External Analysis-Based Regression Model for Robust Soft Sensing of Multimode Chemical Processes

Zhiqiang Ge and Zhihuan Song
Dept. of Control Science and Engineering, State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control, Zhejiang University, Hangzhou 310027, Zhejiang P.R. China

Manabu Kano
Dept. of Systems Science, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan

DOI 10.1002/aic.14253
Published online October 18, 2013 in Wiley Online Library (wileyonlinelibrary.com)

To remove the influence of operation mode changes in the chemical process, the whole variable set is partitioned into external, main, and quality variables. External variables are related to the operation mode. Two regression models are initially developed between external variables and main variables/quality variables, based on which the influence of the operation mode is removed from both input and output of the soft sensor. Then, an additional regression model is constructed for soft sensing, which is robust to the change of the operation mode. Compared to existing methods, the new method has advantages to handle two critical issues: (1) capable of quality estimation in new process modes; (2) able to distinguish changes in operation modes from process faults. Besides, a monitoring and analysis strategy is proposed for performance evaluation of the new soft sensor. Two case studies are provided to illustrate the efficiency of the proposed method. © 2013 American Institute of Chemical Engineers *AIChE J*, 60: 136–147, 2014
Keywords: external analysis, robust soft sensor, quality estimation, multimode chemical processes, performance monitoring

Introduction

Due to the difficulty of measuring quality variables in the chemical process, they are often determined by offline analyses in the laboratory or by some online analyzers. Nevertheless, both offline analyses and online analyzers are expensive or time-consuming which may introduce a significant delay to the control system. Hence, online quality estimation and prediction, which are also known as soft sensing or inference sensor, are crucial to realize feedback control of quality.[1] In the past years, online quality estimation and prediction methods have been widely studied in both academy and industrial areas. Particularly, with the wide utilization of the distributed control system in modern industrial processes, a huge number of process data have been collected, on which the data-based soft sensing methods have gained much attentions.[2–4] Different from the model-based method which greatly relies on the first-principle model of the process, the data-based method is much more knowledge-free. Representative data-based methods includes multiple regression analysis, linear latent regression methods such as principal component regression (PCR) and partial least squares (PLS), nonlinear regression methods such as artificial neural network and support vector machine, and so forth.[5–19]

Correspondence concerning this article should be addressed to Z. Ge at gezhiqiang@zju.edu.cn.

However, a drawback of many existing soft sensors is their static or nonadaptive nature, which makes the soft sensor models remain unchanged after they have been developed, even though process characteristics may change frequently due to set-point changes, operation grade changes, catalyst deactivation, seasonal effects, equipment aging, and so on. As a result, the online quality estimation/prediction performance will be deteriorated if the soft sensor models do not have adaptive nature. In fact, a recent questionnaire survey revealed that the most important problem of current soft-sensors was how to cope with changes in process characteristics and keep high estimation accuracy for a long period of time, that is, model maintenance.[20]

To handle the time-varying behaviors including operation mode change, many model maintenance methods have been proposed in recent years, such as adaptive modeling approaches, model library-based methods, local model-based methods, and so forth. Examples of adaptive modeling approaches for modeling maintenance include recursive and adaptive principal component analysis (PCA), recursive PLS, moving window-based PCA and PLS, exponentially weighted moving average-based PCA, and so on.[21–23] For model maintenance of processes with operation mode changes, the model library-based method was developed, where a model library is defined as a collection of models identified by different operation modes.[24–26] Then, the quality variable can be predicted by its corresponding soft sensor model in each operation mode. Another useful tool for modeling time-varying processes is the local model-based approach, in which local models are trained on limited partitions of the historical dataset.[27–30] By constructing an online model for the current data sample in a specific operation region, this approach provides a set of simple models based on which both of the time-varying and the nonlinear behaviors can be well-addressed. Other methods for model maintenance include fuzzy model-based method, multiblock-based models, and so forth.

Despite of the development of many model maintenance methods, these methods have several shortcomings which limit their applications in industrial processes. For example, the adaptive method is often carried out blindly, which means that continuous updating is performed whether a process change has been identified. Furthermore, there is also a risk that the adaptive model will be accommodated to faulty process conditions if some abnormal data samples are used for model updating. In other words, the adaptive modeling approach does not provide any mechanism to distinguish normal changes in operation condition and faults. The main shortcoming of the model library-based method is its limited ability to cope with new process changes. This is because the model library has been built on existing operation conditions. However, when a new change is introduced into the process, there is no model in the library corresponding to such operation condition. Similarly, the local modeling-based method is also limited in handling the new process change and distinguish between normal and faulty conditions.

In practice, an efficient adaptive soft sensor model should also cover the following two critical issues: (1) to handle new process changes; (2) to distinguish normal process changes and faulty conditions. Recently, a method which is called external analysis has been proposed for monitoring processes with multiple operation modes.[31,32] By removing the influence of the mode-related variables from the main process variables, the mode behavior of the process can be effectively eliminated. Therefore, a single monitoring model can be constructed for different operation conditions. Besides, the faulty condition of the process can also be monitored. However, due to the offline modeling structure of the external analysis-based method, new operation modes of the process may not be modeled very well. Also, the method is specially designed for the monitoring purpose, no quality variables or their relationships to the external variables have been considered for regression modeling and soft sensing.

Motivated by the external analysis method, a novel adaptive soft sensor is proposed in this article for online quality estimation of multimode chemical processes. First, the information of the quality variables is incorporated into the external analysis method, thus the process variables can be partitioned into three categories: external variables, main variables, and quality variables. By removing the influences of the external variables in both of the main and quality variables, an external analysis based regression (EAR) model is developed for soft sensing. Second, a moving-window approach is introduced into the EAR model, depending on which an online implementation procedure can be formulated for external analysis of new process changes. Third, an additional performance monitoring and analysis strategy is developed, which is based on a one-class classification method.

The remainder of this article is structured as follows. In section Preliminaries, preliminaries of the traditional PCR model and the support vector data description (SVDD) method are introduced. A detailed description of the proposed soft sensor is provided in section Methodology, including the development of the EAR model, the moving-window based adaptive soft sensor, and the SVDD-based performance monitoring and analysis strategy. To evaluate the feasibility and efficiency of the proposed soft sensor, two case studies are demonstrated in section Illustration and Results. Finally, conclusions are made.

Preliminaries

Principal component regression

Suppose the measurement matrix can be represented as $\mathbf{X} \in R^{n\times m}$, where n is the number of data samples and m is the number of measured variables. The predicted variable matrix can be given as $\mathbf{Y} \in R^{n\times r}$, where n is the number of data samples and r is the number of predicted variables. The traditional derivation of PCR can be expressed as follows[4]

$$\mathbf{X}=\mathbf{T}\mathbf{P}^T+\mathbf{E} \tag{1}$$

$$\mathbf{Y}=\mathbf{T}\mathbf{C}^T+\mathbf{F} \tag{2}$$

where $\mathbf{P} \in R^{m\times q}$ is the loading matrix, $\mathbf{T} \in R^{n\times q}$ is the principal component matrix, q is the selected number of principal components, $\mathbf{C} \in R^{r\times q}$ is the regression matrix, and $\mathbf{E}$ and $\mathbf{F}$ are the residuals matrices with appropriate dimensions.

Support vector data description

SVDD first maps the data from the original space to the feature space $\Phi : \mathbf{x} \rightarrow \mathbf{F}$, then a hypersphere with the minimum volume which separates the transferred data from the rest of the feature space is found in the feature space. To construct the minimum volume of the hypersphere, SVDD intends to solve the following optimization problem[33]

$$\begin{gathered}\min_{R,a,\xi} R^2+C\sum_{i=1}^{n}\xi_i \\ s.t. \|\Phi(\mathbf{x}_i)-\mathbf{a}\|^2 \le R^2+\xi_i, \xi_i \ge 0 \quad i=1,2,\ldots,n\end{gathered} \tag{3}$$

where R and $\mathbf{a}$ are radius and center of the hypersphere, C gives the trade-off between the volume of the hypersphere and the number of errors, and ξ_i represents the slack variable. The dual form of the optimization problem can be obtained as[33]

$$\begin{gathered}\min_{\alpha_i} \sum_{i=1}^{n}\alpha_i K(\mathbf{x}_i,\mathbf{x}_j)-\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j K(\mathbf{x}_i,\mathbf{x}_j) \\ s.t. \quad 0 \le \alpha_i \le C, \sum_{i=1}^{n}\alpha_i=1\end{gathered} \tag{4}$$

where $K(\mathbf{x}_i,\mathbf{x}_j)=\langle\Phi(\mathbf{x}_i),\Phi(\mathbf{x}_j)\rangle$ is a kernel function which is introduced to compute the inner product in the feature space α_i is a Lagrange multiplier.

Methodology

In this section, the proposed soft sensor is detailed. First, the EAR model is developed, which can effectively incorporate the quality information. Then, the adaptive soft sensor is constructed based on the EAR model. For online implementation of the proposed method, a moving-window approach is employed into the EAR-based soft sensor. Furthermore, a process monitoring scheme is developed for fault detection of the process.

