IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 19, 2017, accepted April 5, 2018, date of publication April 11, 2018, date of current version May 2, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2825451

A Novel Sliding Window PCA-IPF Based Steady-State Detection Framework and Its Industrial Application

YALIN WANG[1], (Member, IEEE), KENAN SUN[1], XIAOFENG YUAN[1], (Member, IEEE), YUE CAO[1], (Student Member, IEEE), LING LI[1], AND HEIKKI N. KOIVO[2], (Senior Member, IEEE)

[1]School of Information Science and Engineering, Central South University, Changsha 410083, China
[2]Department of Electrical Engineering and Automation, Aalto University, 00076 Espoo, Finland

Corresponding author: Xiaofeng Yuan (yuanxf@csu.edu.cn)

This work was supported in part by the Major Program of the National Natural Science Foundation of China under Grant 61590921, in part by the Program of National Natural Science Foundation of China under Grant 61703440, in part by the Foundation for Innovative Research Groups of the National Natural Science Foundation of China under Grant 61621062, in part by the 111 project under Grant B17048, in part by Innovation-driven Plan in Central South University under Grant 2018CX011, and in part by the Fundamental Research Funds for the Central Universities under Grant 222201717006, and in part by the Fundamental Research Funds for the Central Universities of Central South University under Grant 2017zzts023.

ABSTRACT In industrial processes, it is of great significance to carry out steady-state detection (SSD) for effective system modeling, operation optimization, performance evaluation, and process monitoring. Traditional SSD approaches often need to identify process state for each variable and obtain a composite index with sliding window technique, which ignores the variable correlations and is time consuming. Moreover, they can only provide the state of each whole window that slides along data series. To deal with these problems, a novel sliding window principal component analysis-improved polynomial fitting based method is proposed for steady-state detection. In the proposed framework, principal component analysis is first used to eliminate the data correlations and variable noises. Then, the size of sliding window is automatically determined by the data series of the first principal component. After that, SSD is carried out for each selected principal component by 2nd-order improved polynomial fitting. At last, the overall process state is determined by the weighted combination of the SSD results of selected principal components, in which the weight of each principal component is determined by its corresponding contribution of variance. The effectiveness and flexibility of the proposed SSD framework is validated on an industrial hydrocracking process.

INDEX TERMS Steady-state detection, principal component analysis, polynomial fitting, sliding window, hydrocracking process.

I. INTRODUCTION

In modern industrial processes, the real-time detection of steady state is significant for effective process modeling and control. Steady-state models are extensively used for system identification [1]–[3], process modeling and control [4]–[6], data reconciliation [7]–[9], soft sensor and fault diagnosis [10]–[12], etc. At steady state, the process generally runs around certain stable point or within some stationary region. Thus, most of the controlled variables can remain constant or near-constant for a long period of time. However, most industrial processes include both steady and non-stationary states due to reasons like fluctuations in operation conditions and changes in environment, for which the real variable relationships may deviate from the original system design. Deviations from steady-state assumption may lead to wrong real-time process optimization and operation. To keep behavior of the true models close to the corresponding processes, it is necessary to adjust the parameters of steady-state models frequently, which should be performed with only steady, or nearly steady state data. Therefore, steady-state detection is an important step in industrial processes.

With the development of distributed control system (DCS) technologies, a large amount of process data can be collected and recorded for process analysis and modeling, which contains both steady and unsteady state data. It is of great

