

# 基于 SCKF 的 Elman 递归神经网络在软测量建模中的应用

李军, 桑桦

兰州交通大学自动化与电气工程学院, 甘肃 兰州 730070

基金项目: 国家自然科学基金资助项目(51467008); 光电技术与智能控制教育部重点实验室(兰州交通大学)开放课题(KFKT2016-3)

通信作者: 李军, lijun691201@mail.lzjtu.cn 收稿/录用/修回: 2016-04-25/2016-06-29/2016-07-11

## 摘要

针对具有强非线性、复杂的化工过程软测量建模, 提出一种基于平方根容积卡尔曼滤波(SCKF)的递归神经网络方法. 基于 Elman 递归神经网络, 首先构建状态空间模型, 然后应用 SCKF 算法进行训练, 所有网络的权值将作为系统的状态进行更新. 容积卡尔曼滤波(CKF)通过三阶 Spherical-Radial 容积准则生成容积点, 利用容积点逼近状态的后验分布, 使得高维非线性滤波中的多变量积分数值求解成为可能. 在 CKF 的基础上, SCKF 采用预测及后验误差协方差矩阵的平方根因子进行递推运算, 进一步改进了算法的数值稳定性. 将该方法应用于脱丁烷塔底部丁烷组分含量以及硫回收装置尾气中  $\text{SO}_2$  和  $\text{H}_2\text{S}$  含量的软测量动态建模实例中, 在同等条件下, 还与基于 EKF、SCKF 的前馈神经网络, 基于 EKF 的递归神经网络等其它方法对比. 结果表明, 本文的方法能够获得很好的建模精度, 显示出其有效性.

## 关键词

平方根容积卡尔曼滤波  
递归神经网络  
软测量  
建模

中图分类号: TP274

文献标识码: A

## Elman Recurrent Neural Network Method Based on SCKF Algorithm and Its Application to Soft Sensor Modeling

LI Jun, SANG Hua

School of Automation & Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

## Abstract

To achieve soft sensor modeling in a chemical process with strong nonlinearity and complexity, a recurrent neural network method based on square-root cubature Kalman filter (SCKF) training algorithm is proposed. The state-space model of the Elman recurrent neural network is first established, and then the network is trained by the SCKF algorithm; all its weights are considered the state of the system to be updated. Cubature Kalman filter (CKF) generates cubature points by third-order spherical-radial standards and then uses cubature points to approximate the posterior distribution of the state, which makes computing the numerical solution of the multivariate integral of high-dimensional nonlinear filtering possible. On the basis of CKF, SCKF propagates the square root factor of prediction and posteriori error covariance matrix to further improve the numerical stability of the algorithm. The employed method is applied to instances of soft sensor modeling, which include the estimation of the butane concentration in the bottom flow of a debutanizer column and the estimation of the concentrations of  $\text{SO}_2$  and  $\text{H}_2\text{S}$  in sulfur recovery unit tail gas composition. Compared with feedforward neural network method based on extended Kalman filter (EKF) and the SCKF algorithm and recurrent neural network method based on the EKF algorithm, the proposed method can obtain better modeling estimation accuracy under the same condition. Experimental results demonstrate the effectiveness of the proposed method.

## Keywords

square-root cubature Kalman filter (SCKF);  
recurrent neural network;  
soft sensor;  
modelling

## 1 引言

软测量建模是目前化工过程控制领域的热点研究方向<sup>[1-4]</sup>, 对于具有复杂、强非线性特性的动态系统, 它能

够有效地改善和解决生产过程中在线分析仪测量滞后大、价格昂贵、维护保养复杂等问题, 满足实际生产过程中实时性的需求. 目前, 神经网络、支持向量机(support vector machine, SVM)等计算智能方法是主要的建模工具,

在基于数据驱动的软测量建模<sup>[5-9]</sup>中已取得了成功的应用. 与前馈神经网络(feedforward neural network, FNN)相比, 递归神经网络(recurrent neural network, RNN)通过合并网络节点的反馈连接, 具有动态记忆特性, 能够处理具有时延信息的数据, 具有更好的潜力. 文[10]针对氧化铝生产过程中铝酸钠溶液组分的浓度建模, 提出一种基于稳定学习的递归神经网络动态 PLS(partial least square)软测量建模方法, 取得了很好的效果.

RNN 通常采用 BPTT(back propagation through time)、RTRL(real time recurrent learning)学习算法, 收敛速度慢, 易陷入局部最优. 为了加快网络的收敛速度, 文[11-12]给出了无微分的训练学习算法——扩展卡尔曼滤波(EKF)算法, 与 BPTT、RTRL 算法相比, 虽然计算复杂度较高, 但是它显示出更好的学习性能和更强的鲁棒性. 但是, 对于高度非线性系统, 由于 EKF 算法的一阶线性逼近特性, 会导致算法出现“发散”情形, 而且其仅适用于可微函数. 文[13-14]在贝叶斯滤波的框架下, 基于 Spherical-Radial 容积准则, 提出一种容积卡尔曼滤波(CKF)算法, 其在高斯分布的假设下, 将贝叶斯滤波问题的求解转化为非线性函数与高斯概率密度函数的多维积分求解问题, CKF 计算效率高. 容积规则的选择是能够有效求解这类问题的关键. 文[15]将 CKF 算法用于前馈神经网络的训练, 取得较好的训练效果. 但是, CKF 算法在运行过程中无法保证误差协方差矩阵的对称性和正定性, 这会使得算法出现发散现象. 为避免该问题, 平方根容积卡尔曼滤波(SCKF)<sup>[14, 16]</sup>算法被提出, 有效改进了 CKF 算法的数值稳定性.

针对化工过程软测量建模, 本文采用 SCKF 算法训练 Elman 递归神经网络, 并将其应用于脱丁烷塔底部 C4(丁烷)组分浓度的预测和硫回收装置(SRU)中硫化氢(H<sub>2</sub>S)及二氧化硫(SO<sub>2</sub>)气体浓度的估计的实例实验中, 在同等条件下, 还将与基于 BPTT、RTRL 和 EKF 算法训练的 Elman 递归网络, 基于 BPTT、RTRL、EKF 和 SCKF 算法训练的全连接递归神经网络(fully-connected recurrent neural network, FCRNN)<sup>[11]</sup>, 基于 EKF 和 SCKF 算法的多层感知器(multi-layer perceptron, MLP)方法进行比较, 以验证本文方法的有效性.

## 2 简单递归神经网络

Elman 神经网络作为一阶局部 RNN, 是一种简单递归神经网络(SRN), 其拓扑结构如图 1 所示.

SRN 主要由输入层 U、隐含层 R、联系层 C 及输出层 Y 组成, 与前馈神经网络的连接形式相同, 输入层、隐含层和输出层的神经元分别通过权值  $W^{RU}$  和  $W^{YR}$  全连接, 时延连接将隐含层的递归神经元在当前  $k$  时刻的激活函数输出值反馈给联系层, 满足  $c_k = v_{k-1}$ . 因此, 每一个递归单元通过权值  $W^{RC}$  均接受来自前一时刻的所有递归单元的激活函数输出作为输入, 来自前一时刻的递归单元的激活输出可视为递归层输入的扩展, 这使得 SRN 具有短时记忆功能.