EAR model

In most cases, changes in the operating condition can be assumed to be brought from the outside of the process. These changes include process inputs such as a feed flow rate, set-point change of controllers, and change of the material. According to the external analysis method, process variables can be classified into two groups: external variables and main variables. Although the external variable set consists of variables related to operating condition changes, the main variable set is formed by other process variables. The normal variations of operating conditions are reflected in the external variables, thus changes in external variables should not be considered as process faults. To distinguish normal process changes and faulty conditions, both of the changes in external variables and their influences on main variables should be removed. To this end, the information of main variables can be decomposed into two different parts: the part explained by external variables and another unexplained part. The concept of the external analysis method can be illustrated in Figure 1.

For quality estimation purpose, the quality variables should be incorporated into the soft sensor model. Therefore, except for the external and main variables, there is an additional variable set which consists of quality-related variables. To remove the influent of the operation mode change in the soft sensor model, impacts of the external variables on both main variables and quality variables should be considered. After these two impacts have been removed, a regression model can be constructed between the residuals of the main variable set and the quality variable set. The concept of the EAR model can be illustrated in Figure 2.

Given the historical dataset $\mathbf{Z} \in \mathbf{R}^{n\times m}$, where n and m are numbers of samples and variables in the process. Suppose this process dataset contains C operation modes, each operation mode has $n_c, c=1,2,\ldots,C$ data samples, thus $\sum_{c=1}^{C} n_c = n$. Due to the new external analysis method, m process variables can be classified into three different parts: external variables, main variables, and quality variables. Suppose m_u variables are classified as external variables, m_x variables are classified as main variables, and m_y variables are classified as quality variables, thus $m_u+m_x+m_y=m$. Therefore, the historical data matrix can be decomposed as:

$$\mathbf{Z}=[\mathbf{U}\ \mathbf{X}\ \mathbf{Y}]=\begin{bmatrix} \mathbf{U}_1 & \mathbf{X}_1 & \mathbf{Y}_1 \\ \mathbf{U}_2 & \mathbf{X}_2 & \mathbf{Y}_2 \\ & \cdots & \\ \mathbf{U}_C & \mathbf{X}_C & \mathbf{Y}_C \end{bmatrix} \tag{5}$$

where $\mathbf{U}$, $\mathbf{X}$, and $\mathbf{Y}$ consists of external variables, main variables, and quality variables in different operation modes, respectively. To separate the influence of external variables on the main variables dataset, the information of $\mathbf{X}$ should be decomposed into two parts: a part explained by external variables and a nonexplained part. Therefore, a regression model should be built for information decomposition of the main variable dataset. For linear processes, multiple linear regression tools such as least squares, PCR, and PLS methods can be used. A more complicated situation is that the relationships between external and main variables are nonlinear; in this case, a nonlinear regression model should be employed. However, in the present work, we only focus on the linear relationship between different variable sets. Particularly, the widely used PCR model is employed for regression model construction between external and main variables dataset, which is given as follows

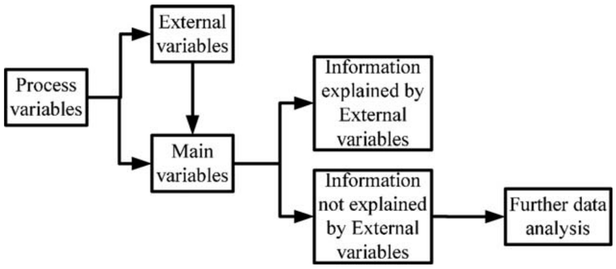

Figure 1. The concept of external analysis method.

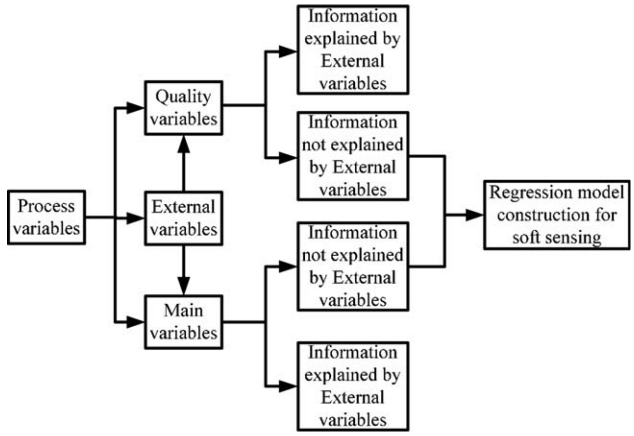

Figure 2. The concept of external analysis-based regression modeling method.

$$\mathbf{U}=\mathbf{T}_x\mathbf{P}_x^T+\mathbf{E}_x \tag{6}$$

$$\mathbf{X}=\mathbf{T}_x\mathbf{C}_x^T+\mathbf{F}_x \tag{7}$$

where $\mathbf{P}_x \in R^{m_u\times q_x}$ is the loading matrix of $\mathbf{U}$, $\mathbf{T}_x \in R^{n\times q_x}$ is the principal component matrix of $\mathbf{U}$, q_x is the selected number of principal components corresponding to the main variable PCR model, which can be determined by the cumulative percentage variance (CPV) method. $\mathbf{C}_x \in R^{m_x\times q_x}$ is the regression matrix, $\mathbf{E}_x$ and $\mathbf{F}_x$ are the residuals matrices with appropriate dimensions. Therefore, based on the PCR model, the information of the main variable dataset $\mathbf{X}$ can be decomposed into the following two parts: $\mathbf{T}_x\mathbf{C}_x^T$ corresponds to the part explained by the external variables, and $\mathbf{F}_x$ is the unexplained information, which has no relation with the mode change of the process.

Similarly, another PCR model can also be built between the external variables and the quality variables, which is given as

$$\mathbf{U}=\mathbf{T}_y\mathbf{P}_y^T+\mathbf{E}_y \tag{8}$$

$$\mathbf{Y}=\mathbf{T}_y\mathbf{C}_y^T+\mathbf{F}_y \tag{9}$$

where $\mathbf{P}_y \in R^{m_u\times q_y}$ is the loading matrix, $\mathbf{T}_y \in R^{n\times q_y}$ is the principal component matrix, q_y is the selected number of principal components corresponding to the quality variable PCR model, $\mathbf{C}_y \in R^{m_y\times q_y}$ is the regression matrix, and $\mathbf{E}_y$ and $\mathbf{F}_y$ are the residuals matrices with appropriate dimensions. As a result, the information of the quality variable dataset $\mathbf{Y}$ can also be decomposed into the two parts: $\mathbf{T}_y\mathbf{C}_y^T$ corresponds to the part explained by the external variables, and $\mathbf{F}_y$ is the unexplained information, which has also no relation with the mode change of the process. For the PCR model, it is noticed that $\mathbf{P}_x$ and $\mathbf{T}_x$ are same as $\mathbf{P}_y$ and $\mathbf{T}_y$, whereas they are maybe different under other model structures.

After the unexplained information of both main variables and quality variables have been separated, an additional PCR regression model can be developed to describe the relationship between them, which is given as

$$\mathbf{F}_x=\mathbf{T}_{x-y}\mathbf{P}_{x-y}^T+\mathbf{E}_{x-y} \tag{10}$$

$$\mathbf{F}_y=\mathbf{T}_{x-y}\mathbf{C}_{x-y}^T+\mathbf{F}_{x-y} \tag{11}$$

where $\mathbf{P}_{x-y} \in R^{m_x \times q_{x-y}}$ is the loading matrix of $\mathbf{F}_x$, $\mathbf{T}_{x-y} \in R^{n \times q_{x-y}}$ is the principal component matrix of $\mathbf{F}_x$, q_{x-y} is the selected number of principal components, $\mathbf{C}_{x-y} \in R^{m_y \times q_{x-y}}$ is the regression matrix, and $\mathbf{E}_{x-y}$ and $\mathbf{F}_{x-y}$ are the residuals matrices of the PCR model. Therefore, a total of three PCR models have been built to construct the external analysis regression model, which are denoted as $\mathrm{PCR}^{\text{x-u}}$, $\mathrm{PCR}^{\text{y-u}}$, and $\mathrm{PCR}^{\text{y-x}}$ in this article.