significance to develop practical techniques for steady-state detection (SSD) to improve process control strategies. By far, researchers have proposed many kinds of steady-state identification methods. Generally, they can be classified into three main categories: model-based, statistical theory based and trend extraction based approaches [13]. Model-based approaches are usually designed to detect process steady state by deeply analyzing the physical and chemical backgrounds of specific processes like mass balance, energy balance, etc. For example, Prabhakar and Kumar [14] proposed an approach for the assessment of voltage stability margins based on the P-Q-V curve technique and Thevenin's equivalent. Dorr *et al.* [15] presented an analytical redundancy technique, which is based on steady-state relationships between measurements. And it is applied for detection, isolation and identification of sensor faults in nuclear power plants. Though model-based techniques can be used to identify steady state in some situations, they are limited mostly to special process plants. They are strongly dependent on the accurate modeling of the processes, which is usually very difficult or costly to obtain, especially for complicated large-scale industrial processes. Moreover, with the running of the processes, the underlying process model may change due to the time-varying problem. However, the process state is usually reflected in the real-time collected process data. It is more reasonable to carry out SSD by data-driven methods. Therefore, statistical test based methods were proposed for steady-state detection by Narasimhan *et al.* [16] and [17], and Maseleno and Hardaker [18] Among them, composite statistical test (CST) [16] and mathematical theory of evidence (MTE) [17], [18] are the two most typically used statistical methods. These methods often assume that the measurements are contaminated by random noise, which obey the Gaussian distribution with mean zero. Then, a window is sliding along the sampling data series. By comparing the mean and covariance between adjacent windows, t-test is used to identify whether the variable is in steady state or not. Also, their improved strategies were developed for practical applications. Then, Rhinehart [19] further proposed a novel R detection method, which utilizes two separate techniques to estimate the variance of data and calculates the ratio of variances estimated by the two techniques for steady-state detection. This method can provide SSD results for variables at each sampling instant. However, it is very sensitive to process noises and easily affected by the selected parameters of filters.

Therefore, another category of SSD approaches was developed with data fitting techniques for data trend extraction, like polynomial function fitting, wavelet transform, particle filtering, etc. Flehmig *et al.* [20] proposed a wavelet-based approach to localize and identify the polynomial trends in noisy data, which is highly computational efficient due to the hierarchical search in the time-frequency plane. Later, Jiang *et al.* [21] developed a wavelet transform based steady-state detection method, in which the process trends are extracted by wavelet-based multi-scale processing from noisy measurements. Wu *et al.* [22] proposed an online SSD strategy using multiple change-point models and particle filters, which can first identify the change points of data and then carry out piecewise linear fitting to extract the data trends. Fu *et al.* [23] proposed an adaptive polynomial filtering method for SSD, in which process steady-state variables are determined by the first-order coefficients of polynomial filtering. This method is easy to implement and faster than other methods. Especially, it is very suitable for online steady-state detection.

As can be seen, for most of traditional SSD methods, they mainly focus on how to detect for a single measured process variable. For multivariate processes, it is necessary to carry out SSD procedure for each variable and then obtain the composite SSD index by weighting on different variables. Hence, they are very computationally complex and time-consuming. This is more difficult for modern industrial processes since there are thousands of measured variables. Also, it is not easy to identify which variables are more important than the others for steady-state detection. Moreover, the correlations between different variables are not considered in the traditional SSD methods. Usually, there are strongly redundancies and correlations between process variables. This may result in false identification results. Hence, it is necessary to eliminate the correlations between variables and capture the main data information before carrying out steady-state detection. As for multivariate processes, the running state can be characterized by the underlying data structure of variable data. To eliminate the correlations and to discover the underlying data structure, it is more desirable to use low-dimensional features to capture the main data information than the original high-dimensional variables. To meet these requirements, principal component analysis (PCA) is adopted to obtain the new latent variables for feature extraction, in which the dimension of latent variables is much lower than the original data dimension. By utilizing PCA, the data information can be mostly retained in the selected principal components while the correlations are largely reduced. Moreover, the state information is mainly kept in these principal components and the process noise is left in the residues. It is more reasonable to detect the steady state in principal components than in the original high-dimensional variables. Therefore, SSD can be simply carried out on the low-dimensional latent variables, which can largely improve the detection efficiency and accuracy.

As a matter of fact, a moving window is needed for most SSD methods. It is very important to select a proper window size. If the window size is too large, then it may fail to detect the steady state and the detection may be delayed. On the other hand, too small a window size may increase the possibility of false detection. Moreover, most of previous works usually detect the window as a whole to be in steady or unsteady state, which is sometimes not accurate since a window may contain both steady- and unsteady- state sampling data simultaneously. It is desirable to provide accurate state detection results for each sampling instant, which is more helpful for real-time process control and optimization.