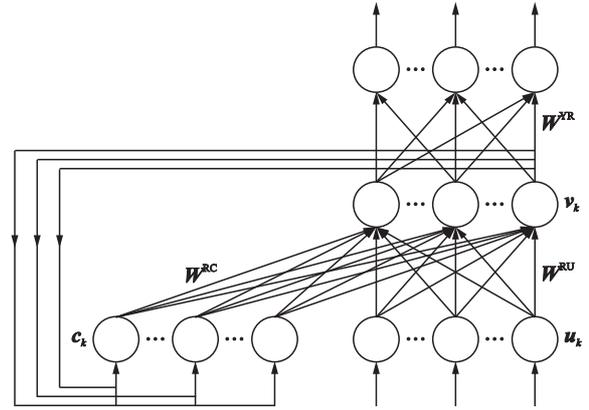


图 1 SRN 拓扑结构图

Fig.1 The topology structure of SRN

给定 SRN 网络在  $k$  时刻的外部输入,  $u_k = (u_{1,k}, \dots, u_{j,k}, \dots, u_{N_i,k})$ ,  $u_k \in \mathbb{R}^{N_i}$ , 隐含层的递归神经元输出  $v_k = (v_{1,k}, \dots, v_{j,k}, \dots, v_{N_v,k})$ ,  $v_k \in \mathbb{R}^{N_v}$ , 则隐含层的第  $i$  个单元输入  $\tilde{v}_{i,k}$  和激活输出  $v_{i,k}$  可以表示为

$$\tilde{v}_{i,k} = \sum_{j=1}^{N_i} W_{ij}^{RU} u_{j,k} + \sum_{j=1}^{N_v} W_{ij}^{RC} v_{j,k-1} \quad (1)$$

$$v_{i,k} = f(\tilde{v}_{i,k}) \quad (2)$$

输出层的神经元输出  $y_k \in \mathbb{R}^{N_o}$ , 其第  $i$  个单元的输出  $y_{i,k}$  可表示为

$$y_{i,k} = f\left(\sum_{j=1}^{N_v} W_{ij}^{YR} v_{j,k}\right) \quad (3)$$

其中,  $f(\cdot)$  表示激活函数, 通常使用 Sigmoid 函数、双曲正切函数. 为进一步提高网络的学习性能, 本文选取如下的特殊双曲正切函数<sup>[11]</sup>:

$$f(x) = 1.7159 \tanh\left(\frac{2}{3}x\right) \quad (4)$$

## 3 基于 SCKF 的 SRN 训练算法

采用 SCKF 算法对 SRN 进行训练, 将训练问题作为一个非线性动力系统的动态参数估计问题来处理, 即将网络的连接权值作为 SCKF 的状态, 随着时序  $k$  不断对网络的权值参数进行更新, 使得网络的实际输出与期望输出之间的均方误差随着时序的增加而逐渐减小.

由图 1 所示的 SRN, 定义  $k$  时刻由  $W^{RU}$ 、 $W^{RC}$  及  $W^{YR}$  中的矩阵元素按列向量的次序重组展开, 所形成的向量  $w \in \mathbb{R}^{N_w}$  为权值状态向量.

此时, 网络的非线性状态空间模型可表示为

$$w_k = w_{k-1} + q_{k-1} \quad (5)$$

$$y_k = h(w_k, u_k, v_k) + r_k \quad (6)$$

其中,  $w_k$  是维数为  $N_w = N_v \times (N_i + N_v) + N_o \times N_v$  的权向量; 过程噪声向量  $q_{k-1}$ 、量测噪声  $r_k$  假定服从具有零均值, 协方差矩阵分别为  $Q_{k-1}$ 、 $R_k$  的高斯噪声过程. 另外, 式(5)人工添加的过程噪声有助于避免算法训练时, 目标函数陷入局部极小.

若给定输入—输出样本集  $D_k = \{u_i, y_i\}_{i=1}^k$ , 在高斯假

设的框架下,从贝叶斯滤波的角度分析网络的监督学习过程,存在如下滤波似然密度函数:

$$p(\mathbf{y}_k | D_{k-1}) = N(\mathbf{y}_k; \hat{\mathbf{y}}_{k|k-1}, \mathbf{P}_{yy, k|k-1}) \quad (7)$$

其中,  $N(\cdot, \cdot)$  为高斯分布的记号. 一步预测的量测值和相应的协方差矩阵为

$$\hat{\mathbf{y}}_{k|k-1} = \int_{\mathbb{R}^{N_w}} \mathbf{h}(\mathbf{w}_k, \mathbf{u}_k, \mathbf{v}_k) N(\mathbf{w}_k; \hat{\mathbf{w}}_{k|k-1}, \mathbf{P}_{k|k-1}) d\mathbf{w}_k \quad (8)$$

$$\mathbf{P}_{yy, k|k-1} = \int_{\mathbb{R}^{N_w}} \mathbf{h}(\mathbf{w}_k, \mathbf{u}_k, \mathbf{v}_k) \mathbf{h}^T(\mathbf{w}_k, \mathbf{u}_k, \mathbf{v}_k) N(\mathbf{w}_k; \hat{\mathbf{w}}_{k|k-1}, \mathbf{P}_{k|k-1}) \cdot d\mathbf{w}_k - \hat{\mathbf{y}}_{k|k-1} \hat{\mathbf{y}}_{k|k-1}^T + \mathbf{R}_k \quad (9)$$

状态和量测变量的联合条件高斯密度函数可表示为

$$p([\mathbf{w}_k^T \ \mathbf{y}_k^T]^T | D_{k-1}) = N\left(\begin{pmatrix} \hat{\mathbf{w}}_{k|k-1} \\ \hat{\mathbf{y}}_{k|k-1} \end{pmatrix}, \begin{pmatrix} \mathbf{P}_{k|k-1} & \mathbf{P}_{wy, k|k-1} \\ \mathbf{P}_{wy, k|k-1}^T & \mathbf{P}_{yy, k|k-1} \end{pmatrix}\right) \quad (10)$$

其中,量测与状态变量的互协方差为

$$\mathbf{P}_{wy, k|k-1} = \int_{\mathbb{R}^{N_w}} \mathbf{w}_k \mathbf{h}^T(\mathbf{w}_k, \mathbf{u}_k, \mathbf{v}_k) N(\mathbf{w}_k; \hat{\mathbf{w}}_{k|k-1}, \mathbf{P}_{k|k-1}) d\mathbf{w}_k - \hat{\mathbf{w}}_{k|k-1} \hat{\mathbf{y}}_{k|k-1}^T \quad (11)$$

一旦接收新的量测值,由式(10)状态变量的后验密度函数计算如下:

$$p(\mathbf{w}_k | D_k) = N(\mathbf{w}_k; \hat{\mathbf{w}}_{k|k}, \mathbf{P}_{k|k}) \quad (12)$$

其中,

$$\hat{\mathbf{w}}_{k|k} = \hat{\mathbf{w}}_{k|k-1} + \mathbf{G}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \quad (13)$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{G}_k \mathbf{P}_{yy, k|k-1} \mathbf{G}_k^T \quad (14)$$

$$\mathbf{G}_k = \mathbf{P}_{wy, k|k-1} \mathbf{P}_{yy, k|k-1}^{-1} \quad (15)$$

另外,网络的提前一步预测输出可以写为

$$\hat{\mathbf{y}}_{k|k-1} = E[\mathbf{h}(\mathbf{w}_k, \mathbf{u}_k, \mathbf{v}_k) | D_{k-1}] \quad (16)$$

由式(5)的线性状态方程,显然,可得到

$$\hat{\mathbf{w}}_{k|k-1} = \hat{\mathbf{w}}_{k-1|k-1} \quad (17)$$

因此,式(13)可等价表示为

$$\hat{\mathbf{w}}_{k|k} = \hat{\mathbf{w}}_{k-1|k-1} + \mathbf{G}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \quad (18)$$

由式(5)可得到预测误差协方差矩阵

$$\mathbf{P}_{k|k-1} = \mathbf{P}_{k-1|k-1} + \mathbf{Q}_{k-1} \quad (19)$$

由于随着训练过程的不断更新,过程噪声协方差  $\mathbf{Q}_{k-1}$  应逐步减小. 因此,可通过含有遗忘因子  $\lambda$  的自适应策略进行退火处理,即令  $\mathbf{Q}_{k-1} = (1/\lambda - 1)\mathbf{P}_{k-1|k-1}$ , 其中遗忘因子  $\lambda \in (0, 1]$ , 这使得过程噪声按指数规律衰减. 这样,式(19)可进一步变形为

$$\mathbf{P}_{k|k-1} = 1/\lambda \mathbf{P}_{k-1|k-1} \quad (20)$$

当所有的条件概率密度函数均假定服从高斯分布时,贝叶斯滤波的一种近似就是对非线性函数与高斯函数乘积的多维积分进行直接数值逼近的 SCKF 算法. 由给出容积点的计算规则,可以实现基于 SCKF 的 SRN 训练算法.