Adaptive soft sensor based on EAR model

Based on the proposed EAR model, an adaptive soft sensor is developed in this subsection. Given a new data sample $\mathbf{r}_{\text{new}}=[\mathbf{u}_{\text{new}}\ \ \mathbf{x}_{\text{new}}]$, which only consists of external and main variables, to estimate the quality variable, the influence of the external variable should be removed from the main variable in the first step. Therefore, based on the $\mathrm{PCR}^{\text{x-u}}$ model, the unexplained main variable information can be extracted as

$$\begin{aligned}\mathbf{t}_{x,\text{new}}&=\mathbf{P}_x^T\mathbf{u}_{\text{new}}\\ \mathbf{f}_{x,\text{new}}&=\mathbf{x}_{\text{new}}-\mathbf{C}_x\mathbf{t}_{x,\text{new}}=\mathbf{x}_{\text{new}}-\mathbf{C}_x\mathbf{P}_x^T\mathbf{u}_{\text{new}}\end{aligned} \tag{12}$$

In the second step, the $\mathrm{PCR}^{\text{y-x}}$ model is employed for calculating of the estimated unexplained quality variable for the new data sample, which is given as

$$\begin{aligned}\mathbf{t}_{x-y,\text{new}}&=\mathbf{P}_{x-y}^T\mathbf{f}_{x,\text{new}}\\ \hat{\mathbf{f}}_{y,\text{new}}&=\mathbf{C}_{x-y}\mathbf{t}_{x-y,\text{new}}=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{f}_{x,\text{new}}\\ &=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{x}_{\text{new}}-\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{C}_x\mathbf{P}_x^T\mathbf{u}_{\text{new}}\end{aligned} \tag{13}$$

Then, the explained part of the quality variable by the external variables should be added back to form the final estimated quality variable, which are formulated as follows

$$\begin{aligned}\mathbf{t}_{y,\text{new}}&=\mathbf{P}_y^T\mathbf{u}_{\text{new}}\\ \hat{\mathbf{y}}_{\text{new}}&=\hat{\mathbf{f}}_{y,\text{new}}+\mathbf{C}_y\mathbf{t}_{y,\text{new}}=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{f}_{x,\text{new}}+\mathbf{C}_y\mathbf{P}_y^T\mathbf{u}_{\text{new}}\\ &=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\left(\mathbf{x}_{\text{new}}-\mathbf{C}_x\mathbf{P}_x^T\mathbf{u}_{\text{new}}\right)+\mathbf{C}_y\mathbf{P}_y^T\mathbf{u}_{\text{new}}\\ &=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{x}_{\text{new}}+\left(\mathbf{C}_y\mathbf{P}_y^T-\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{C}_x\mathbf{P}_x^T\right)\mathbf{u}_{\text{new}}\end{aligned} \tag{14}$$

As can be seen from Eq. 14, the estimated quality variable is actually a weighted combination of the main variables and the external variables, with their corresponding weighted values as $\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T$ and $\mathbf{C}_y\mathbf{P}_y^T-\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{C}_x\mathbf{P}_x^T$.

Moving-window-based EAR model for soft sensing of new operation modes

Up to now, the EAR model-based soft sensor can only address the soft sensing problem for the process with multiple known operation modes. If a new process change happens or the process changes to a new operation mode, the performance of the EAR model-based soft sensor may be degraded. In this part, a moving-window is introduced to the EAR model, which can efficiently catch the real-time information of the process. Therefore, an online EAR modeling scheme is developed.

Given a new moving window with its size as w, which has incorporated the new data information, the real-time moving-window-based data sample can be represented as $\mathbf{Z}^w=[\mathbf{U}^w\ \ \mathbf{X}^w\ \ \mathbf{Y}^w]$. First, the relation between the external variables $\mathbf{U}^w$ and the main variables $\mathbf{X}^w$ should be built, which is given as follows

$$\begin{aligned}\mathbf{U}_x^w&=\mathbf{T}_x^w\mathbf{P}_x^{wT}+\mathbf{E}_x^w\\ \mathbf{X}_x^w&=\mathbf{T}_x^w\mathbf{C}_x^{wT}+\mathbf{F}_x^w\end{aligned} \tag{15}$$

Based on this real-time $\mathrm{PCR}^{\text{x-u}}$ model, the unexplained information of the main variables is stored in the matrix $\mathbf{F}_x^w$. Second, the relation between the external variables $\mathbf{U}^w$ and the quality variables $\mathbf{Y}^w$ can be calculated as

$$\begin{aligned}\mathbf{U}_y^w&=\mathbf{T}_y^w\mathbf{P}_y^{wT}+\mathbf{E}_y^w\\ \mathbf{X}_y^w&=\mathbf{T}_y^w\mathbf{C}_y^{wT}+\mathbf{F}_y^w\end{aligned} \tag{16}$$

In this real-time $\mathrm{PCR}^{\text{y-u}}$ model, the unexplained information of the quality variables is stored in the matrix $\mathbf{F}_y^w$. Different from the EAR model in section EAR model, the relation between the unexplained data information matrices $\mathbf{F}_x^w$ and $\mathbf{F}_y^w$ does not need to be modeled. This is because after the impacts of the external variables have been removed from both main variable and quality variable datasets $\mathbf{X}^w$ and $\mathbf{Y}^w$, the relation between them should remain the same as that in the previous operation modes.

Therefore, given a new data sample $\mathbf{r}_{\text{new}}=[\mathbf{u}_{\text{new}}\ \ \mathbf{x}_{\text{new}}]$, the predicted quality variable can be calculated as follows

$$\mathbf{f}_{x,\text{new}}^w=\mathbf{x}_{\text{new}}-\mathbf{C}_x^w\mathbf{t}_{x,\text{new}}^w=\mathbf{x}_{\text{new}}-\mathbf{C}_x^w\mathbf{P}_x^{wT}\mathbf{u}_{\text{new}} \tag{17}$$

$$\begin{aligned}\mathbf{t}_{x-y,\text{new}}^w&=\mathbf{P}_{x-y}^T\mathbf{f}_{x,\text{new}}^w\\ \hat{\mathbf{f}}_{y,\text{new}}^w&=\mathbf{C}_{x-y}\mathbf{t}_{x-y,\text{new}}^w=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{f}_{x,\text{new}}^w\\ &=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{x}_{\text{new}}-\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{C}_x^w\mathbf{P}_x^{wT}\mathbf{u}_{\text{new}}\end{aligned} \tag{18}$$

$$\begin{aligned}\hat{\mathbf{y}}_{\text{new}}&=\hat{\mathbf{f}}_{y,\text{new}}^w+\mathbf{C}_y^w\mathbf{t}_{y,\text{new}}^w=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{f}_{x,\text{new}}^w+\mathbf{C}_y^w\mathbf{P}_y^{wT}\mathbf{u}_{\text{new}}\\ &=\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{x}_{\text{new}}+\left(\mathbf{C}_y^w\mathbf{P}_y^{wT}-\mathbf{C}_{x-y}\mathbf{P}_{x-y}^T\mathbf{C}_x^w\mathbf{P}_x^{wT}\right)\mathbf{u}_{\text{new}}\end{aligned} \tag{19}$$

Performance monitoring and analysis based on SVDD

To monitor the operation condition of the process, and also to distinguish normal process changes and the faulty condition, a performance monitoring model is developed in this subsection, which is based on the SVDD model. As an efficient one-class classification method, SVDD has gained much attention in recent years, and many applications have already been carried out in different areas, such as pattern recognition, image processing, fault detection, and so forth. The main idea of the SVDD method is to map the input vectors to a feature space, and to find hypersphere with the minimum volume which separates the transferred data from the rest of the feature space. Applications have shown a high generalization performance of SVDD if large reference dataset with few abnormal samples is available.

In the present work, the SVDD model is used to monitor the estimation performance of the EAR model-based soft sensor. Hence, before the SVDD modeling step, the prediction residual the soft sensor for the training dataset should be calculated, which is given as

$$\mathbf{Re}_{tr}=\mathbf{Y}_{tr}-\hat{\mathbf{Y}}_{tr} \tag{20}$$

where $\mathbf{Y}_{tr}$ and $\hat{\mathbf{Y}}_{tr}$ are measurement values and predicted values of the quality variables. Then, SVDD maps each data vector of the residual dataset $\mathbf{Re}_{tr}$ from the original space to the feature space $\Phi : \mathbf{re} \rightarrow F$. To improve the accuracy and generalization of the SVDD model, a nonlinear function is often employed for data transformation. In this case, a kernel function $K(\mathbf{re}_i, \mathbf{re}_j)=\langle\Phi(\mathbf{re}_i), \Phi(\mathbf{re}_j)\rangle$ can be employed to compute the inner product of the data in the feature space. In the present article, the most popular Gaussian kernel is used. Then, a hypersphere can be built in the feature space, the center $\mathbf{a}$ and the radius R of which can be determined as[33]

$$\begin{aligned} \mathbf{a}&=\sum_{i=1}^{n_{tr}}\alpha_i\Phi(\mathbf{re}_i) \\ R&=\sqrt{1-2\sum_{i=1}^{n_{tr}}\alpha_i K(\mathbf{re}_0, \mathbf{re}_i)+\sum_{i=1}^{n_{tr}}\sum_{j=1}^{n_{tr}}\alpha_i\alpha_j K(\mathbf{re}_i, \mathbf{re}_j)} \end{aligned} \tag{21}$$

where $\mathbf{re}_0$ is referred to a support vector of SVDD, with the parameter $0 < \alpha_0 < C$. Then, a performance monitoring statistic based on the distance between the data sample and the center of the SVDD hypersphere can be developed as

$$\mathrm{Dist}_i=d(\Phi(\mathbf{re}_i))=\|\Phi(\mathbf{re}_i)-\mathbf{a}\| \leq \mathrm{Dist}_{\mathrm{lim}}=R \tag{22}$$

Therefore, when a new data sample has been estimated by the REA-based soft sensor, the residual value between the measurement value and the estimated value should be calculated first, based on which the estimation performance and the process condition can be monitored and analyzed by the SVDD model. Following the previous part, if the predicted value of the new data sample is calculated as $\hat{\mathbf{y}}_{\mathrm{new}}$, the estimation residual can be generated as

$$\mathbf{re}_{\mathrm{new}}=\mathbf{y}_{\mathrm{new}}-\hat{\mathbf{y}}_{\mathrm{new}} \tag{23}$$

where $\mathbf{y}_{\mathrm{new}}$ is the measurement value of the new data sample. Then, the value of the performance monitoring statistic corresponding to this new data sample can be calculated as

$$\mathrm{Dist}_{\mathrm{new}}=d(\Phi(\mathbf{re}_{\mathrm{new}}))=\|\Phi(\mathbf{re}_{\mathrm{new}})-\mathbf{a}\| \tag{24}$$

Based on the performance monitoring statistic, if $\mathrm{Dist}_{\mathrm{new}} > R$, the estimation performance of the soft sensor is considered to be degraded. Among the EAR model-based estimation framework, a faulty condition should be judged if the value of $\mathrm{Dist}_{\mathrm{new}}$ has exceeds its confidence limit. However, if the process changes to some new but normal operation condition, the $\mathrm{Dist}_{\mathrm{new}}$ statistic will not exceed its confidence limit, this is because the EAR model-based soft sensor can well track the new operation mode. In other words, the EAR model-based soft sensor can efficiently distinguish the normal process change with the faulty operation condition.

