AN ADAPTIVE SOFT SENSOR DETERIORATION EVALUATION AND MODEL UPDATING METHOD FOR TIME-VARYING CHEMICAL PROCESSES

Article Highlights
• An adaptive soft sensor model deterioration evaluation method is proposed
• An adaptive modeling sample data updating method is proposed
• The RMSE and MAE of the fitting performance of the proposed method in a CSTR are improved
• The RMSE and MAE of the predicted performances of the proposed method in a CSTR have lower values
• The RMSE and MAE of the predicted performances of the proposed method in a DCP are the lowest average values

Abstract
Due to the time-varying nature of chemical processes, soft sensor models deteriorate, and data prediction accuracy decreases. To address this problem, an adaptive soft sensor modeling method is proposed that not only evaluates the model deterioration by an adaptive moving window-constrained statistical hypothesis test, but also adaptively updates the modeling samples using moving window-cosine similarity. First, this method evaluates the model deterioration via positioning by constrained statistical hypothesis testing based on the differences between the prediction performance evaluation index data obtained from moving window stepping and the original prediction performance evaluation indexes. Additionally, the dynamic temporal variation in chemical processes causes changes in the impacts of the auxiliary variables on the dominant variable, and this effect limits the improvement in the prediction accuracy of the soft sensor model by updating only the auxiliary variable data. The moving window-cosine similarity method is combined to propose a strategy that updates both the modeled auxiliary variables and the auxiliary variable data. Finally, the parameters of the soft sensor model are optimized via particle swarm optimization (PSO) to improve the fitting performance. Simulated data of a continuous stirred tank reactor (CSTR) and actual data from a debutanizer column process (DCP) are used for model verification to evaluate the performance of the proposed adaptive soft sensor modeling method, and the results show its effectiveness.

Keywords: chemical process, adaptive soft sensor model, statistical hypothesis test, moving window, cosine similarity, PSO.

During chemical production, several main process variables, such as product quality, have characteristics such as a low sampling rate and delay [1]. It is important and desirable to accurately obtain the values of these quality-related variables in real time because infrequent and inaccurate measurements may result in poor control performance, large production losses and even safety hazards [2]. Therefore, these parameters must be estimated via more easily obtained process variables to ensure the stability of the major process variable data. Because most chem-

L. LI, Y. DAI: AN ADAPTIVE SOFT SENSOR DETERIORATION METHOD

The decline in the prediction accuracy of soft sensor models and obtain accurate chemical process state information, which will achieve optimal control of chemical processes and improve product quality [27,28].

Adaptive learning mechanisms mainly include moving window learning [29,30], recursive learning [31,32], temporal difference modeling [33,34] and offset compensation [35,36]. However, these methods have several limitations that must be analyzed. For example, moving window and recursive models suffer from difficulties in dealing with abrupt variations, such as changes of set point values, because they usually cannot track the process dynamics until sufficient new samples from the new operating conditions have been collected. In addition, recursive learning is unable to track the process state of sudden changes in time, and offset compensation weights are difficult to calculate. Therefore, Xiong et al. [33] combine the temporal difference with moving window method which can get the modeling data set through moving window method with the delay parameters. Liu et al. [37] propose an adaptive model-moving windows (MW), time difference (TD), and locally weighted regression (LWR) under the framework of Bayesian network (BN), which can get lower values of RMSE than moving window method. Ni et al. [36] combine the moving window with GPR, and incorporate the mean and variance updates in the modeling samples, that can effectively capture the process dynamics and to model nonlinearity simultaneously. Moreover, other methods are also applied to adaptive soft sensors. Xu et al. [38] propose an adaptive method with nonlinear differential-algebraic observer for chemical processes, that through output feedback gain matrix realize the adjustment online. The above method effectively realizes the adaptive adjustment of soft sensor model. However, these methods cannot identify the deterioration of soft sensors in time-varying chemical processes adaptively. In recent years, an identification method based on data subsets consisting of consecutive time samples was proposed [39,40]. This method achieves the segmentation of test datasets using a moving window with a fixed length. Based on the test data obtained by moving the window step-by-step, the deterioration of the soft sensor model is determined by the t-test method [41]. Additionally, a method for determining the deterioration of models based on chi-square distribution and Mahalanobis distance is proposed [42]. The above method realizes the deterioration of soft sensor model effectively. However, the partitioning strategies ignore the negative influence of the variance in the predicted residual [43,44]. As a result, if the process characteristics

ical processes are not well understood and have distinct features, such as nonlinearity and variations over time, creating soft sensors via data-driven methods [3,4] is an effective approach to address these issues [5,6].

As an alternative to the mechanism model, soft sensors have increasingly been developed in recent years to deal with the problems of unacceptably high costs and large measurement delays for data acquisition of major process variables [7,8]. The basic idea of soft measurement is to combine the theory of automatic control with knowledge of the production process organically. Because the dominant variables are often difficult to measure or cannot be measured temporally, other auxiliary variables that are easy to measure are selected, inferred and estimated by forming mathematical relationships, and the hardware function is replaced by software.

In recent years, a variety of data-driven algorithms have been applied to develop soft sensors. The most widely used methods are principal component regression (PCR) [9,10], partial least squares (PLS) [11,12], artificial neural networks (ANNs) [13,14], support vector machines (SVMs) [15-17], Gaussian process regression (GPR) [18,19], and extreme learning machines (ELMs) [20,21]. Of these nonlinear soft sensing methods that have been extensively reported in the literature, the least squares support vector control (LSSVM) is becoming increasingly popular for practical applications. LSSVM have superior computing speed, high convergence precision, and generalizable performance. In addition, compared with other methods, LSSVM are more suitable for chemical processes with small samples. However, the regularized parameter C in the objective function and the radial basis function parameter σ² affect the regression performance of the soft sensor method.