### 3.1 Spherical-Radial 准则

考虑到任意非线性函数与任意高斯分布  $N(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$  的乘积的数学期望总是能变换为对标准高斯分布乘积的数学期望,其中  $\boldsymbol{\mu}$  为均值向量,  $\boldsymbol{\Sigma}$  为协方差矩阵.

考虑多维标准高斯积分以如下形式逼近:

$$\int \mathbf{g}(\boldsymbol{\xi}) N(\boldsymbol{\xi} | \mathbf{0}, \mathbf{I}) d\boldsymbol{\xi} \approx M \sum_i \mathbf{g}(c[\mathbf{e}]_i) \quad (21)$$

其中,  $\boldsymbol{\xi} \in \mathbb{R}^n$ ,  $n$  为状态空间的维数,点集  $[\mathbf{e}]_i$  属于单位向

量  $[\mathbf{I}]$  构成的对称集合,即:

$$[\mathbf{I}] = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots \right\}$$

$M$  是权重,  $c$  是待确定的参数.

因为点集合是对称的,若分别给定  $\mathbf{g}_1(\boldsymbol{\xi}) = 1$ ,  $\mathbf{g}_j(\boldsymbol{\xi}) = \xi_j^2$ , 可以构造出精确到三阶的准则<sup>[10]</sup>确定  $M = \frac{1}{2n}$ ,  $c = \sqrt{n}$ .

该变量代换的方法可以延伸至求取非线性函数与服从任意均值向量与协方差矩阵的高斯分布函数乘积的积分. 给定如下形式的多维积分:

$$\int \mathbf{g}(\mathbf{x}) N(\mathbf{x}; \mathbf{m}, \mathbf{P}) d\mathbf{x} \quad (22)$$

其三阶 Spherical-Radial 准则的容积逼近计算如下:

Step 1 计算单位容积点如下:

$$\boldsymbol{\xi}_i = \begin{cases} \sqrt{n} \mathbf{e}_i, & i = 1, 2, \dots, n \\ -\sqrt{n} \mathbf{e}_{i-n}, & i = n+1, n+2, \dots, 2n \end{cases} \quad (23)$$

其中  $\mathbf{e}_i \in \mathbb{R}^n$  是第  $i$  个坐标轴方向上的单位向量.

Step 2 计算积分如下:

$$\int \mathbf{g}(\mathbf{x}) N(\mathbf{x}; \mathbf{m}, \mathbf{P}) d\mathbf{x} = \frac{1}{2n} \sum_{i=1}^{2n} \mathbf{g}(\mathbf{m} + \sqrt{\mathbf{P}} \boldsymbol{\xi}_i) \quad (24)$$

其中  $\sqrt{\mathbf{P}}$  是矩阵的平方根因子,且  $\mathbf{P} = \sqrt{\mathbf{P}} \sqrt{\mathbf{P}}^T$ .

因此,式(16)可进一步化简为

$$\begin{aligned} \hat{\mathbf{y}}_{k|k-1} &= E[\mathbf{h}(\mathbf{w}_k) | D_{k-1}] = \int_{\mathbb{R}^{N_w}} \mathbf{h}(\mathbf{w}) N(\mathbf{w}; \mathbf{0}, \mathbf{I}) d\mathbf{w} \\ &= \frac{1}{2N_w} \sum_{i=1}^{2N_w} \mathbf{h}(\boldsymbol{\xi}_i) \end{aligned} \quad (25)$$

其中  $\boldsymbol{\xi}_i$  是状态向量  $\mathbf{w}_k \in \mathbb{R}^{N_w}$  的容积点表示.

容积规则是无微分的运算,无需 EKF 算法中的 Jacobians 矩阵或 Hessians 矩阵的复杂求导运算,在每次更新时,仅需计算  $2n$  个容积点,一定程度上避免了“维数灾难”,另外,三阶容积规则还具有  $2n$  个容积点的理论下界,基于该规则的 CKF 或 SCKF 可以认为是非线性贝叶斯滤波器较好的次优逼近.

SCKF 算法通过 Spherical-Radial 准则求得系统状态的容积点,然后通过非线性系统将容积点传播,进而可求得较为精确的后验均值和协方差的平方根因子,通过滤波过程中的时间更新和量测更新实现对所需状态的估计.

### 3.2 SRN-SCKF 算法

标准 CKF 算法在传播容积点的过程中,由于在有限精度的计算机上所执行的算术运算所引起的误差效应,误差协方差矩阵的两个基本特性:正定性和对称性常常会丢失,其正定性的丢失可能会使得 CKF 算法的运行被终止. 在每一时刻的 CKF 时间更新和量测更新过程中所涉及到的对矩阵的平方根运算、求逆运算、舍入误差被矩阵平方运算所放大等数值敏感性运算均会破坏协方差矩阵的特性,而且一些非线性滤波问题也可能是数值病态的,这使得协方差矩阵有可能非正定,从而导致算法不稳定甚至是

不收敛. 为了解决上述问题, 文[11]给出了 CKF 的平方根改进算法, 即 SCKF 算法, 它本质上传播预测和后验误差协方差的平方根因子, 避免了矩阵的平方根运算, 改进了算法的数值稳定性. 除此之外, SCKF 还具有保持协方差的对称性和正定性的特点.

假设在  $k-1$  时刻, 权值状态向量的后验概率密度近似服从高斯分布, 其均值为权值向量滤波估计值  $\hat{\mathbf{w}}_{k-1|k-1}$ , 协方差矩阵等于滤波误差协方差矩阵  $\mathbf{P}_{k-1|k-1}$ , 即  $p(\mathbf{w}_{k-1}|D_{k-1}) = N(\hat{\mathbf{w}}_{k-1|k-1}, \mathbf{S}_{k-1|k-1} \mathbf{S}_{k-1|k-1}^T)$ , 其中,  $\mathbf{S}_{k-1|k-1}$  为  $k-1$  时刻的滤波协方差矩阵的平方根因子, 且  $\mathbf{P}_{k-1|k-1} = \mathbf{S}_{k-1|k-1} \mathbf{S}_{k-1|k-1}^T$ .

基于 SCKF 的 SRN 网络的训练学习算法实现步骤如下:

Step 1 容积点的计算. 考虑式(17)、式(20), 由式(24), 计算权值状态向量的容积点  $\mathbf{W}_{i, k-1|k-1}$

$$\mathbf{W}_{i, k|k-1} = \hat{\mathbf{w}}_{k-1|k-1} + \frac{1}{\sqrt{\lambda}} \mathbf{S}_{k-1|k-1} \boldsymbol{\xi}_i, \quad i=1, \dots, m \quad (26)$$

其中  $\boldsymbol{\xi}_i$  的定义同式(23),  $m=2N_w$ .