Implementation procedures of the EAR model-based soft sensor

In summary, the modeling and online quality estimation of the EAR model-based soft sensor can be formulated as follows. A corresponding flowchart of this method is illustrated in Figure 3.

Modeling Phase. Step 1: Data collection from normal operation conditions;

Step 2: Classify the process variable into three categories: external variable set, main variable set, and quality variable set;

Step 3: PCR model ($\mathrm{PCR}^{\text{x-u}}$) development between external variables and main variables, thus the explained data information by the external variable in the main variable dataset is removed;

Step 4: Develop another PCR model ($\mathrm{PCR}^{\text{y-u}}$) to remove the explained data information by the external variable in the quality variable dataset;

Step 5: Based on the unexplained data information of the main variable and the quality variable datasets, an additional PCR model ($\mathrm{PCR}^{\text{y-x}}$) is built;

Step 6: For performance monitoring and analysis, an SVDD model is developed.

Online Prediction Phase. Step 1: Data acquirement by using the moving-window approach;

Step 2: Separate the process variables into external variables, main variables, and quality variables among the moving window;

Step 3: Construct a moving-window-based $\mathrm{PCR}^{\text{x-u}}$ (MW-$\mathrm{PCR}^{\text{x-u}}$) model between the external variables and the main variables, generate the unexplained main data information;

Step 4: Construct a moving-window-based $\mathrm{PCR}^{\text{y-u}}$ (MW-$\mathrm{PCR}^{\text{y-u}}$) model between the external variables and the quality variables and generate the unexplained quality data information;

Step 5: For the new data sample, employ the MW-$\mathrm{PCR}^{\text{x-u}}$ model to calculate the unexplained data information of the main variables;

Step 6: Calculate the estimated explained data information of the quality variables based on the new external variables and the MW-$\mathrm{PCR}^{\text{y-u}}$ model;

Step 7: Calculate the estimated unexplained data information of the quality variables based on the unexplained data information of the main variables and the $\mathrm{PCR}^{\text{y-x}}$ model;

Step 8: Generate the final estimated quality variable by adding the predicted explained data information and the estimated unexplained data information together;

Step 9: Prediction performance analysis and process monitoring by the SVDD-based model.

Illustrations and Results

In this section, the performance of the proposed soft sensor is evaluated through a simulation example and the Tennessee Eastman (TM) benchmark process.

Simulation Example

Consider the following system which contains five process variables and two output variables, it is revised from the Ref. 31

$$\begin{aligned} \mathbf{x}(i)&=\mathbf{s}(i)\mathbf{A}+\mathbf{v}(i) \\ \mathbf{A}&=\begin{bmatrix} 1 & 0.894 & 0.058 & 0.353 & 0.813 \\ 0 & 0.445 & 0.932 & 0.466 & 0.419 \\ 0 & 0.846 & 0.525 & 0.203 & 0.672 \end{bmatrix} \\ \mathbf{s}&=[s_1\ \ s_2\ \ s_3] \end{aligned} \tag{25}$$

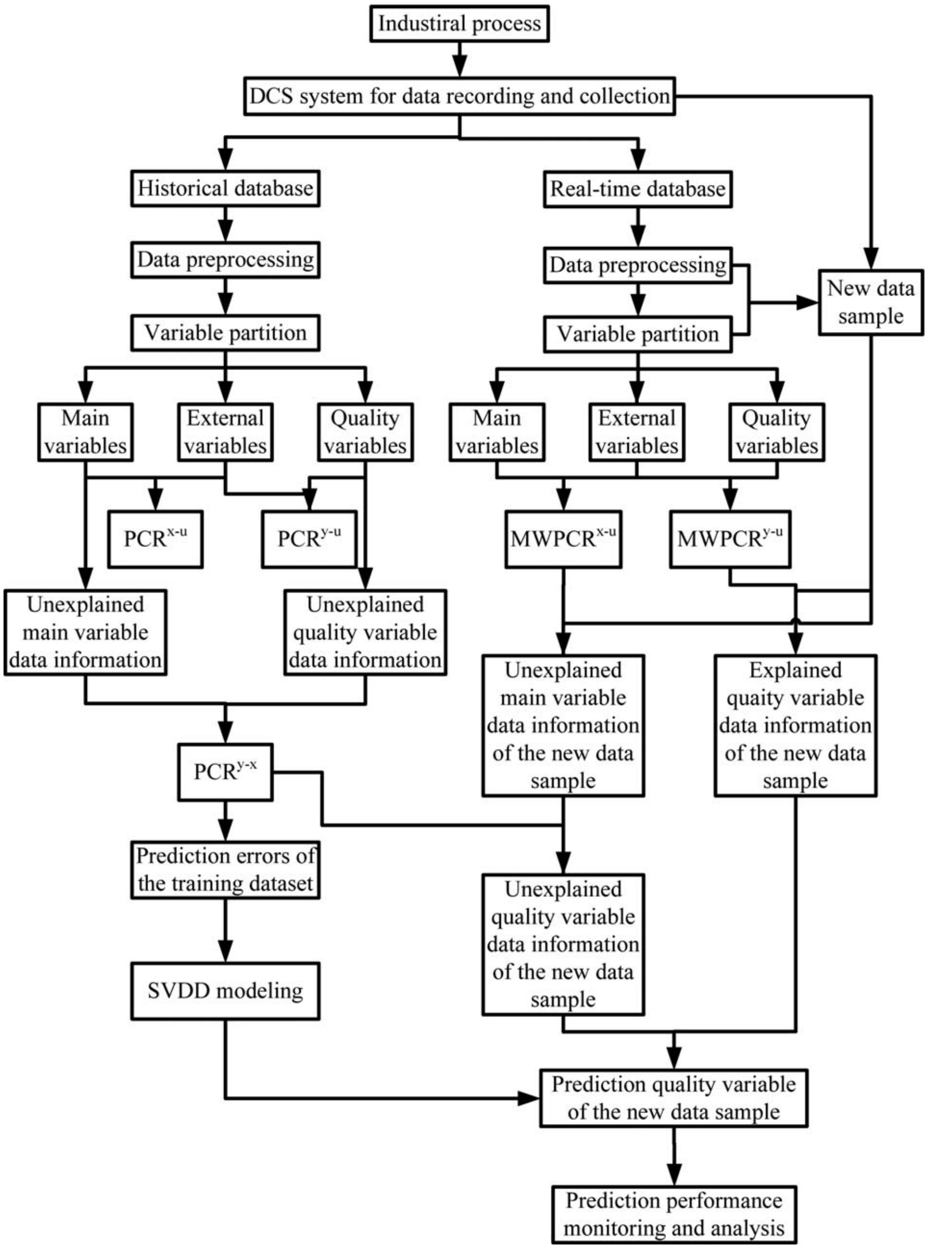

Figure 3. Flowchart of the EAR-based soft sensor.

$$\mathbf{y}(i)=\mathbf{x}(i)\mathbf{B}+\mathbf{w}(i)$$
$$\mathbf{B}=\begin{bmatrix} 0.110 & 1.909 \\ 0.307 & 0.879 \\ 1.159 & -1.094 \\ 0.703 & -0.139 \\ 0.374 & 0.296 \end{bmatrix} \tag{26}$$

where $i=1,2,\ldots,n$, s_1, s_2, and s_3 are uncorrelated random signals following uniform distribution within [0 1], $\mathbf{v}$, and $\mathbf{w}$ are noise vectors of each measured process variable $\mathbf{x}$ and output variable $\mathbf{y}$, respectively, both of which follow normal distributions with deviation of 0.05. When the dataset for the normal process condition has been generated, the additive signal r on s_1 is changed stepwise every 200 steps to simulate the normal operation changes of the process. The value of r is randomly determined between 0 and 10. Meanwhile, random changes of the regression matrix $\mathbf{B}$ have been used to simulate different model structure of each operation mode. As can be inferred from Eq. 25, the changes in r only affect the variable x_1, no impact has been put on other process variables. Therefore, according to the principal of the external analysis method, x_1 should be determined as the external variable, and the other four process variables are determined as the main variables.