Moreover, due to factors such as variations in operating conditions, declines in catalyst activity, mechanical wear, compositional variations in the feed materials and seasonal variations, most industrial processes are time-varying [22]. Conventional soft sensor models cannot adapt to new conditions and deteriorate; thus, these models cannot accurately predict the current state [23-25]. Adaptive learning mechanisms enable online learning for soft sensor models and achieve model "self-maintenance", which effectively prevents time-varying factor-induced model performance deterioration and extends the life of soft sensor models [26]. Therefore, it is important to investigate methods to adaptively identify model deterioration, promptly update soft sensor models, prevent
change, invalid assumptions may remain valid. Nevertheless, a potential problem associated with this approach is that a model decision based on a single piece of data is vulnerable to abnormal process data. Therefore, even if the process characteristics do not change, it is possible that model deterioration may be identified.

The subsequent part of an adaptive soft sensor is the model updating method using the upd-modeling samples. Several methods like improved empirical mode decomposition (IEMD) [45], statistical approach of stepwise linear regression [46], mutual information [42] and auto-associative neural network (AANN) [47] are used to select the modeling variables, which can improve the prediction performance of the model effectively. However, the moving window method has several advantages over other methods, such as high online computational efficiency and ease of integration with modeling methods, and this mechanism is more suitable for processes with gradual temporal variations, such as chemical processes [48,49]. However, the moving window mechanism updates only the auxiliary variable data and not the soft sensor-modeled auxiliary variable, and it is limited in improving the prediction accuracy. How to combine the basic moving window with the method of selecting auxiliary variables in modeling is an important problem that deserves in-depth research. A sample data updating strategy for handling the time-varying characteristics of chemistry processes is the topic of this paper.

Because the soft sensor model parameters affect the model fitting performance and are difficult to select, nonlinear identification and robustness of the soft sensor model are normally implemented via an intelligent optimization algorithm [50]. Due to its advantages, including ease of implementation and fewer numbers of adjustable parameters [51,52], particle swarm optimization (PSO) [53] has attracted attention for this issue [54,55]. In this paper, the parameters of the soft sensor model are optimized via PSO to improve its accuracy.

Based on the discussion presented above, this paper proposes an adaptive soft sensor method based on a moving window-constrained statistical hypothesis test and moving window-cosine similarity that is focused on solving the aforementioned issues remaining in the two main steps of the adaptive soft sensor modeling method to develop a high-accuracy soft sensor for time-varying chemical processes. First, a model deterioration detection method based on a moving window-constrained statistical hypothesis test with dual detected data is presented to improve the detection of sensor model deterioration for chemical process states in previous work to reduce the inaccuracy. Subsequently, an updating mechanism that explicitly utilizes information and quantifies the correlation between the auxiliary variables and the dominant variable for upd-samples is proposed to build a set of modeling samples via the moving window-cosine similarity method. Moreover, the performance of the proposed soft sensor is evaluated in detail.

The remainder of this paper is structured as follows: the section on adaptive soft sensor development details the improved local learning-based soft sensor development, including the soft sensor model deterioration detection, sample data updating, and soft sensor modeling method; in the section on simulation and discussion, case studies using simulated data for a continuous stirred tank reactor (CSTR) and actual data from a debutanizer column process (DCP) are conducted, and performance evaluations of the proposed method are reported. Finally, the results of the paper are given in the conclusion.

The SPSS and MATLAB softwares are used to perform this work.

Adaptive soft sensor development

The modeling samples of the adaptive soft sensor are composed of “offline” data and “online” data. In the offline part, a local dataset is constructed using a moving window, and a soft sensor model is based on the local dataset. In the online part, to estimate the primary variable, the original soft sensor model gives a response with the samples updated by the moving window. However, when the soft sensor model deteriorates, it is updated by the upd-samples. To handle the soft sensor modeling of time-varying chemical processes, this paper proposes an adaptive soft sensor modeling method. The proposed method applies an LSSVM as a basic modeling tool, integrates the statistical hypothesis test to determine the deterioration of the soft sensor model, employs the moving window-cosine similarity to update the modeling samples of the soft sensor, and finally uses the LSSVM to build the soft sensor model. A schematic diagram of the proposed method is given in Figure 1.

Soft sensor model deterioration detection method based on a moving window-constrained statistical hypothesis test

Analyses of chemical processing equipment and production states show that due to the dynamic temporal variations in chemical processes, the performance of soft sensor models gradually degrades, resulting in their deterioration. To test the deterioration of a soft sensor model, we assume that the number of
samples in a soft sensor model of a sequential pipeline chemical process is \( n (n \geq 1) \), the number of modeled sample auxiliary variables is \( p (p \geq 1) \), the model sample dataset is \( D_{\text{ori}} = \{X_{\text{ori}}, Y_{\text{ori}}\} \), the test dataset is \( D_{\text{test}} = \{X_{\text{test}}, Y_{\text{test}}\} \), the length of the test dataset is the moving window width \( w (w \geq 1) \), the moving step is 1, and the sample data in the current area when the moving window proceeds \( l \) steps are \( D_{\text{ori}} = \{X_{\text{ori}}, Y_{\text{ori}}\} \).

Due to the excellent universality of the root mean squared error (RMSE) and mean absolute error (MAE) in evaluating the prediction performance of soft sensor models, they can effectively solve the problem of state change and model deterioration. In this paper, the RMSE and MAE evaluation indexes are used as basic data to evaluate the deterioration of the soft sensor model.

The prediction error of the modeling data is:

\[
R_{\text{ori}} = \frac{1}{w} \sum_{i=1}^{w} [f_{\text{ori}}(X_{\text{ori}}(i)) - Y_{\text{ori}}(i)]^2
\] (1)

\[
M_{\text{ori}} = \frac{1}{w} \sum_{i=1}^{w} [f_{\text{ori}}(X_{\text{ori}}(i)) - Y_{\text{ori}}(i)]
\] (2)

The prediction error data in the current area are:

\[
R_{\text{ori}} = \frac{1}{w} \sum_{i=1}^{w} [f_{\text{ori}}(X_{\text{ori}}(i)) - Y_{\text{ori}}(i)]^2
\] (3)

\[
M_{\text{ori}} = \frac{1}{w} \sum_{i=1}^{w} [f_{\text{ori}}(X_{\text{ori}}(i)) - Y_{\text{ori}}(i)]
\] (4)

If there are no significant differences between \( R_{\text{ori}} \) and \( R_{\text{ori}} \), or between \( M_{\text{ori}} \) and \( M_{\text{ori}} \), the original soft sensor model has experienced no deterioration and can be used. If there are significant differences between \( R_{\text{ori}} \) and \( R_{\text{ori}} \), or between \( M_{\text{ori}} \) and \( M_{\text{ori}} \), the original soft sensor model has deteriorated, and the soft sensor model should be updated. The problem is thus converted into determining how to identify significant differences between \( R_{\text{ori}} \) and \( R_{\text{ori}} \), or between \( M_{\text{ori}} \) and \( M_{\text{ori}} \).