Step 2 结合 SRN 的状态空间模型式(6), 计算容积点的传播:

$$\mathbf{Y}_{i, k|k-1} = \hat{\mathbf{h}}(\mathbf{W}_{i, k|k-1}, \mathbf{u}_k, \mathbf{v}_k), \quad i=1, \dots, m \quad (27)$$

Step 3 由式(25), 可估计网络的提前一步预测输出, 即预测均值向量  $\hat{\mathbf{y}}_{k|k-1}$

$$\hat{\mathbf{y}}_{k|k-1} = \frac{1}{m} \sum_{i=1}^m \mathbf{Y}_{i, k|k-1} \quad (28)$$

Step 4 估计量测输出的误差协方差矩阵  $\mathbf{P}_{yy, k|k-1}$  的平方根因子  $\mathbf{S}_{yy, k|k-1}$ , 即新息协方差矩阵的平方根因子.

$$\mathbf{S}_{yy, k|k-1} = \text{Tri}a([\mathbf{Y}_{k|k-1} \mathbf{S}_{R, k}]) \quad (29)$$

其中, 加权的中心化矩阵  $\mathbf{Y}_{k|k-1}$  为

$$\mathbf{Y}_{k|k-1} = \frac{1}{\sqrt{m}} [\mathbf{Y}_{1, k|k-1} - \hat{\mathbf{y}}_{k|k-1}, \mathbf{Y}_{2, k|k-1} - \hat{\mathbf{y}}_{k|k-1}, \dots, \mathbf{Y}_{m, k|k-1} - \hat{\mathbf{y}}_{k|k-1}]$$

矩阵  $\mathbf{S}_{R, k}$  是量测噪声协方差矩阵  $\mathbf{R}_k$  的平方根因子. 算子  $\text{Tri}a(\cdot)$  表示对矩阵实施正交三角分解的运算, 返回一个下三角矩阵.

若  $\mathbf{P}$  为预测误差协方差矩阵, 由

$$\mathbf{P} = \mathbf{A} \mathbf{A}^T = \mathbf{R}^T \mathbf{Q}^T \mathbf{Q} \mathbf{R} = \mathbf{R}^T \mathbf{R} = \mathbf{S} \mathbf{S}^T \quad (30)$$

的结果可知, 对  $\mathbf{P}$  的平方根因子  $\mathbf{A}^T$  做 QR 分解, 则返回上三角矩阵  $\mathbf{R}$ , 由式(30)可知  $\mathbf{S} = \mathbf{R}^T$  是下三角矩阵.

Step 5 估计权值状态向量与量测输出向量的互协方差矩阵

$$\mathbf{P}_{wy, k|k-1} = \mathbf{W}_{k|k-1} \mathbf{Y}_{k|k-1}^T \quad (31)$$

其中, 加权中心化矩阵

$$\mathbf{W}_{k|k-1} = \frac{1}{\sqrt{m}} [\mathbf{W}_{1, k|k-1} - \hat{\mathbf{w}}_{k-1|k-1}, \mathbf{W}_{2, k|k-1} - \hat{\mathbf{w}}_{k-1|k-1}, \dots, \mathbf{W}_{m, k|k-1} - \hat{\mathbf{w}}_{k-1|k-1}]$$

Step 6 计算卡尔曼增益

$$\mathbf{G}_k = \mathbf{P}_{wy, k|k-1} \mathbf{P}_{yy, k|k-1}^{-1} = (\mathbf{P}_{wy, k|k-1} / \mathbf{S}_{yy, k|k-1}^T) / \mathbf{S}_{yy, k|k-1} \quad (32)$$

其中, 符号“/”表示矩阵的右除算子. 若矩阵  $\mathbf{B}$  为上三角矩阵,  $\mathbf{A}/\mathbf{B}$  执行回代算法; 如果  $\mathbf{B}$  为下三角矩阵, 则执行前向替代算法.

Step 7 在获取  $k$  时刻的网络输出值  $\mathbf{y}_k$  时, 根据式(18)的状态更新表达式, 计算权值状态向量  $\hat{\mathbf{w}}_{k|k}$  的滤波值:

$$\hat{\mathbf{w}}_{k|k} = \hat{\mathbf{w}}_{k-1|k-1} + \mathbf{G}_k (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \quad (33)$$

Step 8 估计相应权值状态向量的滤波误差协方差矩阵的平方根因子, 即:

$$\mathbf{S}_{k|k} = \text{Tri}a([\mathbf{W}_{k|k-1} - \mathbf{G}_k \mathbf{Y}_{k|k-1} \mathbf{G}_k \mathbf{S}_{R, k}]) \quad (34)$$

从浮点运算次数分析, SCKF 算法的计算复杂度随着状态维数的立方而增长, 因此其计算复杂度与 CKF 算法、EKF 算法相当, 为  $O(N_w^3)$ .

另外, 与 CKF 算法不同, SCKF 算法在更新过程中, 需要对先验及后验误差协方差矩阵进行正交三角化分解, 传播三角化的协方差矩阵的平方根因子, 以避免矩阵求逆运算. 式(34)是其算法核心, 为得到该式的结果, 可首先由式(32)得到:

$$\mathbf{G}_k \mathbf{P}_{yy, k|k-1} \mathbf{G}_k^T = \mathbf{G}_k \mathbf{P}_{wy, k|k-1} \quad (35)$$

由于  $\mathbf{P}_{yy, k|k-1}$  具有对称性,  $\mathbf{P}_{yy, k|k-1} = \mathbf{P}_{yy, k|k-1}^T$ , 故可得

$$\mathbf{G}_k \mathbf{P}_{yy, k|k-1} \mathbf{G}_k^T - \mathbf{G}_k \mathbf{P}_{wy, k|k-1} = \mathbf{0} \quad (36)$$

其次, 由文[9-10]的 CKF 算法推导过程, 可得误差协方差矩阵的滤波更新过程为

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{G}_k \mathbf{P}_{yy, k|k-1} \mathbf{G}_k^T \quad (37)$$

考虑式(32), 式(35)可变形为

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{P}_{wy, k|k-1} \mathbf{G}_k^T \quad (38)$$

最后, 将式(36)与式(38)相加, 可得:

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{P}_{wy, k|k-1} \mathbf{G}_k^T + \mathbf{G}_k \mathbf{P}_{yy, k|k-1} \mathbf{G}_k^T - \mathbf{G}_k \mathbf{P}_{wy, k|k-1} \quad (39)$$

由于  $\mathbf{P}_{k|k-1} = \mathbf{S}_{k-1|k-1} \mathbf{S}_{k-1|k-1}^T$ , 故可得

$$\begin{aligned} \mathbf{P}_{k|k} &= \mathbf{S}_{k|k-1} \mathbf{S}_{k|k-1}^T - \mathbf{S}_{k|k-1} \mathbf{S}_{k|k-1}^T \mathbf{H}_k^T \mathbf{G}_k^T + \\ &\mathbf{G}_k (\mathbf{S}_{k|k-1} \mathbf{S}_{k|k-1}^T + \mathbf{S}_{R, k} \mathbf{S}_{R, k}^T) \mathbf{G}_k^T - \mathbf{G}_k \mathbf{S}_{k|k-1} \mathbf{S}_{k|k-1}^T \\ &= [\mathbf{S}_{k|k-1} - \mathbf{G}_k \mathbf{S}_{k|k-1} \mathbf{G}_k \mathbf{S}_{R, k}] \times \\ &[\mathbf{S}_{k|k-1} - \mathbf{G}_k \mathbf{S}_{k|k-1} \mathbf{G}_k \mathbf{S}_{R, k}]^T \end{aligned} \quad (40)$$

由式(40)的结果显然可知, 式(34)成立.

算法运行时, 令  $\hat{\mathbf{w}}_0 = E[\mathbf{w}_0]$ ,  $\mathbf{w}_0$  可在区间  $[-0.5, 0.5]$  之间随机取值,  $\mathbf{P}_0 = E[(\mathbf{w}_0 - \hat{\mathbf{w}}_0)(\mathbf{w}_0 - \hat{\mathbf{w}}_0)^T]$ ,  $\mathbf{P}_0 = \mathbf{S}_0 \mathbf{S}_0^T$ ,  $\mathbf{S}_0 = \delta \mathbf{I}_{N_w}$ ,  $\delta$  可取较小的正数,  $\mathbf{I}_{N_w}$  为单位矩阵, 量测噪声协方差矩阵可取为固定值, 即  $\mathbf{R}_k = \sigma^2 \mathbf{I}_{N_w}$ ,  $\sigma$  取较小的正数即可.

