISA Transactions*, 2023, 133: 285 – 301.
- [8] CHIANG Y M, CHANG L C, CHANG F J, et al. Comparison of static-feedforward and dynamic-feedback neural networks for rainfall-runoff modeling. *Journal of Hydrology*, 2004, 290(3/4): 297 – 311.
- [9] XIE Yongfang, LI Li, XIE Shiwen, et al. Method of root cause analysis based on semi-quantitative probabilistic graphical model. *Control Theory & Applications*, 2023, 40(3): 419 – 429. (谢永芳, 李理, 谢世文, 等. 基于半定量概率图模型的溯因分析方法. *控制理论与应用*, 2023, 40(3): 419 – 429.)
- [10] RUBIO J J, WEN Y. Nonlinear system identification with recurrent neural networks and dead-zone Kalman filter algorithm. *Neurocomputing*, 2007, 70(13/15): 2460 – 2466.
- [11] YAO L Z, LI T F, LI Y Y, et al. An improved feed-forward neural network based on UKF and strong tracking filtering to establish energy consumption model for aluminum electrolysis process. *Neural Computing and Applications*, 2019, 31: 4271 – 4285.
- [12] LI Taifu, YAO Lizhong, YI Jun, et al. An improved UKFNN based on square root filter and strong tracking filter for dynamic evolutionary modeling of aluminum reduction cell. *Acta Automatica Sinica*, 2014, 40(3): 522 – 530. (李太福, 姚立忠, 易军, 等. 强跟踪平方根UKFNN的铝电解槽工耗动态演化模型. *自动化学报*, 2014, 40(3): 522 – 530.)
- [13] ZHOU W, LI X L, YI J, et al. A novel UKF-RBF method based on adaptive noise factor for fault diagnosis in pumping unit. *IEEE Transactions on Industrial Informatics*, 2022, 15(3): 1415 – 1424.
- [14] YAO L Z, DING W, HE T T, et al. A multiobjective prediction model with incremental learning ability by developing a multi-source filter neural network for the electrolytic aluminium process. *Applied Intelligence*, 2022, 52: 17387 – 17409.
- [15] DING W, YAO L Z, LI Y Y, et al. Incremental learning model based on an improved CKS-PFNN for aluminium electrolysis manufacturing. *Neural Computing and Applications*, 2022, 34: 2083 – 2102.
- [16] DING W, YAO L Z, LI Y Y, et al. Dynamic evolutionary model based on a multi-sampling inherited HAPFNN for an aluminium elec-

- trolysis manufacturing system. *Applied Soft Computing*, 2021, 99: 106925.
- [17] ZHANG J, YANG X H, ZHANG W A. A Progressive bayesian filtering framework for nonlinear systems with heavy-tailed noises. *IEEE Transactions on Automatic Control*, 2023, 68(3): 1918 – 1925.
- [18] TIAN M C, BO Y B, CHEN Z M, et al. Multi-target tracking method based on improved firefly algorithm optimized particle filter. *Neuro-computing*, 2019, 359: 438 – 448.
- [19] ZHANG Z R, HUANG C Q, DING D. et al. Hummingbirds optimization algorithm-based particle filter for maneuvering target tracking. *Nonlinear Dynamics*, 2019, 97(2): 1227 – 1243.
- [20] DENG Z Y, WANG H B, WANG R X. An improved unscented Kalman filter for interrupted and drift sensor faults of aircrafts. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1 – 10.
- [21] HAO Gang, YE Xiufen, CHEN Ting. Weighted measurement fusion algorithm for nonlinear unscented Kalman filter. *Control Theory & Applications*, 2011, 28(6): 753 – 758.
(郝钢, 叶秀芬, 陈亭. 加权观测融合非线性无迹卡尔曼滤波算法. 控制理论与应用, 2011, 28(6): 753 – 758.)
- [22] KNUDSEN T, LETH J. A new continuous discrete unscented Kalman filter. *IEEE Transactions on Automatic Control*, 2019, 64(5): 2198 – 2205.
- [23] CAI Anjiang, LIU Kaifeng, GUO Shihong, et al. Quaternion derivative unscented Kalman filter-based twostep attitude estimation algorithm for multi-rotor unmanned aerial vehicle. *Control Theory & Applications*, 2020, 37(2): 365 – 373.
(蔡安江, 刘凯峰, 郭师虹, 等. 基于四元数衍生无迹卡尔曼滤波的二段式多旋翼无人机姿态估计算法. 控制理论与应用, 2020, 37(2): 365 – 373.)
- [24] LIU Q, TIAN Y M, CHAI Y I, et al. Design of unscented Kalman filter based on the adjustments of the number and placements of the sampling points. *ISA Transactions*, 2021, 108: 188 – 195.
- [25] YANG Feng, ZHENG Litao, WANG Jiaqi, et al. Double layer unscented Kalman filter. *Acta Automatica Sinica*, 2019, 45(7): 1386 – 1391.
(杨峰, 郑丽涛, 王家琦, 等. 双层无迹卡尔曼滤波. 自动化学报, 2019, 45(7): 1386 – 1391.)
- [26] YI J, BAI J J, ZHOU W, et al. Operating parameters optimization for the aluminum electrolysis process using an improved quantum-behaved particle swarm algorithm. *IEEE Transactions on Industrial Informatics*, 2018, 14(8): 3405 – 3415.
- [27] WANG Lanwei, ZHANG Xingguo, ZHANG Yu, et al. Application of fractal interpolation method to geo-electric field interference data process. *Acta Seismologica Sinica*, 2021, 43(3): 350 – 358.
(王兰炜, 张兴国, 张宇, 等. 分形插值法在地电场干扰数据处理中的应用. 地震学报, 2021, 43(3): 350 – 358.)
- [28] CUI L G, TAO Y Q, DENG J, et al. BBO-BPNN and AMPSO-BPNN for multiple-criteria inventory classification. *Expert Systems with Applications*, 2021, 175: 114842.
- [29] HASSANIEN A E, ABOUELMAGD L M, MAHMOUD A S, et al. An optimized backpropagation neural network models for the prediction of nanomaterials concentration for purification industrial wastewater. *Engineering Applications of Artificial Intelligence*, 2023, 126: 107010.
- [30] DOUCET A, GODSILL S, ANDRIEU C. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 2000, 10(1): 197 – 208.

作者简介:

方小燕 硕士研究生, 主要从事铝电解过程建模研究, E-mail: fangxiaoyan1009@163.com;

姚立忠 博士, 副教授, 主要从事铝电解过程建模与多目标优化研究, E-mail: lizhong_yao@cqnu.edu.cn;

罗海军 博士, 教授, 主要从事人工智能、系统测试研究, E-mail: luohaijun@cqnu.edu.cn;

张玉泽 硕士研究生, 主要从事流程工艺建模研究, E-mail: xiaowo970510@163.com;

易军 博士, 教授, 主要从事铝电解过程建模与多目标优化研究, E-mail: laoyifrcq@163.com.