The error of the soft sensor model is defined as follows:

\[
D_R = R_{\text{ori}} - R_{\text{ori}}
\] (5)

\[
D_M = M_{\text{ori}} - M_{\text{ori}}
\] (6)

Based on the detection method of distributions, the moving window proceeds \( l \) steps, and the data distributions \( \{D_R(i)\}_{i=1}^{l} \) and \( \{D_M(i)\}_{i=1}^{l} \) can be ascertained.

The test problems are as follows:

\[
H_0 : \mu_{\text{ori}} = \mu_{\text{ori}} \land \{S_{\text{ori}}^2 = S_{\text{ori}}^2\}
\]

\[
H_1 : \mu_{\text{ori}} \neq \mu_{\text{ori}} \lor \{S_{\text{ori}}^2 \neq S_{\text{ori}}^2\}
\] (7)

The statistical variable is as follows:

\[
T = \frac{\hat{\mu}_{\text{ori}} - \mu_{\text{ori}}}{\sqrt{S_{\text{ori}}^2 / L}}
\] (8)

Normally, \( \mu_{\text{ori}} = 0 \), and \( \hat{\mu}_{\text{ori}} \) is the average of \( \{D_R(i)\}_{i=1}^{l} \) and \( \{D_M(i)\}_{i=1}^{l} \), as follows:

\[
\hat{\mu}_{\text{ori}} (R) = \frac{1}{L} \sum_{i=1}^{L} D_R(i)
\] (9)

\[
\hat{\mu}_{\text{ori}} (M) = \frac{1}{L} \sum_{i=1}^{L} D_M(i)
\] (10)

\( S_{\text{ori}}^2 \) is the variance of \( \{D_R(i)\}_{i=1}^{l} \) and \( \{D_M(i)\}_{i=1}^{l} \), as follows:

\[
S_{\text{ori}}^2 (R) = \frac{1}{L-1} \sum_{i=1}^{L} (D_R(i) - \hat{\mu}_{\text{ori}} (R))^2
\] (11)

\[
S_{\text{ori}}^2 (M) = \frac{1}{L-1} \sum_{i=1}^{L} (D_M(i) - \hat{\mu}_{\text{ori}} (M))^2
\] (12)

Thus, the rejection region of this test problem is:

\[
T \geq \frac{t_{n-1} (n-1)}{\sqrt{S_{\text{ori}}^2 / L}}
\] (13)

When \( \hat{\mu}_{\text{ori}} (R) = \hat{\mu}_{\text{ori}} (M) = 0 \), based on the characteristics of the \( t \) distribution, the statistical
variables \( T(D_R) \) and \( T(D_M) \) follow the \( t \) distribution with 
\( L-1 \) degrees of freedom [26] as follows:
\[
T_{D_R} \sim T(L-1) \quad (14)
\]
\[
T_{D_M} \sim T(L-1) \quad (15)
\]

Therefore, the model deterioration is identified via a \( t \)-test of the \( RMSE \) and \( MAE \) by finding
the proper thresholds \( \lambda_i(R) \) and \( \lambda_i(M) \) for \( D_R \) and \( D_M \), respectively, in the given confidence ranges \( \alpha_i(R) \) and \( \alpha_i(M) \), respectively. Hypothesis \( H_1 \) is accepted if
and only if the following condition is true:
\[
(T_{D_R} > \lambda_i(R)) \& (T_{D_M} > \lambda_i(M)) \quad (16)
\]

This result means there is a significant difference between \( R_{ori} \) and \( R_{upd} \), and between \( M_{ori} \) and \( M_{upd} \), thus, the soft sensor model has deteriorated. In contrast, when the condition is false, hypothesis \( H_2 \) is accepted. This result means that the soft sensor model has not deteriorated.

Because a sequential pipeline chemical process has random abnormal data, to ensure the practicality of the proposed soft sensor model deterioration detection method, constraints are applied to the statistical hypothesis test. These constraints are as follows:

The values of the statistical variables \( T(D_R) \) and \( T(D_M) \) should exceed the thresholds \( \lambda_i(R) \) and \( \lambda_i(M) \), respectively, three times consecutively to identify soft sensor model deterioration. The formula is as follows:
\[
\text{s.t.} \quad (T_{D_R}(w_{i,i}) > \lambda_i(R)) \& (T_{D_M}(w_{i,i}) > \lambda_i(M)), \quad i = 0, 1, 2 \quad (17)
\]

**Modeling sample data updating based on moving window-cosine similarity**

When the soft sensor model deteriorates due to the time-varying nature of chemical processes, the influence of each auxiliary variable on the dominant variable changes. Because the updated model sample data significantly affect the accuracy of the newly created soft sensor model [42], the influence of the original modeled auxiliary variable on the dominant variable of the current chemical process changes. The updated soft sensor model, which is created from the originally modeled auxiliary variable, results in limited improvement in the prediction performance. Additionally, high-dimensional auxiliary variables in chemical processes result in highly complex models and often incomplete model structures [56]. The moving window method can obtain the latest data representing the current process [57] and adaptively update that data [58], and the cosine similarity correlation calculation method can effectively measure the correlations between vectors [59] and identify the effectiveness of the influences [60]. Therefore, a sample data updating method using moving window-cosine similarity-based soft sensor modeling is proposed to update the sample datasets of soft sensor models for chemical processes and improve their prediction performance.