## 4 化工过程软测量建模实验

将基于 SCKF 的 SRN 方法应用于化工过程的软测量建模实例中, 利用相关辅助变量及主导变量的历史数据, 结合具有外部输入的非线性时间序列分析模型进行软测量建模, 同等条件下, 还将与基于 BPTT、RTRL、EKF 的 SRN, 基于 EKF、SCKF 的 MLP, 基于 BPTT、RTRL、EKF、SCKF 的 FCRNN 等其它方法进行比较. 性能指标评价选取均方误差 (MSE) 和相关系数<sup>[17]</sup>.

### 4.1 脱丁烷塔底 C4 浓度的软测量估计

脱丁烷塔是炼油厂炼油过程中脱硫和石脑油分离装置的必要组成部分, 需要满足在塔底部 C4 组分含量最小化的质量控制要求, 使用在线监测分析仪测量 C4 浓度时, 由于时延的原因, 整个周期大约需 30~75 min. 因此, 建立软测量模型, 对 C4 浓度进行实时动态监测是十分必要的.

为实现软测量动态建模,在图2所示的分馏装置中需安装传感器进行辅助测量.图2中的7个灰圈表示可测辅助变量: $u_1$ 是塔顶温度, $u_2$ 是塔顶压力, $u_3$ 是回流量, $u_4$

是流向下一过程的流量, $u_5$ 是第6塔板温度, $u_6$ 、 $u_7$ 分别是塔底不同区域的温度,采样周期为12 min.主导变量为C4的浓度.

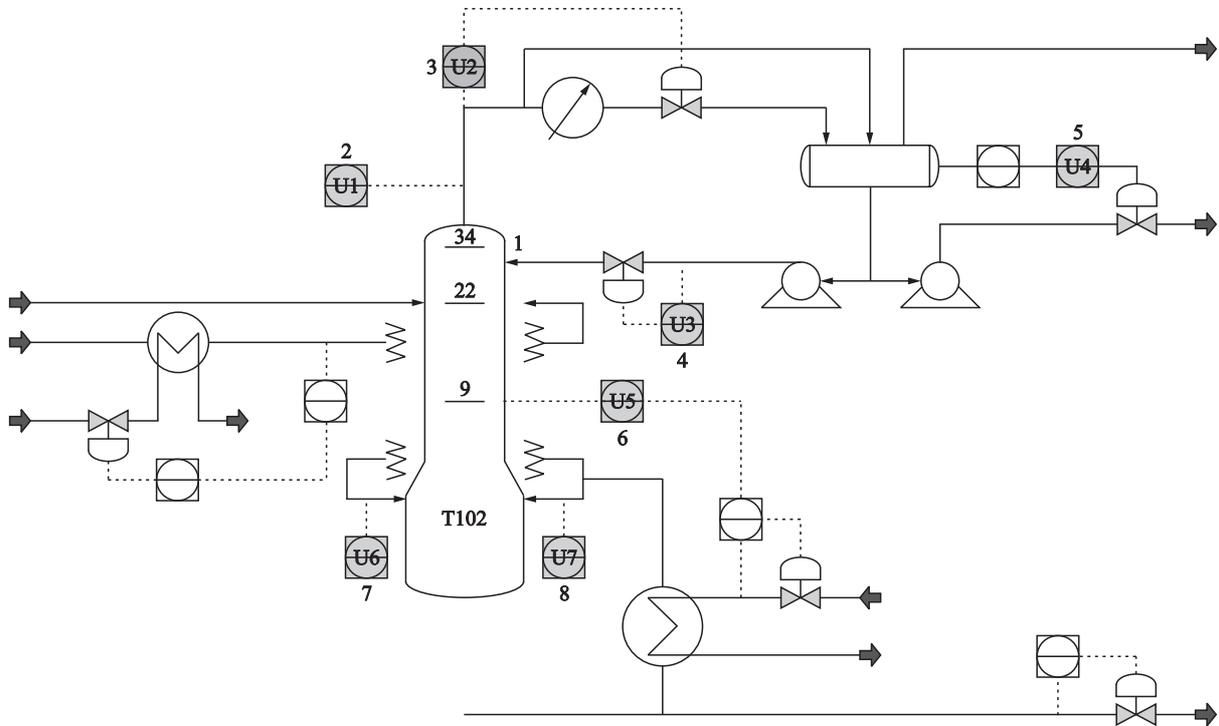


图2 脱丁烷塔的操作流程

Fig.2 Flowchart of debutanizer column

该过程具体的描述见文[2, 5],可收集2394组数据,均归一化为[0, 1]之间的数值.数据集由如下网址: <http://www.springer.com/us/book/9781846284793> 获得.为应用本文方法进行软测量动态建模,结合非线性时间序列分析,考虑如下的NARX模型<sup>[2]</sup>,即

$$y(k) = f(u_1(k), u_2(k), u_3(k), u_4(k), u_5(k), u_5(k-1), u_5(k-2), u_5(k-3), (u_6(k) + u_7(k))/2, y(k-1), y(k-2), y(k-3), y(k-4)) \quad (41)$$

对 $f$ 的逼近采用结构为13-5-1的SRN网络,即采用基于SCKF算法的SRN构建软测量模型,权值初始化为(-0.5, 0.5)之间的随机数,遗忘因子 $\lambda = 0.9995$ ,权值误差协方差矩阵的平方根因子 $S_0 = \sqrt{I \times 10^{-1}}$ , $R_k$ 固定为 $3 \times 10^{-3}$ .

在对比方法中, $f$ 的逼近可采用单隐层的MLP及FCRNN,其具体网络结构与SRN相同.

为了检验模型的性能,将所收集的数据的前一半作为训练数据集,其余的一半为测试数据集.图3给出基于各种不同的方法进行训练时,MSE随训练代数变化的收敛曲线.从图3中可以看出,SRN-SCKF方法的收敛速度最快,经过约10次迭代后可达到较高的精度.图4给出了测试数据集上,基于SRN-SCKF方法的软测量输出估计值与实际值的对比结果.图5相应给出了基于SRN-SCKF方法预测C4浓度的误差结果.从图4与图5的预测效果可以看出,本文方法具有很好的估计精度.

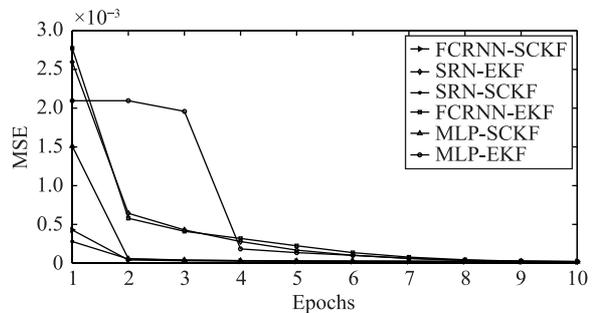


图3 随训练代数变化的不同方法的MSE曲线比较

Fig.3 Comparison of mean square error curve changing with training epochs based on different methods

在测试数据集上,应用不同软测量方法对C4浓度进行估计的具体性能指标则由表1给出.由表1可看出,RNN的建模精度均高于MLP网络,其中,SRN的精度略高于FCRNN,而且,与传统的训练算法相比,采用EKF及SCKF算法进行训练的SRN网络,其收敛速度更快、精度更高,而且SRN-SCKF方法的预测精度比SRN-EKF方法的精度高一个数量级左右,这也进一步表明了所提出方法的有效性.

同时,本文方法还与文[2, 8, 17, 18]的结果进行了比较,其中,文[2]采用了一种具有12个隐含层节点的MLP方法,文[8]采用了一种具有自适应过程状态分区的

局部偏最小二乘法, 文[17]采用了一种具有离线及在线 RLS 学习算法的改进 ESN 网络, 文[18]采用了一种贝叶斯网络建立了估计 C4 浓度的软测量模型. 本文方法的估计精度明显优于文[2, 18]的结果, 略好于文[8, 17]的结果.