In this example, except for the basic operation condition which is described in Eqs. 25 and 26, four different normal operation modes under different r values are also simulated, with each dataset consisting of 200 samples. For modeling training, another normal dataset has been generated under the basic operation condition. The number of the principal components in the PCR model is selected as 1 in PCR$^{x\text{-}u}$

and PCR$^{y\text{-}u}$ models, and 2 in the PCR$^{y\text{-}x}$ model. For comparison, the traditional PCR-based soft sensor is also developed. The data characteristics of both input and output data are presented in Figure 4, in which we can clearly find that different data behaviors have been exhibited in different operation conditions. Initially, the size of the moving window is selected as 30. Because the size of the moving window is important to the EAR model based soft sensor, the estimation performance of the soft sensor under different sizes of the moving window is studied.

First, the result of the EAR model-based soft sensor under the window size 30 is generated, which is shown in Figure 5a. As can be seen, all of the data samples under four new operation modes have been well-estimated, thus the operation condition of the process has been correctly tracked. It is noticed that the estimation performance of several data samples when the new operation mode starts has degraded, which have been highlighted by ellipses in Figure 5a. This is due to the moving window method, because the updating procedure will incorporate data information in two adjacent operation modes. In contrast, the traditional PCR-based soft sensor cannot give accurate estimation results when the operation condition of the process has been changed, the result of which is given in Figure 5b. This is because the PCR model has only incorporated the data information of the basic operation mode. When the process changes to new operation modes, the PCR-based soft sensor has no experience to support this new data information, which renders a very poor prediction performance. Similarly, the model library-based soft sensor cannot provide satisfactory results neither, because the information of the new operation modes has also excluded in the training dataset. Although the adaptive method-based soft sensor may give a good estimation result, it will probably accommodate to the faulty operation condition.

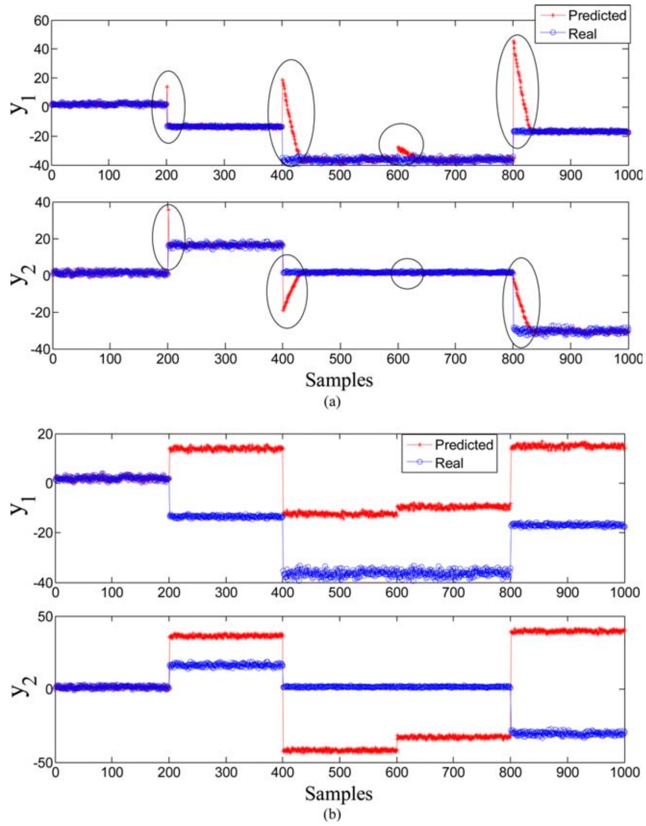

Figure 5. Quality estimation results of datasets under multiple operation conditions, (a) EAR-PCR; (b) PCR.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

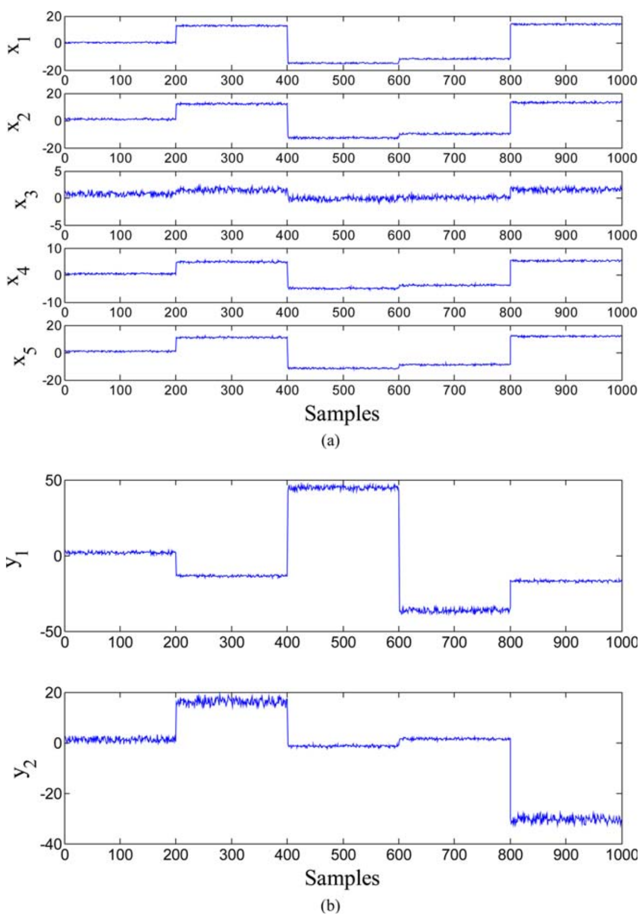

Figure 4. Data characteristics under different operation conditions, (a) input data; (b) output data.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Next, a faulty operation condition is introduced to the process, in which we have changed the correlation matrix between the unexplained main variable data information and the unexplained quality variable data information. Therefore, the relationship between the input and out variables has been changed. Precisely, two elements ([1, 1] and [2, 2]) of the correlation matrix have been changed, which are shown as follows

$$\boldsymbol{\Theta}=\mathbf{P}_{x-y}\mathbf{C}_{x-y}^{T}=\begin{bmatrix}0.330 & 0.711\\ 1.215 & -0.919\\ 0.605 & -0.535\\ 0.324 & 0.507\end{bmatrix}$$

$$\xrightarrow{\text{correlation changes to}}\boldsymbol{\Theta}'=\begin{bmatrix}5.330 & 0.711\\ 1.215 & -5.919\\ 0.605 & -0.535\\ 0.324 & 0.507\end{bmatrix}$$

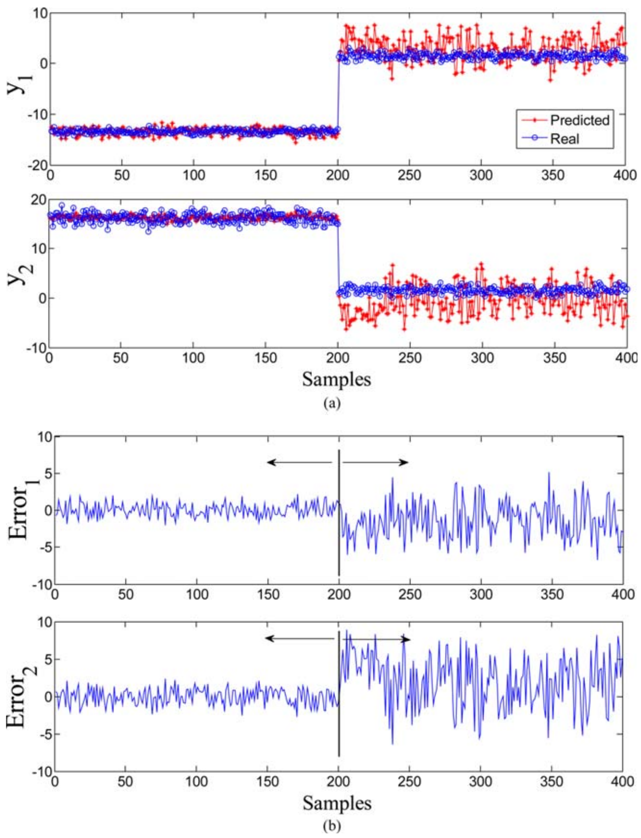

Figure 6. Estimation results of the EAR model-based soft sensor when the process changes to a faulty operation condition, (a) quality estimation results; and (b) prediction errors.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

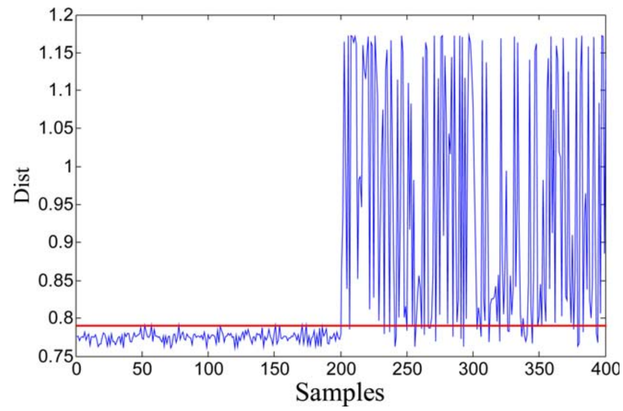

Figure 7. SVDD monitoring results of the faulty operation condition.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The estimation results of the EAR model based soft sensors is given in Figure 6a, where the first 200 data samples are normal and the last 200 ones are abnormal. Clearly, the estimation performance of the last 200 data sample has been degraded, which can also be seen from the prediction errors given in Figure 6b. Besides, if we examine the monitoring results of this dataset, the monitoring statistic values of the abnormal data samples will significantly exceeds its confidence limit. The monitoring results based on the SVDD method are given in Figure 7.