The width of the currently modeled auxiliary variable sample dataset window is the number of modeled samples in the original soft sensor model \( n(n \geq 1) \), the number of modeled auxiliary variables \( p(p \geq 1) \), and the number of chemical process auxiliary variables \( m(m \geq 1) \). The chemical process sample data in this window are as follows:
\[
DS = \{ u_i(k) | k = 1, 2, ..., m, i = 1, 2, ..., n \}
\]
\[
U(k) = \begin{bmatrix} u_{11}, u_{21}, \cdots, u_{m1} \\ u_{12}, u_{22}, \cdots, u_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1p}, u_{2p}, \cdots, u_{mp} \end{bmatrix}
\]

that is, the new chemical process sample data contain \( m \) sets of \( p \)-dimensional eigenvectors.

To properly select \( p \) modeled auxiliary variables from \( m \) auxiliary variables in a chemical process, the cosine similarity-based correlation calculation method is employed to obtain \( p \) modeled auxiliary variables with the strongest correlations with the dominant variable.

The cosine similarity [61] maps individual index data into a vector space and calculates the cosine of the angle between two vectors as the correlation between the two variables. A larger cosine of the intersection angle means that the two vectors are more similar, and a smaller value means that the two vectors are less similar [62].

The formula used to calculate the cosine similarity between the variables \( X = (x_1, x_2, \cdots, x_n) \) and \( Y = (y_1, y_2, \cdots, y_n) \) is as follows:
\[
\cos \langle X,Y \rangle = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}} \quad (20)
\]
where \( x_i \) and \( y_i \) represent the \( i \)-th elements of the variables \( x \) and \( y \), respectively.

The cosine similarity between the sample data of \( m \times n \) groups of auxiliary variables
\[(u_j), 1 \leq i \leq m; 1 \leq j \leq n \] and the sample data of the dominant variable \[y_j, 1 \leq j \leq n\] is calculated to obtain \[\frac{\partial}{\partial t} \xi = \nabla \sum_{i=1}^{n} \xi_i\].

Based on the number of soft sensor model samples \(p(p \geq 2)\), \(p\) auxiliary variables with the largest cosine similarities and auxiliary variable data updated via the moving window are selected as the samples for the soft sensor modeling:

\[
DS_{\text{new}} = \left\{ u_j \mid j = 1, 2, \ldots, n; i = 1, 2, \ldots, p \right\}
\]

\[
U(k) = \begin{bmatrix}
u_{11}, u_{12}, \ldots, u_{1p} \\
u_{21}, u_{22}, \ldots, u_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
u_{m1}, u_{m2}, \ldots, u_{mp}
\end{bmatrix}
\]

Based on the soft sensor model sample dataset \(DS_{\text{new}}\), a soft sensor model is created via the soft sensor modeling method.

The current model sample dataset is \(DS_{\text{updm}} = \{X_{\text{updm}}, Y_{\text{updm}}\}\), and the test dataset is \(DS_{\text{updm}} = \{X_{\text{updm}}, Y_{\text{updm}}\}\). The width of the current auxiliary variable test sample dataset window is set to \(w(w \geq 1)\); the number of modeled auxiliary variables is \(p(p \geq 1)\), and the number of chemical process auxiliary variables is \(m(m \geq 1)\). \(R_{\text{updm}}\) and \(M_{\text{updm}}\) are calculated for comparison with the prediction performance indexes \(R_{\text{ori}}\) and \(M_{\text{ori}}\), respectively, of the original soft sensor model to determine whether the prediction performance of the soft sensor model has recovered.

**Soft sensor modeling based on LSSVM**

An LSSVM is a machine learning method proposed by Suykens et al. [63] to solve function estimation problems. This method has excellent computational speed and convergence precision and is suitable for small data samples.

The LSSVM model is as follows:

\[
y(x) = \omega^T \varphi(x) + b
\]

where \(\varphi(x)\) is a nonlinear transformation function, \(\omega\) is the adjustable weight vector, and \(b\) is the offset value.

The objective function of the LSSVM [64] is as follows:

\[
\min J(\omega, \xi) = \frac{1}{2} \omega^T \omega + \frac{C}{2} \sum_{i=1}^{l} \xi_i^2
\]

s.t. \[y_i = \omega^T \varphi(x_i) + b + \xi_i \quad (i = 1, 2, \ldots, l)
\]

where \(x_i \in \mathbb{R}^p\) is the input vector, \(y_i \in \mathbb{R}\) is the corresponding output vector, \(\xi_i\) is the error between the system output value and the actual value, \(C > 0\) is the regularization parameter, \(\varphi(x)\) is the nonlinear mapping from the input space to the feature space, \(\omega\) is the system weight, \(b\) is the offset value, and \(s.t.\) represents a constraint.

Based on the Karush-Kuhn-Tucker (KKT) conditions, the Lagrange polynomial function for the optimization is solved. The LSSVM model for the function estimate is as follows:

\[
\hat{y} = f(x) = \sum_{i=1}^{l} \alpha_i K(x_i, x) + \beta
\]

The universal radial basis function is employed in this paper [65] as follows:

\[
k(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right)
\]

\(\sigma^2\) is the kernel radius. It is difficult to determine the regularization parameter \(C\) and radial basis function parameter \(\sigma^2\) in the LSSVM objective function, and this difficulty affects the regression performance of the LSSVM method. In this paper, PSO is employed. Minimizing the sum of the squared error between the actual sample output data \(y_k(t_k)\) and the predicted sample output \(\hat{y}(t_k)\) is defined as the optimization objective to optimize the parameters \(C\) and \(\sigma^2\). The objective function is as follows:

\[
\min J = \frac{1}{M} \sum_{k=1}^{M} \left\| y_k(t_k) - \hat{y}_k(t_k) \right\|^2
\]

**Model parameter optimization based on PSO**

PSO [53] is a swarm intelligence algorithm that simulates the predatory behaviors of flocks of birds and schools of fish.

In the particle swarm algorithm, numerous particles with individual positions, velocities and optimization objective fitnesses are configured. The particle positions and velocities are initialized randomly, and the parameters are optimized via iteration.