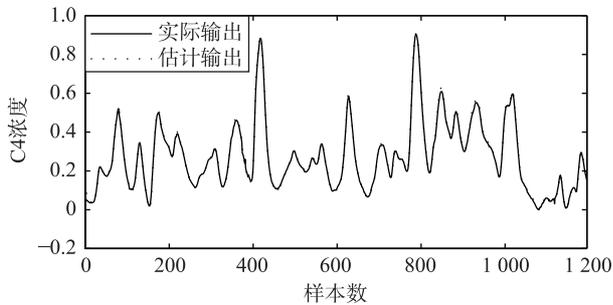


图4 测试集上基于 SRN-SCKF 方法预测 C4 浓度的曲线

Fig.4 Prediction curve of concentration for C4 using SRN-SCKF method in the test dataset

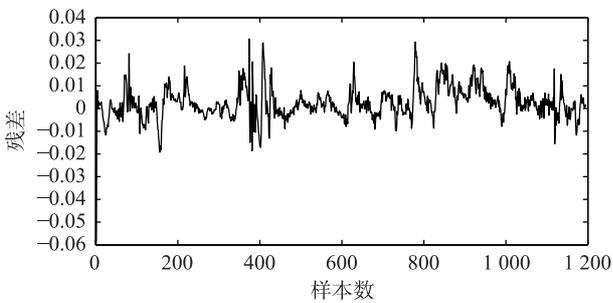


图5 测试集上基于 SRN-SCKF 方法预测 C4 浓度的误差

Fig.5 Prediction error curve for the estimation of the C4 concentration based on SRN-SCKF in the test dataset

#### 4.2 SRU 中 H<sub>2</sub>S 和 SO<sub>2</sub> 浓度的软测量估计

SRU 的作用是在排放酸性气体流之前, 移除其中的

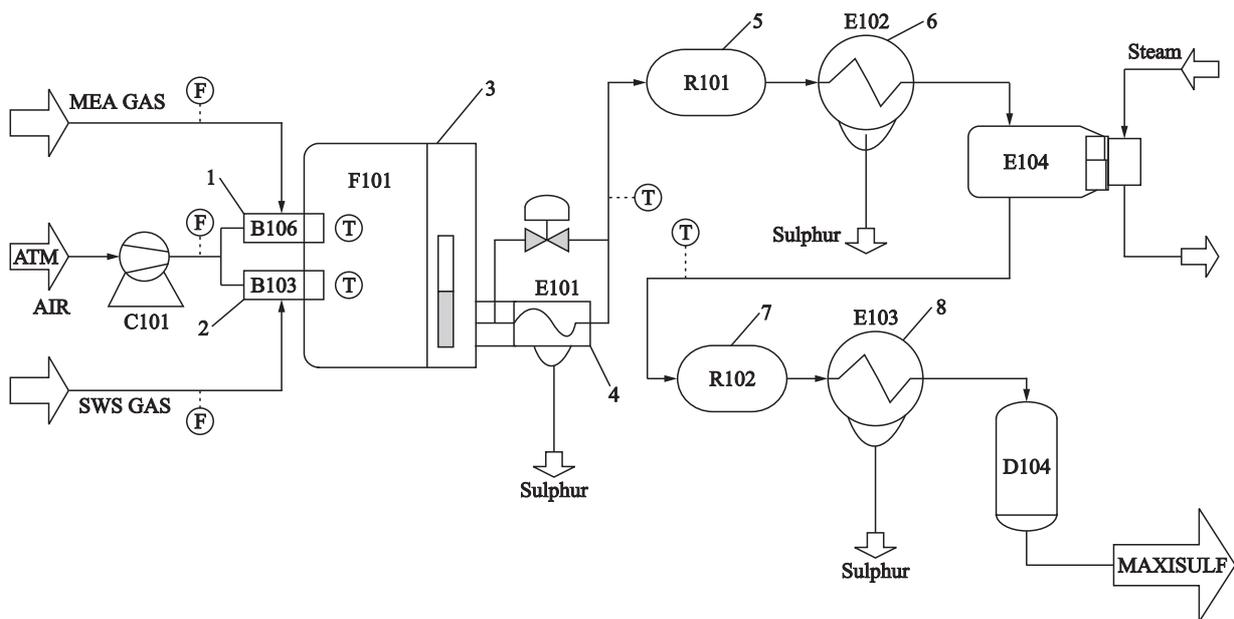


图6 硫回收装置简化流程

Fig.6 Simplified flowchart scheme of SRU

表1 在测试集上不同方法对 C4 浓度估计的性能指标  
Tab.1 The performance indicators of the estimation concentration of C4 using different methods in test dataset

不同方法	相关系数	MSE	Epochs
FCRNN-BPTT	0.985 7	$9.19 \times 10^{-4}$	300
SRN-BPTT	0.986 1	$9.86 \times 10^{-4}$	300
FCRNN-RTRL	0.990 3	$6.72 \times 10^{-4}$	300
SRN-RTRL	0.990 3	$5.06 \times 10^{-4}$	300
MLP-EKF	0.995 7	$3.52 \times 10^{-4}$	10
FCRNN-EKF	0.998 8	$2.08 \times 10^{-4}$	10
SRN-EKF	0.998 2	$1.41 \times 10^{-4}$	10
MLP-SCKF	0.999 1	$7.82 \times 10^{-5}$	10
FCRNN-SCKF	0.999 1	$6.10 \times 10^{-5}$	10
SRN-SCKF	0.999 7	$2.70 \times 10^{-5}$	10

SO<sub>2</sub>、H<sub>2</sub>S 等环境污染物, 同时将硫元素作为一种副产品进行回收, 4 条完全相同的硫回收线子单元将并行工作, 主要对两种气体进行处理, 第一种产生于洗气装置中, 主要包含 H<sub>2</sub>S, 称为 MEA 气体. 第二种产生于酸水汽提装置, 富含 NH<sub>3</sub> 和 H<sub>2</sub>S, 也称为 SWS 气体. SRU 的某一条硫回收线的简化流程如图 6 所示.

由于酸性气体较强的腐蚀性, 使得在线分析仪的故障频率很高, 当在线分析仪发生故障时, 使用“软传感器”则可以监测 H<sub>2</sub>S 和 SO<sub>2</sub> 的浓度, 使得装置正常运行. 由图 6 可分别选取 5 个辅助变量, 即:  $u_1$  是 MEA 气体流量 (MEA\_GAS);  $u_2$  是空气流量 (AIR\_MEA);  $u_3$  是第二空气流量 (AIR\_MEA\_2);  $u_4$  是 SWS 区域的气体流量 (SWS\_GAS + MEA\_SPILLING);  $u_5$  是 SWS 区域的空气流量 (AIR\_SWS + MEA\_SPILLING\_AIR), 辅助变量的采样周期为 1 min. 主导变量分别为 H<sub>2</sub>S 和 SO<sub>2</sub> 的浓度输出.

SRU 过程的描述由文[2]给出,数据集由网址 <http://www.springer.com/us/book/9781846284793> 获得. 为应用本文的方法进行动态软测量模型建模,结合非线性时间序列分析,考虑如下的 NMA 模型<sup>[2]</sup>:

$$y_1(k) = f_1(u_1(k), u_1(k-5), u_1(k-7), u_1(k-9), \dots, u_5(k), u_5(k-5), u_5(k-7), u_5(k-9)) \quad (42)$$

$$y_2(k) = f_2(u_1(k), u_1(k-5), u_1(k-7), u_1(k-9), \dots, u_5(k), u_5(k-5), u_5(k-7), u_5(k-9)) \quad (43)$$

其中,  $y_1(k)$  表示  $H_2S$  的浓度输出,  $y_2(k)$  表示  $SO_2$  的浓度输出,未知的  $f_1(\cdot)$  及  $f_2(\cdot)$  可采用 SRN 网络、FCRNN 网络或 MLP 进行逼近,分别构建用于  $H_2S$  和  $SO_2$  浓度估计的软测量模型.