To deal with these problems, a novel sliding window PCA-IPF based steady-state detecting method is proposed in this paper. First, PCA is applied to process data for dimensionality reduction, in which the principal components carry on the main trend and information of the process state. As the first principal component usually contains the most variance of data, it is used to adaptively determine the size of the sliding window. Then, for each selected principal component, a 2^{nd}-order polynomial function is used to fit the data series in each sliding window. In the detection step, the state for each sampling instant is related to the data trend that are determined by its previous and subsequent data. Hence, the sampling instants is not detected at the two ends of each window. This is detected in its previous (fore-end) or next window (back-end), in which their fitted curve contains trend information of both sides. Then, by calculating the distance between the fitted value and the maximum/minimum value of the fitted curve, the state can be classified as steady or unsteady by defining a novel threshold, which is related to the standard deviation of fitted data in the corresponding detection window and to all other windows. Then, a new composite detection index is designed by weighting for all the selected principal components. Since different principal components give different contributions to the variance of data, they have different importance in the final detection index. The weight for each component is determined by the contribution of covariance in representing the whole data information. The proposed SSD strategy is computationally more efficient and can give more accurate detection results since it can capture the main data trend first and then carry out SSD on each useful principal component. The industrial application also shows the efficiency of the proposed SSD framework.

The remainder of this paper is structured as follows. In Section II, preliminaries about polynomial least squares fitting and principal component analysis are introduced. Then, the proposed PCA-IPF based steady-state detection strategy is described in detail in Section III. In Section IV, the effectiveness and flexibility of the proposed method is evaluated on the industrial hydrocracking process. At last, conclusion and prospect are given in Section V.

II. PRELIMINARIES

A. POLYNOMIAL LEAST SQUARES FITTING

Polynomial least squares function is used to estimate the underlying structure that can describe a set of observations. Given the observed sampling data, it is usually necessary to find a proper fitting curve for them. Polynomial least squares fitting is one of such approaches, which fits the observed data with a polynomial function of time by minimizing the sum of the squares of the offsets.

Suppose the polynomial least squares function with Kth-order degree for a target variable y with time t as [24]

$$y(t) = c_0 + c_1 t + \cdots + c_K t^K \tag{1}$$

where $c_0, c_1, \ldots, c_K$ are the unknown coefficients. Usually, the observed function for the variable is corrupted by an additional stochastic measuring noise, which can be written as

$$\tilde{y}(t) = y(t) + e(t) \tag{2}$$

where $e(t)$ is the measuring noise and $\tilde{y}(t)$ is the measured variable. Given a set of observed samples $\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_N$, where n is the sample index, the aim is to estimate the coefficients of the fitted polynomial function.

Let $\mathbf{c} = [c_0, c_1, \ldots, c_K]^T$ and $\mathbf{r}(t) = [1, t, \ldots, t^K]^T$. Eq. (1) can be rewritten as $y(t) = \mathbf{c}^T\mathbf{r}(t)$. The polynomial exponents for the observed samples are denoted as $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$. By minimizing the sum of the squares of estimated errors, the optimal estimation for parameter $\mathbf{c}$ is

$$\hat{c} = \left(\mathbf{R}^T\mathbf{R}\right)^{-1}\mathbf{R}^T\tilde{\mathbf{y}} \tag{3}$$

where $\mathbf{R} = [\mathbf{r}_1^T, \mathbf{r}_2^T, \ldots, \mathbf{r}_N^T]^T$ and $\tilde{\mathbf{y}} = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_N]^T$. For purposes of simplicity and robustness, it is more common to select the order of polynomial function to be $K = 2$ in SSD studies.

B. PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA [12], [25], [26] is one of the most popular data dimensionality reduction methods used in numerous areas. It aims to find low-dimensional representations for high-dimensional observed data by maintaining the main variance of data. The detailed procedure of PCA is illustrated as follows.

Given a data set of high-dimensional observations $\mathbf{x}_i \in R^M, i = 1, 2, \ldots, N$, where M is the total number of observed variables and N is the number of data samples. We can denote the observed data matrix as $\mathbf{X}$, whose ith row is observation $\mathbf{x}_i$. First, the mean value vector is calculated as

$$\bar{\mathbf{x}} = \sum_{i=1}^{N}\mathbf{x}_i \Big/ N \tag{4}$$

Then the data covariance matrix is obtained as

$$\mathbf{S} = \sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \Big/ N \tag{5}$$