The iterative process includes the personal best value \((pbest)\), which represents the optimal fitness of an individual particle, and the global best value \((gbest)\), which represents the optimal fitness of all particles. These two extreme values \((pbest, gbest)\) are tracked for the particles, and the individual posit-
ions and velocities are updated. The update formula is as follows:

\[
v_{id}(t+1) = w \times v_{id}(t) + c_1 \times \text{rand()} \times (p_{id} - x_{id}(t)) + c_2 \times \text{rand()} \times (p_{gd} - x_{id}(t))
\]

(29)

\[
x_{id}(t+1) = x_{id}(t) + v_{id}(t+1)
\]

(30)

where \(v_{id}, x_{id},\) and \(p_{id}\) represent the velocity, position and \(p\text{best},\) respectively, of the \(i\)-th particle in the \(t\)-th iteration, \(\text{rand}()\) is a random number in the range \([0, 1]\), \(c_1\) and \(c_2\) are learning factors that represent the weights of the statistical acceleration terms that push each particle toward \(p\text{best}\) and \(g\text{best},\) \(w\) is the inertia weight, and \(t\) is the iteration number.

SIMULATION AND DISCUSSION

To test the proposed adaptive soft sensor modeling method, a nonlinear dynamic time-varying CSTR and a DCP are defined as the objects of study for the data simulation and analysis.

CSTR simulation and discussion

CSTRs have nonlinear dynamic characteristics \cite{66} and are widely employed to test the capabilities of soft sensor modeling methods for solving nonlinear and time-varying problems.

The definitions and steady-state values of the CSTR parameters \cite{67} are listed in Table 1.

The catalyst activity \(k_0\) in the reactor varies periodically. The obtained simulation dataset is normalized.

The soft sensor model structure is as follows:

\[
y_{id}(t_k) = f_{id}(u_i(t_k)), i = 1, 2, \ldots, p
\]

(31)

where \(u_i(t_k)\) represents the sample of the \(i\)-th auxiliary variable at moment \(t_k\), and \(y(t_k)\) represents the predicted dominant variable via the soft sensor model.

In the CSTR process, the concentration of raw material A in reactor \(C_A\) is selected as the dominant variable in the soft sensor model, and the feed flow rate \(F_f\), cooling water flow rate \(F_c\), and reactor interior temperature \(T_r\) are selected as the auxiliary variables of the soft sensor model for the simulation.

The sampling cycle of the auxiliary variables and dominant variable is set to 1 h, the simulation time is set to 90 h, and 20 groups of model sample data and 70 groups of test data are obtained. Gaussian white noise is applied to the simulation data. The PSO parameters are as follows: 100 iterations, 20 particles, \(C = 130\), and \(\sigma^2 = 2.167\).

The training performance of the original soft sensor modeling method is listed in Table 2.

### Table 2. Training performance of the soft sensor model

<table>
<thead>
<tr>
<th>Evaluating indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0256</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0298</td>
</tr>
</tbody>
</table>

The performance evaluation indexes in Table 2 show that the developed soft sensor model has good degrees of fit.

When equipment experiences mechanical wear and sensor deterioration, the prediction performance of the soft sensor model gradually degrades, and the evaluation index gradually increases. Although the RMSE and MAE curves fluctuate, the curves generally increase continuously, and the performance of the soft sensor model gradually deteriorates.

The \(\{D_{R}(l)\}^{l}_{l-1}\) and \(\{D_{G}(l)\}^{l}_{l-1}\) datasets are analyzed using the SPSS software, and the results are shown in Figure 2.

Based on the identifying conditions for the distribution and the Lindberg-Levy central limit theorem (CLT), the data distributions \(\{D_{R}(l)\}^{l}_{l-1}\) and \(\{D_{G}(l)\}^{l}_{l-1}\) are normal distributions.

### Table 1. Definitions and steady-state values of the CSTR parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Steady-state value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(F_f)</td>
<td>feed flow rate</td>
<td>100 L/min</td>
</tr>
<tr>
<td>(C_{A0})</td>
<td>reactant concentration in the feed</td>
<td>1 mol/L</td>
</tr>
<tr>
<td>(T_r)</td>
<td>feed temperature</td>
<td>350 K</td>
</tr>
<tr>
<td>(V)</td>
<td>volume of the reactor</td>
<td>100 L</td>
</tr>
<tr>
<td>(k_0)</td>
<td>reaction rate</td>
<td>7.2×1010 min^-1</td>
</tr>
<tr>
<td>(\rho)</td>
<td>reactant density</td>
<td>1000 g/L</td>
</tr>
<tr>
<td>(C_p)</td>
<td>specific heat capacity of the reactant</td>
<td>1 cal/g/K</td>
</tr>
<tr>
<td>(T_w)</td>
<td>temperature at the cooling water inlet</td>
<td>350 K</td>
</tr>
<tr>
<td>(C_{pw})</td>
<td>specific heat capacity of the cooling water</td>
<td>1 cal/g/K</td>
</tr>
</tbody>
</table>
In this study, the width of the moving window is 10, $\alpha = 0.01$, and there are 9 degrees of freedom. The soft sensor model deteriorates when $T(D_R) > 2.8214$ and $T(D_M) > 2.8214$; otherwise, the soft sensor model experiences no deterioration.

For the statistical hypothesis test based on the soft sensor model deterioration detection method proposed, the curves of $T(D_R)$ and $T(D_M)$ are shown in Figure 3. Figure 3 shows that although there are abnormal fluctuations, $T(D_R)$ and $T(D_M)$ increase steadily. Based on the selected threshold $\lambda_t(R) = \lambda_t(M) = 2.8214$, when the moving window proceeds to step $l = 7$, $T(D_R) = 2.870572$, and $T(D_M) = 4.44494$; both values exceed the threshold of 2.8214. However, when the moving window proceeds to step $l = 8$, $T(D_R) = 1.595584$, and $T(D_M) = 2.735865$; neither value exceeds the threshold of 2.8214. At this point, the original data are abnormal, and the soft sensor model experiences no deterioration. When the moving window proceeds to step $l = 3$, $T(D_R) = 2.846904$, and $T(D_M) = 2.954331$; both values exceed the threshold of 2.8214. Based on the detection method, the moving window proceeds to step $l = 14$, where $T(D_R) = 3.435871$, and $T(D_M) = 3.537272$; both values exceed the threshold of 2.8214. The moving window then proceeds to step $l = 15$, where $T(D_R) = 3.938894$, and $T(D_M) = 3.948832$; both values again exceed the threshold of 2.8214.