该过程收集 10 081 组数据,均归一化为  $[0, 1]$  之间的数值,前一半数据选取为训练数据集,其余的为测试数据集. SRN 网络的结构为 20-6-1,权值初始化为  $(-0.5, 0.5)$  之间的随机数,遗忘因子  $\lambda = 0.999 5$ ,权值误差协方差矩阵的平方根因子  $S_0 = \sqrt{I \times 10^{-1}} I$ ,  $R_k$  固定为  $5 \times 10^{-3}$ . 对比方法中,MLP 及 FCRNN 的结构为 20-6-1.

图 7、图 8 分别给出了对  $H_2S$  和  $SO_2$  的浓度进行估计时,在训练数据集上,基于不同方法 MSE 随训练代数变化的曲线. 从图 7、图 8 可看出,SRN-SCKF 方法的收敛速度较快,经过约 10 次迭代后可获取满意的精度. 图 9 和图 10 则分别给出了基于 SRN-SCKF 方法对  $H_2S$  的浓度进行估计时,在测试集上模型的软测量估计值与实际输出值的对比效果及相应的误差结果曲线. 图 11 和图 12 分别给出了基于 SRN-SCKF 方法对  $SO_2$  浓度进行估计时,在测试集上模型的软测量估计值与实际输出值的对比效果及相应的误差结果曲线,从图 9 ~ 图 12 的结果看出,SRN-SCKF 方法的估计效果非常好.

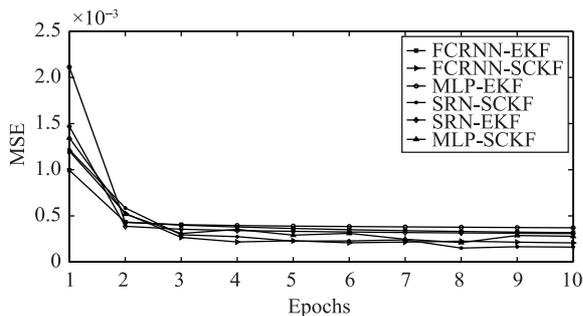


图 7 估计  $H_2S$  浓度时,基于不同方法 MSE 曲线比较  
Fig.7 Comparison of MSE curve for the estimation of the  $H_2S$  concentration based on different methods

表 2 和表 3 给出了基于不同方法估计  $H_2S$  和  $SO_2$  浓度时,在测试集上具体性能指标数值的对比. 可以看出,SRN-SCKF 方法的估计精度与相关系数均最好,这表明了本文方法的有效性. 同时,还与文[2, 17]的结果进行了对比. 在文[2]中,基于 MLP、RBF 神经网络,自适应神经模糊系统和非线性最小二乘四种方法分别对  $H_2S$  和  $SO_2$  的浓度进行估计,在数据集中分别随机选取 1 000 组数据作

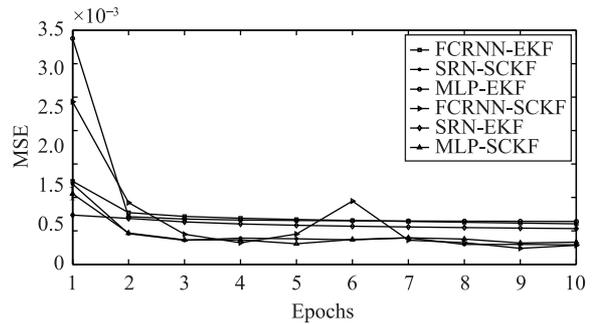


图 8 估计  $SO_2$  浓度时,基于不同方法 MSE 曲线比较  
Fig.8 Comparison of MSE curve for the estimation of the  $SO_2$  concentration based on different methods

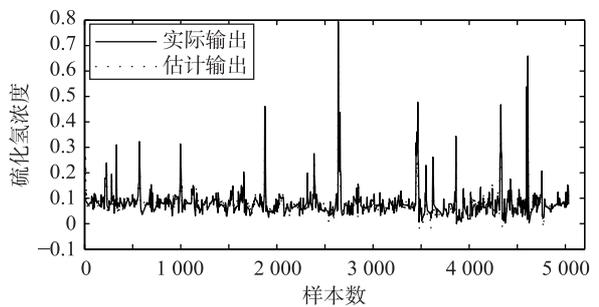


图 9 测试集上基于 SRN-SCKF 方法预测  $H_2S$  浓度曲线  
Fig.9 Prediction curve for the  $H_2S$  concentration based on SRN-SCKF method in the test dataset

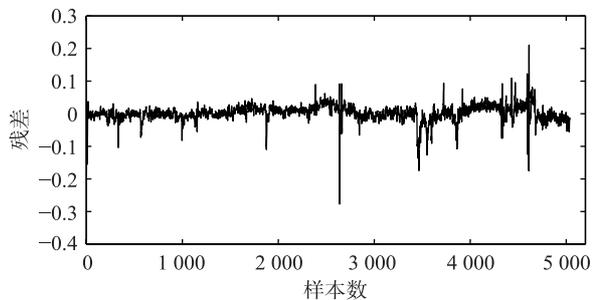


图 10 测试集上基于 SRN-SCKF 方法预测  $H_2S$  浓度的误差  
Fig.10 Prediction error curve for the  $H_2S$  concentration based on SRN-SCKF method in the test dataset

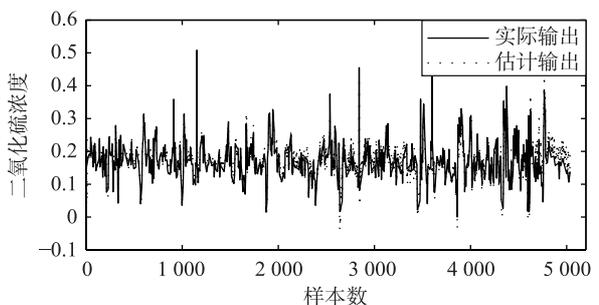


图 11 测试集上基于 SRN-SCKF 方法预测  $SO_2$  浓度曲线  
Fig.11 Prediction curve for the  $SO_2$  concentration based on SRN-SCKF method in the test dataset

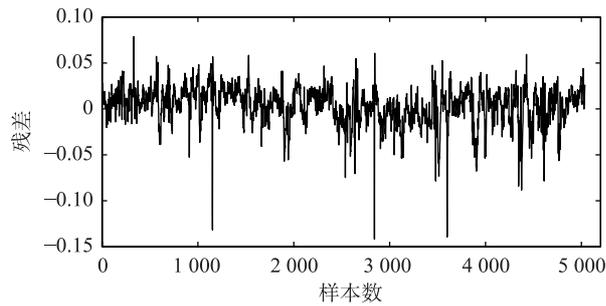


图 12 测试集上基于 SRN-SCKF 方法预测 SO<sub>2</sub> 浓度的误差

Fig.12 Prediction error curve for the SO<sub>2</sub> concentration based on SRN-SCKF in the test dataset

表 2 在测试集上 H<sub>2</sub>S 性能评价指标对比结果

Tab.2 Performance indicators for the estimation of H<sub>2</sub>S concentration using different methods in test dataset

不同方法	相关系数	MSE	Epochs
FCRNN-BPTT	0.720 1	$1.29 \times 10^{-3}$	300
SRN-BPTT	0.732 6	$1.78 \times 10^{-3}$	300
FCRNN-RTRL	0.673 8	$1.52 \times 10^{-3}$	300
SRN-RTRL	0.680 1	$1.92 \times 10^{-3}$	300
MLP-EKF	0.774 5	$1.11 \times 10^{-3}$	10
FCRNN-EKF	0.795 2	$1.07 \times 10^{-3}$	10
SRN-EKF	0.782 8	$1.02 \times 10^{-3}$	10
MLP-SCKF	0.933 0	$3.50 \times 10^{-4}$	10
FCRNN-SCKF	0.946 3	$3.27 \times 10^{-4}$	10
SRN-SCKF	0.950 6	$3.15 \times 10^{-4}$	10

为训练样本和测试样本. 其中, 非线性最小二乘方法的预测精度最高, 对 H<sub>2</sub>S 的 MSE 为  $8 \times 10^{-4}$ , 相关系数为 0.848;

对 SO<sub>2</sub> 的 MSE 为  $4 \times 10^{-4}$ , 相关系数为 0.905. 文[17]采用  $\epsilon$ -SVM、LSSVM 和 LiESN 网络分别对 H<sub>2</sub>S 和 SO<sub>2</sub> 的浓度进行估计. 由比较可知, 本文方法与 LiESN 方法获取的结果相当, 均优于文献中其他方法.