By applying the Eigen decomposition on the covariance matrix

$$\mathbf{SP} = \mathbf{\Lambda}P \tag{6}$$

where $\mathbf{\Lambda}$ is a diagonal eigenvalue matrix with it diagonal eigenvalues as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$, which are arranged in decreasing order for the Eigenvalues of covariance matrix $\mathbf{S}$; $\mathbf{P}$ is the Eigen matrix with its columns being the eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M$ of covariance matrix $\mathbf{S}$ corresponding to its eigenvectors. $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M$ are also the new directions of the principal components. The changes of data are mainly captured in the first few principal components while the redundancy and the noises are left in the last few components. Moreover, the first principal component often carries the most information of the original data and then the second one. The contribution of variance is usually used to measure the data

information that contains in the principal component. The contribution of variance for the dth principal component is calculated as follows.

$$CV_d = \lambda_d \Big/ \sum_{j=1}^{M} \lambda_j \tag{7}$$

To reduce the noise, collinearity and redundancy of data, only the first few components that capture the underlying structure of data are kept for further data analysis. Hence, several techniques can be used to determine the number of principal components in PCA. Among them, the cumulative contribution of variance (CCV) technique is more often used, which is defined as

$$CCV_D = \sum_{i=1}^{D} \lambda_i \Big/ \sum_{j=1}^{M} \lambda_j \tag{8}$$

where D is the number of components to be kept in PCA, which is determined by certain threshold for the index of CCV_D.

III. SLIDING WINDOW PCA-IPF BASED STEADY-STATE DETECTION

For a single variable system, if the variable tends to be stable for a certain period of time, the system is considered to be at steady-state, and the sampling data in this time interval is steady-state data. As the variable data have complicated characteristics like nonlinearities, a single polynomial function is not sufficient to accurately model the trend of variable data. Sliding windows are often used to fit the whole variable curve with piece-wise polynomials. However, as for multivariate systems, the operating variables cannot be in steady state for a long period of time in the actual industrial process due to switching of operation instructions and the adjustment of equipment condition. Therefore, not all the variables can be steady, and they will vary with time to some degree. Usually, when most of the operating variables are in steady state, the multivariate system is regarded to be at steady state. Due to the large number of operating variables, steady-state detection of variable one by one at a time will lead to a heavy computational burden. Moreover, steady-state detection variables are often strongly correlated as a result of redundant sensors and mechanism relationships. Hence, before curve fitting for the trend of variable, it is necessary to carry out dimension reduction to eliminate the redundant data information and capture the main data structure. PCA is able to reduce the dimension effectively and maintain data information in the first few principal components. Hence, it is more reasonable to first carry out PCA on data to extract the main data features. Then, polynomial fitting can be used to model the trend of each principal component by sliding window technique. After that, steady-state index is calculated for each principal component and a synthetic index is obtained for steady-state detection. Fig. 1 shows the basic flowchart of PCA-IPF based steady-state detection method.

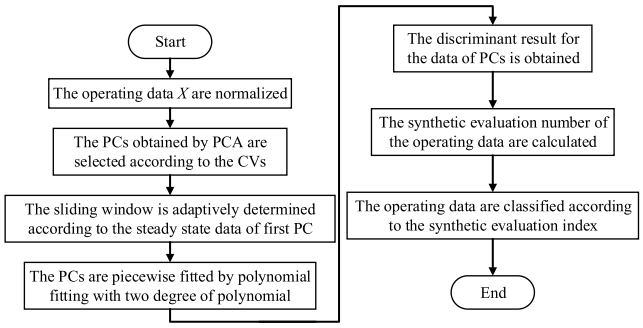


FIGURE 1. The flowchart of the proposed steady-state detection framework.

The detailed procedure is summarized as follows:

1) Assume the data series are $\mathbf{X} = [\mathbf{x}_1, \ldots \mathbf{x}_i, \ldots, \mathbf{x}_N]$, where $\mathbf{x}_i = [x_{i,1}, \ldots, x_{i,j} \ldots, x_{i,M}]^T$. Here, N is the number of samples and M is the number of variables for steady-state detection; i and j are the sample and variable indices, respectively.
2) For each variable j, calculate its mean value $\bar{x}_j$ and standard deviation δ_j from the observed data. Normalize each of the sample as follows