Finally, the influence of the moving window size on the estimation performance is studied. Generally speaking, the method will track the process change better if the size of the moving window is selected as a small value, because the new data information is updated more quickly with a small moving window. On the other hand, if a big moving window has been selected, the tracking speed is much slower, and more estimation errors may be generated between the changes of different operation modes. Therefore, the mean estimation accuracy of the soft sensor may be deteriorated. However, if the number of the process variables is large, the size of the moving window should be selected as an appropriate value. If a small value is selected, the modeling performance of the PCR may not be guaranteed. Table 1 lists the root-mean-squared error (RMSE) values of different selections of the moving window size for the 5 normal operation datasets. The index of RMSE is defined as

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{j=1}^{L}\|\mathbf{y}_j-\hat{\mathbf{y}}_j\|^2}{L}} \tag{27}$$

where $j=1,2,\ldots,L$, $\mathbf{y}_j$ and $\hat{\mathbf{y}}_j$ are real and estimation values, respectively, L is the total number of test data samples. In this table, we can find that with the increase of the moving window size, the mean prediction performance of the soft sensor is decreased. In our opinion, this is mainly due to the reason that more prediction errors have been generated based on the large moving window.

TE Benchmark Process

In last several decades, TE process has been widely used for algorithm testing and performance evaluation,[34–39] which consists of five major unit operations: a reactor, a condenser, a compressor, a separator, and a stripper. In this process, two simultaneous gas-liquid exothermic reactions are generated to produce two different product, which are given as

$$\begin{aligned} A(g)+C(g)+D(g) &\rightarrow G(\mathrm{liq}) \\ A(g)+C(g)+E(g) &\rightarrow H(\mathrm{liq}) \end{aligned} \tag{28}$$

There are 41 measured variables and 12 manipulated variables. The flowchart of TE process is shown schematically in Figure 8. Among all of the 41 process variables, there are 19 component variables, which are difficult to be measured online. In contrast, the rest 22 process variables are very easy to be measured. In this case study, all of the 22 process variables are selected for soft sensor modeling, and five

Table 1. RMSE Values of the EAR Soft Sensor under Different Moving Window Sizes

Size	10	20	30	40	50	60	70	80	90	100
RMSE_y_1	5.8339	7.6556	9.2292	10.4988	11.6381	12.6401	13.6584	14.5571	15.4155	16.1703
RMSE_y_2	2.7038	3.1986	3.7108	4.1604	4.5701	4.8549	5.2241	5.6010	5.8672	6.1605

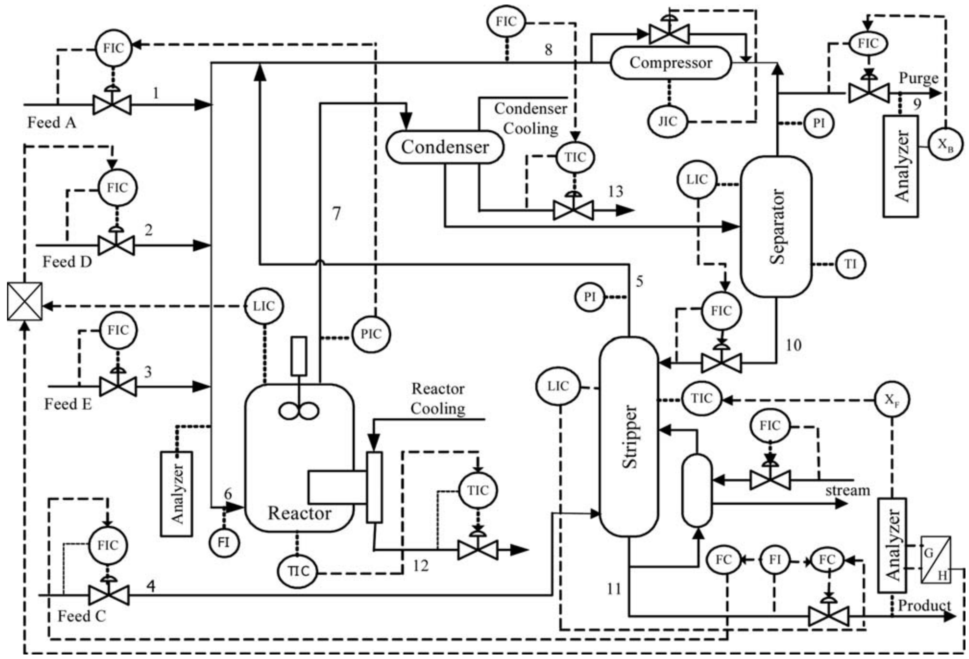

Figure 8. Tennessee Eastman process.

components in the final quality stream are used as the quality variables. In the TE benchmark process, six different operation modes can be simulated, descriptions of which are listed in Table 2.

Among the 22 measured process variables, the first four variables are feeds of the process, which come out of the system. Based on the two reactions given in Eq. 28, they are highly related with the final components of the product, which means they can be considered as driven factors of the operation mode. Therefore, the first four measured variables are considered as external variables, and the rest 18 variables are considered as main variables. Detailed descriptions of the external variables, main variables, and quality variables are tabulated together in Table 3. For soft sensor modeling construction and testing, datasets under different operation modes are generated, each of which consists of 200 data samples. Previously, the normal dataset is partitioned into two equal parts: training dataset and testing dataset. Based on the CPV criterion which is selected as 85% in the present work, the retained numbers of principal components in $PCR^{x\text{-}u}$ and $PCR^{y\text{-}u}$ models are both determined as 3, while the value is 8 in the $PCR^{y\text{-}x}$ model. Similarly, the traditional PCR model is also developed for performance comparison.

The data characteristics of the input and output variables are given in Figure 9. Again, we can clearly find six different data behaviors in both Figure 9a, b, which are corresponding to the six different operation modes in this process. To test the performance of the proposed soft sensor for new operation modes, only the training dataset of the first operation mode is used for model constructions of both soft sensors. The estimation results of both two soft sensors for the combined dataset of all six operation modes are given in Figure 10, where the red line corresponds to the estimated values and the blue line corresponds to the measured value of each data sample. It can be seen from Figure 10a that all data samples in different operation modes (new operation mode or not) have been predicted very well, except for several data samples in adjacencies of two operation modes. Again, this is due to the moving-window approach which has an updating nature to accommodate the new operation mode. However, the estimation performance of the PCR-based soft sensor has been greatly deteriorated when the process changes to new operation modes. Except for the basic operation depending on which the PCR-based soft sensor is developed, the estimation errors of data samples in other five operation modes are very significant. In Figure 11, estimation errors of the last two quality variables by the two soft sensors are presented. It can be seen in Figure 11a that five operation change behaviors have been found, which meets the real situation of the process condition. In contrast, the estimation errors of the PCR-based soft sensor for the five new operation modes are very large, which means it has no ability for quality estimation in new operation conditions.

Through analyzing the SVDD-based monitoring statistic, we can also judge the operation condition of the process. To build the SVDD model, the generally used Gaussian kernel is employed, with its width parameter selected as 3.2, and the confidence limit of the SVDD monitoring statistic is determined as 95%. As a result, five data samples have been selected as support vectors in the SVDD model. The

Table 2. Descriptions of Different Operation Modes in TE Process

Mode	1	2	3	4	5	6
G/H mass ratio	50/50	10/90	90/10	50/50	10/90	90/10
Production rate in stream 11	7038 kg h^{-1} G and 7038 kg h^{-1} H	1408 kg h^{-1} G and 12,669 kg h^{-1} H	100,00 kg h^{-1} G and 1111 kg h^{-1} H	Maximum production rate	Maximum production rate	Maximum production rate

Table 3. Variable Descriptions of the TE Process

External Variables		Main Variables				Quality Variables	
1	A feed	1	Recycle flow	10	Product separator underflow	1	Component D
2	A feed	2	Reactor feed rate	11	Stripper level	2	Component E
3	E feed	3	Reactor pressure	12	Stripper pressure	3	Component F
4	A and C feed	4	Reactor level	13	Stripper underflow	4	Component G
		5	Reactor temperature	14	Stripper temperature	5	Component H
		6	Purge rate	15	Stripper steam flow		
		7	Product separator temperature	16	Compressor work		
		8	Product separator level	17	Reactor cooling water outlet temperature		
		9	Product separator pressure	18	Separator cooling water outlet temperature		

Lagrange multiplier corresponding to each training data point is illustrated in Figure 12, through which it can be found that only five data samples have significant values, all of other multiplier values are zeros which have no impact to the monitoring statistic. Therefore, we can just use these five support vectors to calculate the monitoring statistic value for the new data sample, which will greatly reduce the computationally complexity of online analysis. The results of the SVDD monitoring statistic for different operation modes are shown in Figure 13, in which six operation modes have also been clearly presented.