表 3 在测试集上 SO<sub>2</sub> 性能评价指标对比结果

Tab.3 Performance indicators for the estimation of SO<sub>2</sub> concentration using different methods in the test dataset

不同方法	相关系数	MSE	Epochs
FCRNN-BPTT	0.816 1	$1.07 \times 10^{-3}$	300
SRN-BPTT	0.800 1	$2.05 \times 10^{-3}$	300
FCRNN-RTRL	0.690 7	$1.84 \times 10^{-3}$	300
SRN-RTRL	0.663 2	$2.79 \times 10^{-3}$	300
MLP-EKF	0.791 8	$1.18 \times 10^{-3}$	10
FCRNN-EKF	0.845 1	$7.81 \times 10^{-4}$	10
SRN-EKF	0.860 2	$7.09 \times 10^{-4}$	10
MLP-SCKF	0.878 1	$6.84 \times 10^{-4}$	10
FCRNN-SCKF	0.901 7	$5.81 \times 10^{-4}$	10
SRN-SCKF	0.941 2	$4.02 \times 10^{-4}$	10

## 5 结论

针对具有强非线性特性的化工过程, 利用所收集的相相关辅助变量及主导变量的历史数据, 基于 NARX 和 NMA 非线性时间序列模型, 提出了一种基于 SRN-SCKF 的动态软测量建模方法. 所提出方法的优点是: 训练过程中应用 SCKF 算法对网络的权值进行更新, 避免了常规梯度下降算法易陷入局部极小的不足, 也无需 EKF 算法或 CKF 算法所涉及的矩阵求导或求逆运算; 算法的收敛速度快, 精度高. 通过两个化工过程实例的应用, 验证了本文方法的有效性.

## 参考文献

- [1] Kadlec P, Grbić R, Gabrys B. Review of adaptation mechanisms for data-driven soft sensors[J]. Computers & Chemical Engineering, 2011, 35(1): 1–24.
- [2] Fortuna L, Graziani S, Rizzo A, et al. Soft sensors for monitoring and control of industrial processes[M]. Berlin, Germany: Springer-Verlag, 2007.
- [3] 孙茂伟, 杨慧中. 局部加权混合核偏最小二乘算法及其在软测量中的应用[J]. 信息与控制, 2015, 44(4): 481–486.  
Sun M W, Yang H Z. Local weighted mixed kernel partial least squares algorithm and its applications in to soft-sensing[J]. Information and Control, 2015, 44(4): 481–486.
- [4] 熊伟丽, 张伟, 徐保国. 一种基于 EGMM 的高斯过程回归软测量建模[J]. 信息与控制, 2016, 45(1): 14–19.  
Xiong W L, Zhang W, Xu B G. A Soft sensor modeling method based on EGMM using gaussian process regression[J]. Information and Control, 2016, 45(1): 14–19.
- [5] Fortuna L, Graziani S, Xibilia M G. Soft sensors for product quality monitoring in debutanizer distillation columns[J]. Control Engineering Practice, 2005, 13(4): 499–508.
- [6] Lin B, Recke B, Knudsen J K H, et al. A systematic approach for soft sensor development[J]. Computers & Chemical Engineering, 2007, 31(5): 419–425.
- [7] Graziani S, Napoli G, Xibilia M G. Soft sensor design for a sulfur recovery unit using a clustering based approach[C]//Proceedings of Instrumentation and Measurement Technology. Piscataway, NJ, USA: IEEE, 2008: 1162–1167.
- [8] Shao W M, Tian X M, Wang P, et al. Online soft sensor design using local partial least squares models with adaptive process state partition[J]. Chemometrics and Intelligent Laboratory Systems, 2015, 144: 108–121.
- [9] He Y L, Xu Y, Zhu Q X. Soft-sensing model development using PLSR-based dynamic extreme learning machine with an enhanced hidden layer[J]. Chemometrics and Intelligent Laboratory Systems, 2016, 154: 101–111.

(下转第 357 页)

- tion and Control, 2014, 43(5): 637–640.
- [18] 张建桃, 张铁民, 梁莉. 超声电机非线性建模和广义预测控制[J]. 电机与控制学报, 2011, 15(6): 50–56.  
Zhang J T, Zhang T M, Liang L. Nonlinear modeling and generalized predictive control of ultrasonic motor[J]. 2011, 15(6): 50–56.
- [19] Eliasi H, Menhaj M B, Davilu H. Robust nonlinear model predictive control for a PWR nuclear power plant[J]. Process in Nuclear Energy, 2012, 54(9): 177–185.
- [20] 李靖, 张向文. 汽车 EPS 预测控制仿真研究[J]. 计算机仿真, 2015, 32(11): 188–192.  
Li J, Zhang X W. Simulation of predictive control of vehicle electric power steering system[J]. Computer Simulation, 2015, 32(11): 188–192.

---

## 作者简介

张帆(1990–), 男, 硕士生. 研究领域为汽车电子, 电机控制.

张向文(1976–), 男, 博士, 研究员, 硕士生导师. 研究领域为汽车电子, 智能控制与信号处理.

---

(上接第 349 页)

- [10] 王魏, 柴天佑, 赵立杰. 带有稳定学习的递归神经网络动态偏最小二乘建模[J]. 控制理论与应用, 2012, 29(3): 337–341.  
Wang W, Chai T Y, Zhao L J. Dynamic partial least squares modeling with recurrent neural networks of stable learning[J]. Control Theory & Applications, 2012, 29(3): 337–341.
- [11] Haykin S. Neural networks and learning machines [M]. Upper Saddle River, NJ, USA: Pearson, 2009.
- [12] Čerňanský M, Makula M, Beňušková L. Organization of the state space of a simple recurrent network before and after training on recursive linguistic structures[J]. Neural Networks, 2007, 20(2): 236–244.
- [13] Arasaratnam I, Haykin S. Cubature kalman filters[J]. IEEE Transactions on Automation Control, 2009, 54(6): 1254–1269.
- [14] Arasaratnam I. Cubature Kalman filtering: Theory & applications[D]. Hamilton, ON, Canada: McMaster University, 2009.
- [15] 胡振涛, 袁光耀, 胡玉梅, 等. 基于容积卡尔曼滤波的神经网络训练算法[J]. 控制与决策, 2016, 31(2): 355–360.  
Hu Z T, Yuan G Y, Hu Y M, et al. Training method of neural network based on cubature Kalman filter[J]. Control and Decision, 2016, 31(2): 355–360.
- [16] Wang S Y, Feng J C, Tse C K. Novel cubature Kalman filtering for systems involving nonlinear states and linear measurements[J]. AEU-International Journal of Electronics and Communications, 2015, 69(1): 314–320.
- [17] 李军, 岳文琦. 基于泄漏积分型回声状态网络的软测量动态建模方法及应用[J]. 化工学报, 2014, 65(10): 4004–4014.  
Li J, Yue W Q. Dynamic soft sensor modeling and its application using leaky-integrator ESN[J]. CIESC Journal, 2014, 65(10): 4004–4014.
- [18] 李雅芹, 周开武, 杨慧中. 基于贝叶斯网络的软测量建模方法[J]. 计算机与应用化学, 2010, 27(10): 1391–1394.  
Li Y Q, Zhou K W, Yang H Z. A soft sensor based on Bayesian network [J]. Computers and Applied Chemistry, 2010, 27(10): 1391–1394.

---

## 作者简介

李军(1969–), 男, 博士, 教授. 研究领域为计算智能与非线性系统建模、控制等.

桑桦(1991–), 男, 硕士生. 研究领域为计算智能与软测量建模.