In summary, when the moving window proceeds from step $l = 14$ to step $l = 16$, $T(D_R)$ and $T(D_M)$ exceed the threshold $\lambda_t(R) = \lambda_t(M) = 2.8214$, which satisfies the soft sensor model deterioration criterion; i.e., the current soft sensor model has deteriorated.

From Figure 3, the value of $T(D_M)$ in steps 5, 6 and 7 exceeds the threshold of 2.8214. However, the value of $T(D_R)$ exceeds the threshold of 2.8214 in step 7, and the values of $T(D_R)$ are lower than the threshold of 2.8214 in steps 5 and 6. The model of soft sensor does not deteriorate. The deterioration detection method only under $MAE$ is inaccurate. Instead, the constrained deterioration detection method based on $RMSE$ and $MAE$ can be more accurate to evaluate the deterioration of soft sensor.

After the soft sensor model deterioration, based on the cosine similarity-based auxiliary variable correlation identification method proposed, three new auxiliary variables are selected based on the auxiliary variables of the current sample data, including the feed flow rate $F_i$, reactor interior temperature $T_r$, cooling water flow rate $F_c$, feed temperature $T_i$ and cooling water temperature $T_c$. The modeling dataset of the current sample is created.
Based on Eq. (20), the cosine similarities between the following auxiliary variables and the dominant variable are calculated and listed in Table 3.

**Table 3. Cosine similarities of the auxiliary variables**

<table>
<thead>
<tr>
<th>Auxiliary variable</th>
<th>Cosine similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fi</td>
<td>0.8326</td>
</tr>
<tr>
<td>Tr</td>
<td>0.8650</td>
</tr>
<tr>
<td>Fc</td>
<td>0.8382</td>
</tr>
<tr>
<td>Ti</td>
<td>0.8030</td>
</tr>
<tr>
<td>Tci</td>
<td>0.8559</td>
</tr>
</tbody>
</table>

The cosine similarities between the auxiliary variables and the dominant variable in Table 3 show that the three auxiliary variables that are strongly correlated with the dominant variable concentration $C_A$ are the reactor interior temperature $T_r$, cooling water flow rate $F_c$ and cooling water temperature $T_{ci}$.

The soft sensor model is created based on the updated training sample set.

The training performance of the updating soft sensor modeling method is listed in Table 4.

**Table 4. Training performance of the soft sensor model**

<table>
<thead>
<tr>
<th>Evaluating indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0030</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

Based on the RMSE and MAE before the model deterioration, the curves of the RMSE and MAE for when the model has deteriorated and when the model has been updated, are shown in Figure 4.

The RMSE and MAE curves in Figure 4a and b show that the prediction performance of the soft sensor model has recovered and that the data prediction accuracy has improved significantly compared with the case in which deterioration occurs.

To compare the fitting and prediction performance of the soft sensor model created from the method with the cosine similarity-based auxiliary variable selection and that without the auxiliary variable selection, the feed flow rate $F_i$, cooling water flow rate $F_c$, and feed temperature $T_i$ are selected as the auxiliary variables, and the concentration $C_A$ is selected as the dominant variable to create the soft sensor model.

The training performance of the soft sensor modeling method is listed in Table 5.

**Table 5. Training performance of the soft sensor model**

<table>
<thead>
<tr>
<th>Evaluating indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0074</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0061</td>
</tr>
</tbody>
</table>

The abilities of the soft sensor models to predict data and obtain the relevant RMSE and MAE are compared based on the test sample set. A comparison of the prediction performance evaluation indexes of the updated soft sensor model is shown in Figure 5.

Tables 4 and 5, and Figure 5a and b show that the degree of fit and data prediction accuracy of the soft sensor model created by the method that uses cosine similarity-based auxiliary variable selection are superior to those of the soft sensor model created using random auxiliary variable selection.

The CSTR is simulated. In the first stage, a normal production state is simulated. In the later stage, temperature sensor performance decline-induced feed temperature and cooling water temperature fluctuations are simulated to cause soft sensor model deterioration.

The soft sensor model evaluation indexes in Table 2 show that the soft sensor model created from the original training sample has a superior degree of fit. The $\mathcal{T}(D_i)$ and $\mathcal{T}(D_{ii})$ curves in Figure 3 show that the performance of the soft sensor model gradually deteriorates.
deteriorates, and this deterioration gradually becomes more severe. When the moving window proceeds to step \( l = 14 \), the data prediction produces severe errors; the model performance deterioration exceeds the critical point, the soft sensor model deteriorates, and the model should be updated.

The \( RMSE \) and \( MAE \) performance evaluation curves in Figure 4 show that the performance of the soft sensor model has recovered.

Comparisons of the soft sensor model evaluation indexes in Tables 4 and 5 show that the modeled auxiliary variables selected based on the cosine similarity have better degrees of fit. Compared with the soft sensor model created from random selection of the modeled auxiliary variables, the \( RMSE \) improves by 59.46%, the \( MAE \) improves by 68.85%, and the fitting performance improves.

A comparison of the \( RMSE \) and \( MAE \) curves in Figure 5a and b shows that the soft sensor model created from the modeled auxiliary variables selected based on the cosine similarity has better prediction results and is significantly superior to the prediction curve without the cosine similarity.

### DCP simulation and discussion

In this subsection, the results of the application of the proposed adaptive soft sensor to an actual dataset from a debutanizer column process are presented. Usually, the concentration of bottom butane is measured online by a gas chromatography analyzer mounted at the top of the tower. However, because it takes time for the bottom butane vapor to reach the top of the column and for the analysis to be performed, the online measurement has a time delay.