$$\tilde{x}_{i,j} = (x_{i,j} - \bar{x}_j)/\delta_j \tag{9}$$

Denote the data matrix after normalization as $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1, \ldots \tilde{\mathbf{x}}_i, \ldots, \tilde{\mathbf{x}}_N]$.
3) Apply PCA on the normalized data matrix $\tilde{\mathbf{X}}$, and determine the number D of principal components to be kept. To select single principal component that contains enough information in itself, a threshold θ_{CV} for contribution of variance is set to choose the first few components satisfying

$$CV_d \geq \theta_{CV}, \quad d = 1, 2, \ldots, D \tag{10}$$

Then, the cumulative contribution of variance of the selected principal components is calculated to test if it is greater than the predefined threshold θ_{CCV}. If so, then the final number of principal components is D. if not, then new principal component is added one by one until the following condition is reached

$$CCV_D \geq \theta_{CCV} \tag{11}$$

where CCV_D is the cumulative contribution of variance for the first D principal components; θ_{CCV} is the predefined threshold. Thus, the directions of the D principal components are $\mathbf{p}_1, \ldots, \mathbf{p}_d, \ldots, \mathbf{p}_D$. Then, the score of the ith sample on the dth principal direction is calculated as

$$z_{i,d} = \tilde{\mathbf{x}}_i^T \mathbf{p}_d \tag{12}$$

After applying PCA on the whole dataset, we can obtain D pieces of principal component time series as $z_{1,d}, z_{2,d}, \ldots, z_{N,d}, (d = 1, 2, \ldots, D)$.

4) Determine the size H of the sliding window by the first PC time series, which is described in detail in Section III.A.
5) Fit each selected principal component by a 2nd-order polynomial function with sliding window technique. The fitted time series become $\tilde{z}_{1,d}, \tilde{z}_{2,d}, \ldots, \tilde{z}_{N,d}$, $(d = 1, 2, \ldots, D)$ for each principal component. Moreover, calculate the discriminant result of steady state index for each PC series. Details are described in Section III.B.
6) Compute the synthetic steady-state evaluation index by a weighted sum of the principal component indices, which are described in Section III.

A. DETERMINATION OF WINDOW SIZE

The concept of the steady-state detector [11] initially originates from the theory of noise filter. As one of the simplest and most common methods, sliding window technique is often used for steady-state detectors by analyzing statistical characteristics of data. A predefined time interval is established over which the data are fitted by methods like mean filter or polynomial function. This produces an array of fitting data, which are much smoother than the original data. Moreover, they can better represent the data trend. Hence, fitting data in the sliding window can be used to replace each data point within the timespan for steady-state detection. Since the original data of detection variables contains noise and correlations, they are not suitable to be used for steady state detection directly. In order to effectively eliminate noise in data, correlations between variables and better extract data trend, principal components of data are used to detect steady state by sliding window technique in this paper. To utilize the sliding window technique, the first step needed to determine is the window size. Here, a novel window size selection method is adopted.

The window size is strongly related to the main data trend of steady state. Traditional methods usually determine it with certain critical variables, which may not represent the main data information of the steady state. To alleviate this problem, the first principal component of data is used to set the proper window size for steady-state detection instead. As mentioned before, the first principal component contains the main information of data, which represents the underlying structure of steady-state data. Hence, it can reflect the main data trend of the process state. First, by artificially checking the time series of the first principal component of data, a piece of it that remain steady are manually selected as the standard learning time series for deciding the window size. Denote the first PC standard learning series as $z_{s_1,1}, z_{s_2,1}, \ldots, z_{s_L,1}$, where L is the total number of samples in the standard series. Denote the window size is H. Therefore, the first window is constructed and a 2nd-order polynomial function is used to fit the first principal component data in it. After that, the window is moved forward with step of H time intervals and the first PC data can be fitted by the 2nd-order polynomial function for the second window. By sliding the window along this time series of the first principal component by step of H sequentially, we can repeat this procedure until the polynomial fitting is finished for the first PC data of the whole standard learning time series. Assume the time series are $\tilde{z}_{s_1,1}, \tilde{z}_{s_2,1}, \ldots, \tilde{z}_{s_L,1}$ after 2nd-order polynomial fitting with sliding windows. Then, standard deviations can be calculated for $z_{s_1,1}, z_{s_2,1}, \ldots, z_{s_L,1}$ and $\tilde{z}_{s_1,1}, \tilde{z}_{s_2,1}, \ldots, \tilde{z}_{s_L,1}$, which are denoted δ_s and $\tilde{\delta}_s$, respectively. The normalized standard deviation is determined by