To test the estimation performance of proposed soft sensor under the faulty operation condition, one fault is introduced into the process, which is a random variation of the condenser cooling water inlet temperature. As a result, both of the two soft sensors cannot provide satisfactory estimation performance. Actually, by analyzing the results of SVDD monitoring statistic, we can find that the operation condition of the process has been moved outside of the normal region, which is demonstrated in Figure 14. Therefore, under the faulty operation condition, the EAR model-based soft sensor can successfully distinguish it from normal process changes. Based on the SVDD monitoring strategy, the process fault can also be detected and thus further analysis can be carried out.

To examine the estimation performance of the soft sensor under different sizes of the moving window, the value is changed from 30 to 70. Quality estimation results under different cases are tabulated in Table 4. Besides, the mean RMSE value of the five quality variables in each case is also

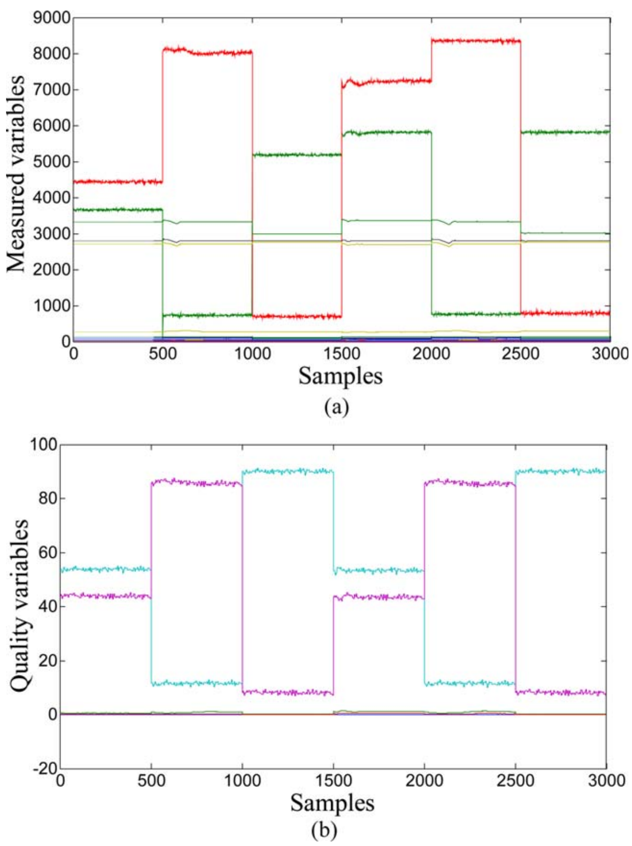

Figure 9. Data characteristics of measured and quality variables, (a) measured variables; and (b) quality variables.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

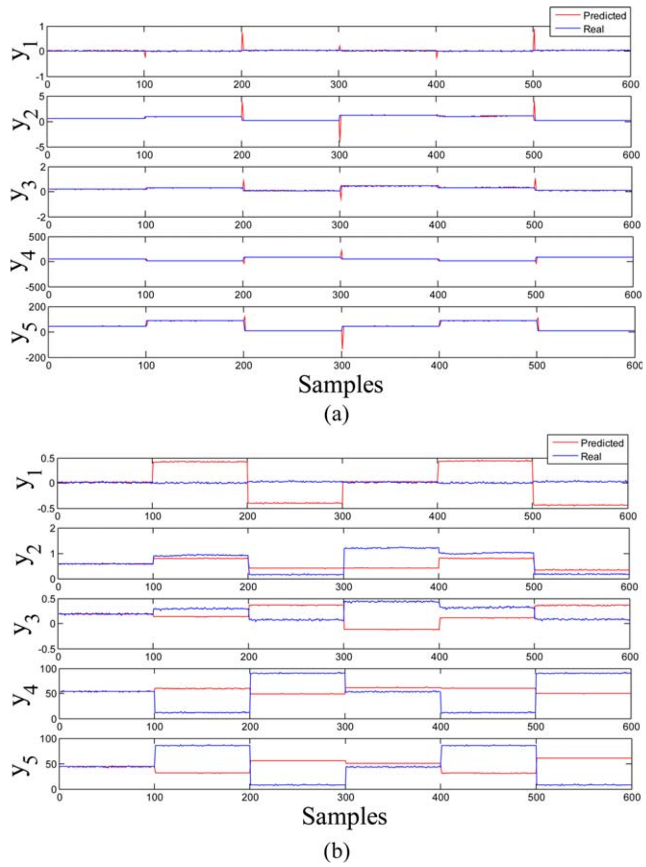

Figure 10. Quality estimation results of datasets under multiple operation conditions, (a) EAR-PCR; and (b) PCR.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

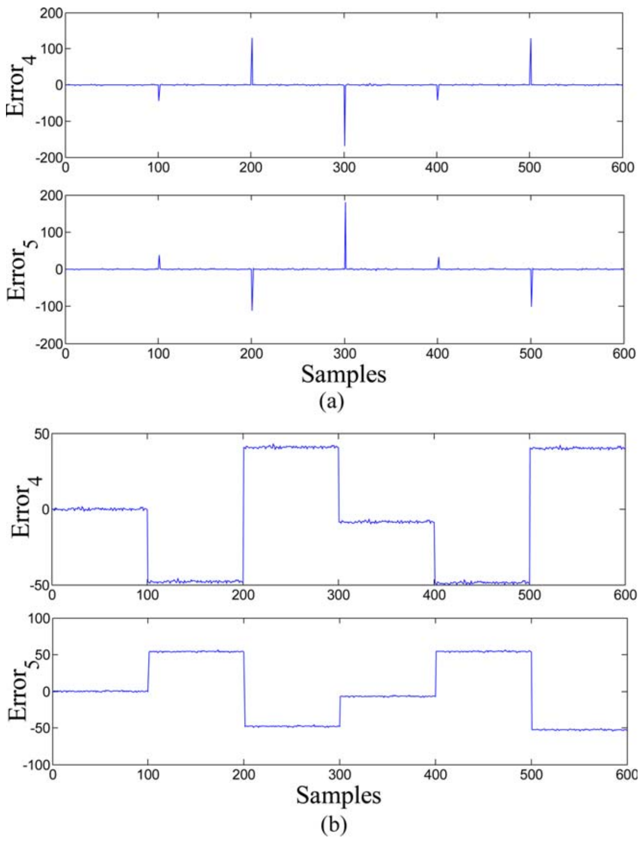

Figure 11. Quality estimation error of the last two quality variables under multiple operation conditions, (a) EAR-PCR; and (b) PCR.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

provided in the last row of the table. Based on this mean RMSE result, it can be conclude that the high estimation performance of the EAR model-based soft sensor will be obtained when a small size of the moving window has been selected. However, as we have mentioned, if the size of the moving window is too small, the modeling performance of the soft senor cannot be guaranteed, which will influence the final estimation performance. Therefore, in practice, the size of the moving window should be carefully chosen. Although this is still an open question, to our experience, it somehow

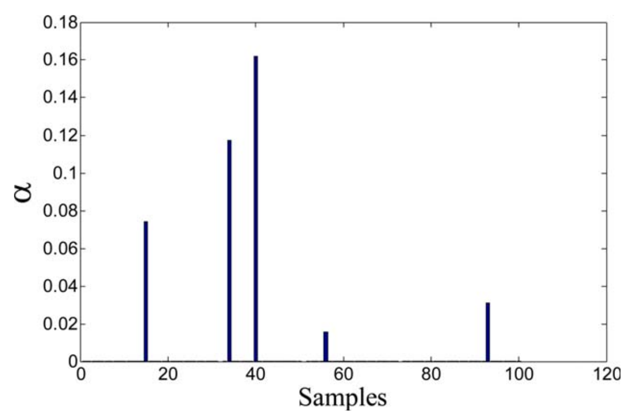

Figure 12. SVDD modeling coefficients of different data samples.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

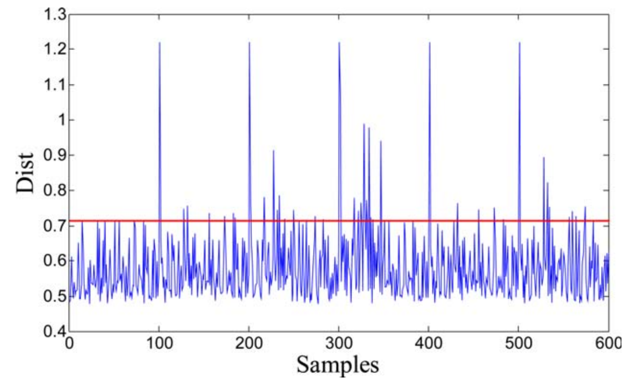

Figure 13. Results of the SVDD-based monitoring statistic for six normal operation modes.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

depends on the type of the modeling method. For example, if a modeling method which needs a large number of training data samples such as the density estimation-based method, the size of the moving window should also be selected as a big value, otherwise the built model may be useless. On the other hand, if some limited data-based modeling method has been employed, such as the support vector machine-based modeling approach, a relatively small size of the moving window can be selected, as long as the modeling performance of the soft sensor could not be influenced.