The dataset from the DCP can be downloaded from Springer; this dataset has become a standard for testing the performance of adaptive soft sensor models. Table 6 shows the definitions of the DCP parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u_1 )</td>
<td>Tower top temperature</td>
</tr>
<tr>
<td>( u_2 )</td>
<td>Tower top pressure</td>
</tr>
<tr>
<td>( u_3 )</td>
<td>Backflow</td>
</tr>
<tr>
<td>( u_4 )</td>
<td>Flow to the next process</td>
</tr>
<tr>
<td>( u_5 )</td>
<td>Layer 6 tray temperature</td>
</tr>
<tr>
<td>( u_6 )</td>
<td>Tower bottom temperature</td>
</tr>
<tr>
<td>( u_7 )</td>
<td>Tower bottom pressure</td>
</tr>
<tr>
<td>( y )</td>
<td>Concentration of butane</td>
</tr>
</tbody>
</table>

In this section, 450 samples are selected from the first half of the dataset as the historical database; 400 samples are the modeled dataset, and the other 50 samples are used as the original test dataset. A total of 500 samples are selected from the second half of the dataset to test the model deterioration and updating effect. The modeling variables of the original soft sensor model are \( u_1, u_2, u_3 \) and \( u_4 \) and were selected by experts.

The LSSVM model parameters are set as \( C = 129.86 \) and \( \sigma^2 = 2.115 \).

The training performances and predicted performances of the soft sensor by the modeling data are given in Table 7. The results show that the soft sensor model has good fitting performance. Notably, because of the noise in the DCP data, the predicted performance of the soft sensor model is not very good. Based on the test samples with the moving window, the \( RMSE \) and \( MAE \) values increase with increasing test samples. These results show that the prediction performance of the soft sensor model det-

### Table 6. Definitions of the DCP parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u_1 )</td>
<td>Tower top temperature</td>
</tr>
<tr>
<td>( u_2 )</td>
<td>Tower top pressure</td>
</tr>
<tr>
<td>( u_3 )</td>
<td>Backflow</td>
</tr>
<tr>
<td>( u_4 )</td>
<td>Flow to the next process</td>
</tr>
<tr>
<td>( u_5 )</td>
<td>Layer 6 tray temperature</td>
</tr>
<tr>
<td>( u_6 )</td>
<td>Tower bottom temperature</td>
</tr>
<tr>
<td>( u_7 )</td>
<td>Tower bottom pressure</td>
</tr>
<tr>
<td>( y )</td>
<td>Concentration of butane</td>
</tr>
</tbody>
</table>

### Table 7. Performance of the soft sensor model

<table>
<thead>
<tr>
<th>Performance of soft sensor</th>
<th>( RMSE )</th>
<th>( MAE )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training performance</td>
<td>0.0390</td>
<td>0.0248</td>
</tr>
<tr>
<td>Predicted performance</td>
<td>0.0903</td>
<td>0.0730</td>
</tr>
</tbody>
</table>
iorates gradually. However, the sig values of 0.017 and 0.021 from SPSS indicate that $\{D_r(j)\}_{j=1}^L$ and $\{D_m(j)\}_{j=1}^L$ are normal distributions.

As shown in Figure 6, the $T(D)$ values of the RMSE and MAE are both greater than 2.8214 at steps 22, 23 and 24. Based on the soft sensor model deterioration detection method, deterioration of the original soft sensor model has been identified. Although the RMSE and MAE values exceed 2.8214 at some points in the early steps of the testing process, they do not exceed the threshold for three continuous steps, so they do not meet the deterioration criteria.

However, without the constrained conditions, the deterioration of the original soft sensor model will be determined by the statistical hypothesis test in step 5. In contrast, the RMSE and MAE have normal values in steps 9 to 17. The soft sensor model has not deteriorated, which is demonstrated by these values. A misjudgment occurs with the statistical hypothesis test.

To update the original soft sensor model, the moving window is used to update the DCP samples. Based on the updated DCP data samples in the moving window, the cosine similarity method is used to select the modeling variables of the updated soft sensor due to changes in the influence of the auxiliary variables on the principal variables. The values of the cosine similarity between the process variables and dominant variables are shown in Table 8.

According to the values of the process variables in Table 8, $u_2$, $u_4$, $u_6$ and $u_7$ are chosen to establish the soft sensor model. Moreover, to demonstrate the superiority of the soft sensor model established using the proposed method, additional groups are added for tests.

The performance of the updated soft sensor model is shown in Table 9. The numbers of modeling samples and test samples are 400 and 50, respectively. A comparison of Table 9 and Table 7 shows that the performances of the soft sensor models are similar. Additionally, based on the RMSE and MAE values over 10 stepping tests, the RMSE and MAE curves for the undeteriorated, deteriorated and updated models are shown in Figure 7. The simulation process and analysis indicate that the performance of the soft sensor model has been restored.

### Table 8. Values of the cosine similarity between the process variables and dominant variables

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Cosine similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_1$</td>
<td>0.8146</td>
</tr>
<tr>
<td>$u_2$</td>
<td>0.8466</td>
</tr>
<tr>
<td>$u_3$</td>
<td>0.8183</td>
</tr>
<tr>
<td>$u_4$</td>
<td>0.8996</td>
</tr>
<tr>
<td>$u_5$</td>
<td>0.7868</td>
</tr>
<tr>
<td>$u_6$</td>
<td>0.8205</td>
</tr>
<tr>
<td>$u_7$</td>
<td>0.8232</td>
</tr>
</tbody>
</table>

### Table 9. Performance of the soft sensor model

<table>
<thead>
<tr>
<th>Performance of soft sensor</th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training performance</td>
<td>0.0385</td>
<td>0.0244</td>
</tr>
<tr>
<td>Predicted performance</td>
<td>0.0701</td>
<td>0.0527</td>
</tr>
</tbody>
</table>

To further validate the proposed adaptive modeling strategy and analyze the influence of statistical fluctuations, we use K-fold cross-validation to perform several experiments. In the cross-validation, $u_1$, $u_2$, $u_3$ and $u_4$ updated data samples based on the moving window, $u_4$, $u_5$, $u_6$ and $u_7$ updated data samples selected randomly, updated data samples based on the AANN and $u_2$, $u_3$, $u_4$ and $u_5$ updated data samples based on the proposed moving window-cosine similarity are used to test the model effect. The parameters of AANN method are $goal = 0.001$, $epochs = 5000$, $lr = 0.05$, and $MN = 4$.

![Fig. 6. Curves of $T(D_r)$ and $T(D_m)$ of the DCP.](image-url)
In order to reduce the fitting effects of soft sensor models, the soft sensor models in K-fold cross-validation have the similar values of RMSE and MAE.

Based on 10 stepping tests, the error bars for the RMSE and MAE of the DCP are shown in Figure 8. Over 10 stepping tests, the RMSEs and MAEs of u2, u4, u6, and u7 based on moving window-cosine similarity are 0.07697 and 0.05887, respectively, which are the lowest average values of the three modeling variables. Moreover, only the RMSE value of u2, u4, u6, and u7 is lower than the original value of the soft sensor. Although the MAE value of u2, u4, u6, and u7 is the smallest, the fluctuation is larger than that of u1, u2, u4, and u7, which may be caused by noise in the data.

Therefore, the 10 stepping tests demonstrated that the performance of the soft sensor model has been restored and that the soft sensor model based on variables selected using moving window-cosine similarity can achieve the best data prediction.

CONCLUSIONS

To address the time-varying issues in chemical processes, which impede the construction of high-accuracy soft sensors, an adaptive soft sensing methodology has been proposed. This study provides two main contributions.

First, we presented a soft sensor model deterioration detection method that uses the moving window-constrained statistical hypothesis test, which generates an up-to-date local model test set and adaptively evaluates the model deterioration. In addition, the decision data of the RMSE and MAE and the decision constraint strategy of exceeding the threshold three consecutive times can effectively reduce the misidentification of model deterioration for real data from a DCP and simulated data from a CSTR model.

Second, a moving window-cosine similarity method was proposed to update modeling samples due to the time-varying characteristics of chemical processes. Compared with the basic moving window method, the soft sensor model based on the sample data obtained by the proposed method provides a better fit and higher prediction accuracy for real data from a DCP and simulated data from a CSTR model. Further validation of the DCP through multiple experiments using the cross-validation technique demonstrated that the soft sensor model based on samples selected by moving window-cosine similarity can achieve the best data prediction.
The results indicate that the developed soft sensors can be used for time-varying chemical processes instead of traditional soft sensor modeling methods. Furthermore, the developed soft sensors can be beneficial to the implementation of advanced process control systems.

REFERENCES


[10] Z. Ge, B. Huang, Z. Song, Mixture semisupervised principal component regression model and soft sensor application, AIChE J. 60 (2014) 533-545


[16] M. Zhang, X. Liu, A real-time model based on optimized least squares support vector machine for industrial polypropylene melt index prediction, J. Chemom. 30 (2016) 324-331


[26] W. Shao, Adaptive soft sensor modeling based on local learning, China University of Petroleum, 2016, pp. 12,13

[27] X. Yuan, Soft sensing for complicated nonlinear processes based on just-in-time learning, Zhejiang University, 2016, pp. 76,77


[29] M. Galiaskarov, V. Kurkina, L. Rusinov, Online diagnostics of time-varying nonlinear chemical processes using moving window kernel principal component analysis and fisher discriminant analysis, J. Chemometr. 31 (2017) e2866


[32] Y. Fu, W. Yang, O. Xu, L. Zhou, Soft sensor modelling by time difference, recursive partial least squares and


[48] H. Kaneko, and K. Funatsu, Classification of the degradation of soft sensor models and discussion on adaptive models., AICHE J. 59 (2013) 2339-2347


[56] P. Cao, X. Luo, Modeling of soft sensor for chemical process, CIESC J. 64 (2013) 788-800


LONGHAO LI
YONGSHOU DAI
College of Information and Control Engineering, China University of Petroleum (East China), Qingdao, Shandong, China

NAUČNI RAD


PROCENA POGORŠANJA ADAPTIVNOG SOFTVERSKOG SENZORA I MODEL KOJI AZURIRA METODO ZA VREMENSKI PROMENLJIVE PROCESE

Zbog vremenski promenljive prirode hemijskih procesa, metode modelovanja adaptivnog softverskog senzora se pogoršavaju, a tačnost predviđanja podataka smanjuje. Da bi se rešio ovaj problem, predlaže se metoda adaptivnog modelovanja softverskog senzora koja ne samo da procenjuje pogoršanje modela pomoću statističkog testa hipoteze ograničene adaptivno pokretnih prozora, već i adaptivno ažurira uzorke modelovanja korišćenjem sličnosti kosinusa pokretnih prozora. Prvo, ova metoda procenjuje pogoršanje modela putem pozicioniranja pomoću ograničenog statističkog testiranja hipoteze zasnovane na razlikama između podataka o indeksu performansi predviđanja dobivenih iz pomeraja pokretnih prozora i originalnih indeksa procene performansi predviđanja. Uz to, dinamička vremenska promena u hemijskim procesima izaziva promene uticaja pomoćnih promenljivih na dominantnu promenljivu, a taj efekat ograničava poboljšanje tačnosti predviđanja modela softverskog senzora ažuriranjem samo podataka za pomoćne promenljive. Metoda sličnosti kosinusa pokretnih prozora se kombinuje da bi se predložila strategija koja ažurira i modelovane pomoćne promenljive i podatke za pomoćne promenljive. Napokon, parametri softverskog senzora su optimizovani pomoću optimizacije rojeva čestica kako bi se poboljšale performanse fitovanja. Simulirani podaci kontinualnog reaktora sa mešanjem i stvarni podaci za kolonu za uklanjanje butana korišćeni su za verifikaciju modela za procenu performansi predložene metode modelovanja softverskog senzora, a rezultati pokazuju njegovu efikasnost.

Ključne reči: hemijski proces, adaptivni model softverskog senzora, statistički test hipoteze, pokretni prozor, sličnost kosinusa, optimizacija rojeva čestica.