Conclusions

In the present article, an EAR model is proposed for adaptive soft sensor modeling of multimode chemical processes. Different from existing soft sensor modeling methods, the EAR-model based soft sensor can successfully handle new process changes, and also has ability to distinguish the normal operation mode and the process fault. To extend the new method for online quality estimation, a moving window approach has been incorporated into the EAR model. Furthermore, an SVDD-based model has been developed for estimation performance monitoring and analysis, depending on which the process fault can be successfully detected. Two case studies have both demonstrated the feasibility and efficiency of the proposed method. Based on the EAR modeling framework, both of the nonlinear and dynamic counterparts of the developed method can be obtained easily. Therefore,

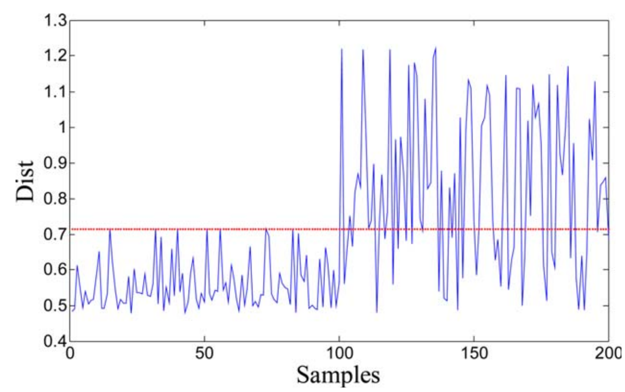

Figure 14. Results of the SVDD-based monitoring statistic for the faulty operation mode.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 4. RMSE Values of the EAR Soft Sensor under Different Moving Window Sizes

Size	30	40	50	60	70
RMSE_y_1	0.0267	0.0267	0.0138	0.0140	0.0162
RMSE_y_2	0.1755	0.1933	0.2141	0.1878	0.1978
RMSE_y_3	0.0196	0.0352	0.0278	0.0217	0.0277
RMSE_y_4	6.4813	7.6377	7.3478	7.5456	6.8109
RMSE_y_5	7.6878	7.4429	7.7188	7.5489	7.7600
Mean	2.8782	3.0672	3.0645	3.0636	2.9625

although only the linear PCR model has been used for soft sensor in the present article, the EAR modeling method is actually a general modeling approach for multimode processes. One can develop different types of robust soft sensors for corresponding multimode cases.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (NSFC) (61370029), Project National 973 (2012CB720500), and the Fundamental Research Funds for the Central Universities (2013QNA5016).

Literature Cited

1. Joseph B, Brosilow CB. Inferential control of processes-1: steady state analysis and design. *AIChE J*. 1978;124:485–508.
2. Fortuna L, Graziani S, Rizzo A, Xibilia MG. *Soft Sensors for Monitoring and Control of Industrial Processes*. London: Springer, 2007.
3. Kano M, Nakagawa Y. Data-based process monitoring, process control and quality improvement: recent developments and applications in steel industry. *Comput Chem Eng*. 2008;32:12–24.
4. Kadlec P, Gabrys B, Strandt S. Data-driven soft sensors in the process industry. *Comput Chem Eng*. 2009;33:795–814.
5. Kruger U, Chen Q, Sandoz DJ, McFarlane RC. Extended PLS approach for enhanced condition monitoring of industrial processes. *AIChE J*. 2001;47:2076–2091.
6. Wang X, Kruger U, Lennox B. Recursive partial least squares algorithms for monitoring complex industrial processes. *Control Eng Pract*. 2003;11:613–632.
7. Kruger U, Antory D, Hahn J, Irwin GW, McCullough G. Introduction of a nonlinearity measure for principal component models. *Comput Chem Eng*. 2005;29:2355–2362.
8. Ge ZQ, Gao FR, Song ZH. Mixture probabilistic PCR model for soft sensing of multimode processes. *Chemometr Intell Lab Syst*. 2011; 105:91–105.
9. Yu J. Multiway Gaussian mixture model based adaptive kernel partial least squares regression method for soft sensor estimation and reliable quality prediction of nonlinear multiphase batch processes. *Ind Eng Chem Res*. 2012;51:13227–13237.
10. Galicia HJ, He QP, Wang J. Comparison of the performance of a reduced-order dynamic PLS soft sensor with different updating schemes for digester control. *Control Eng Pract*. 2012;20:747–760.
11. Gonzaga JCB, Meleiro LAC, Kiang C, Filho RM. ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process. *Comput Chem Eng*. 2009;33:43–49.
12. Yan XF. Hybrid artificial neural network based on BP-PLSR and its application in development of soft sensors. *Chemometr Intell Lab Syst*. 2010;103:152–159.
13. Bhattacharya S, Pal K, Pal SK. Multi-sensor based prediction of metal deposition in pulsed gas metal arc welding using various soft computing models. *Appl Soft Comput*. 2012;12:498–505.
14. De Canete JF, del Saz-Orozco P, Gonzalez S, Garcia-Moral I. Dual composition control and soft estimation for a pilot distillation column using a neurogenetic design. *Comput Chem Eng*. 2012;40:157–170.
15. Liu Y, Hu NP, Wang HQ, Li P. Soft chemical analyzer development using adaptive least-squares support vector regression with selective pruning and variable moving window size. *Ind Eng Chem Res*. 2009; 48:5731–5741.
16. Ge ZQ, Song ZH. Nonlinear soft sensor development based on relevance vector machine. *Ind Eng Chem Res*. 2010;49:8685–8693.
17. Jin X, Wang SY, Huang B, Forbes F. Multiple model based LPV soft sensor development with irregular/missing process output measurement. *Control Eng Pract*. 2012;20:165–172.
18. Yu J. A Bayesian inference based two-stage support vector regression framework for soft sensor development in batch bioprocesses. *Comput Chem Eng*. 2012;41:134–144.
19. Deng J, Xie L, Chen L, Khatibisepehr S, Huang B, Xu F, Espejo A. Development and industrial application of soft sensor with on-line Bayesian model updating strategy. *J Process Control*. 2012;23:317–325.
20. Kano M, Ogawa M. The state of the art in chemical process control in Japan: Good practice and questionnaire survey. *J Process Control*. 2010;20:969–982.
21. Qin SJ. Recursive PLS algorithms for adaptive data monitoring. *Comput Chem Eng*. 1998;22:503–514.
22. Wang X, Kruger U, Irwin GW. Process monitoring approach using fast moving window PCA. *Ind Eng Chem Res*. 2005;44:5691–5702.
23. Jin HD, Lee YH, Lee G, Han CH. Robust recursive principal component analysis modeling for adaptive monitoring. *Ind Eng Chem Res*. 2006;45:696–703.
24. Hwang DH, Han C. Real-time monitoring for a process with multiple operating modes. *Control Eng Pract*. 1999;7:891–902.
25. Singhai A, Seborg DE. Evaluation of a pattern matching method for the Tennessee Eastman challenge process. *J Process Control*. 2006; 16:601–613.
26. Liu JL. On-line soft sensor for polyethylene process with multiple production grades. *Control Eng Pract*. 2007;15:769–778.
27. Ge ZQ, Song ZH. Online monitoring of nonlinear multiple mode processes based on adaptive local model approach. *Control Eng Pract*. 2008;16:1427–1437.
28. Ge ZQ, Song ZH. A comparative study of Just-In-Time-Learning based methods for online soft sensor modeling. *Chemometr Intell Lab Syst*. 2010;104:306–317.
29. Fujiwara K, Kano M, Hasebe S. Development of correlation-based clustering method and its application to software sensing. *Chemometr Intell Lab Syst*. 2010;101:130–138.
30. Kadlec P, Gabrys B. Local learning-based adaptive soft sensor for catalyst activation prediction. *AIChE J*. 2011;57:1288–1301.
31. Kano M, Hasebe S, Hashimoto I, Ohno H. Evolution of multivariate statistical process control: application of independent component analysis and external analysis. *Comput Chem Eng*. 2004;28:1157–1166.
32. Ge ZQ, Yang CJ, Song ZH, Wang HQ. Robust online monitoring for multimode processes based on nonlinear external analysis. *Ind Eng Chem Res*. 2008;47:4775–4783.
33. Tax DMJ, Duin RPW. Support vector domain description. *Mach Learn*. 2004;54:45–66.
34. Downs JJ, Vogel EF. A plant-wide industrial process control problem. *Comput Chem Eng*. 1993;17:245–255.
35. Maurya MR, Rengaswamy R, Venkatasubramanian V. Fault diagnosis by qualitative trend analysis of the principal components. *Chem Eng Res Des*. 2005;83:1122–1132.
36. Ge ZQ, Yang CJ, Song ZH. Improved kernel PCA-based monitoring approach for nonlinear processes. *Chem Eng Sci*. 2009;64:2245–2255.
37. Ge ZQ, Xie L, Kruger U, Lamont L, Song ZH, Wang SQ. Sensor fault identification and isolation for multivariate non-Gaussian processes. *J Process Control*. 2009;19:1707–1715.
38. Zhou DH, Li G, Qin SJ. Total projection to latent structures for process monitoring. *AIChE J*. 2010;56:168–178.
39. Yin S, Ding SX, Haghani A, Hao HY, Zhang P. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *J Process Control*. 2012;22:1567–1581.

Manuscript received Feb. 22, 2013, and revision received Sept. 19, 2013.