$$\delta_H = \tilde{\delta}_s / \delta_s \tag{13}$$

For small values of H, there are very few samples in each window. Thus, a 2nd-order polynomial function is easy to be over-fitted, which will result in a large $\tilde{\delta}_s$ in the fitted standard deviation. Too large a value of H leads to under-fitting for first PC data in the window. In this case, the fitted standard deviation tends to be very small. Therefore, a reasonable value of H is determined by a predefined threshold of δ_H, the procedure is shown in Table 1.

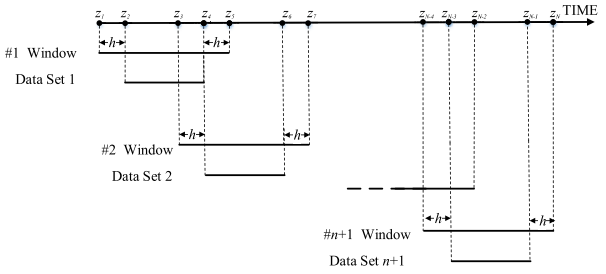
TABLE 1. The procedure of determination of the window size H.

1) Select a steady time series of first PC data as the standard learning samples $z_{s_1,1}, z_{s_2,1}, ..., z_{s_L,1}$
2) Set the initial window size H to be a small value like 2 sampling intervals.
3) Fit the standard learning samples by 2nd-order polynomial function with sliding window of size H, the time series after polynomial fitting are denoted as $\tilde{z}_{s_1,1}, \tilde{z}_{s_2,1}, ..., \tilde{z}_{s_L,1}$
4) Compute the standard variance of the original first PC series and the fitted first PC series, which are denoted as δ_s and $\tilde{\delta}_s$.
5) Calculate the normalized standard deviation δ_H .
6) Compare the normalized standard deviation δ_H with the predefined threshold θ_δ . If $\delta_H > \theta_\delta$, set H=H+1 and return to step 3); else if $\delta_H \le \theta_\delta$, H is the optimal selected window size.

B. THE WINDOW FITTING AND DETECTION STRATEGY

Different from traditional sliding window-based steady-state detection methods, which estimate the whole window as a steady or non-steady region, the polynomial fitting approach can evaluate the steady state for each sampling instant, which extracts the general trend from data for steady-state detection. Hence, for each sampling instant, its state is related to the data trend that is determined by its previous and subsequent data. Hence, the fitting and state detection procedures for each component are slightly different from that used in determining the window size.

Here, we use $z_1, z_2, \ldots, z_N$ to represent one general PC series data. For each window, the data are fitted by a quadratic function to extract the data trend. Then, the state detection can be carried out for these samples in this window. However,



#1 Window :the 1st sliding window; Data Set 1:reserve the data for judgment in Window 1
#2 Window :the 2nd sliding window; Data Set 2:reserve the data for judgment in Window 2
#n+1 Window :the (n+1)th sliding window; Data Set n+1:reserve the data for judgment in Window n+1

FIGURE 2. The illustration of window fitting and detection strategy.

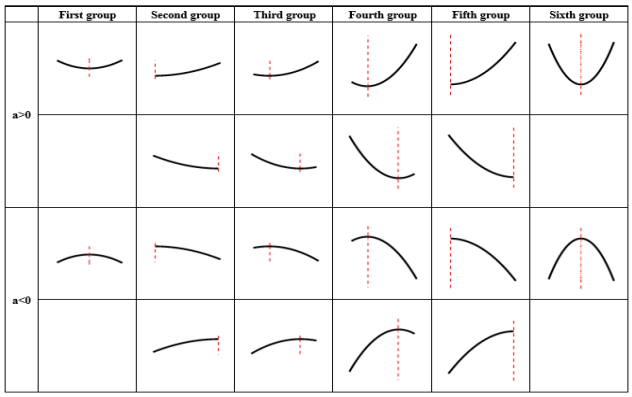


FIGURE 3. The groups of fitted curves.

for the first and last few samples, the data trends are mainly determined by its latter and previous samples, respectively. This indicates that the trends of these samples are not completed. Hence, for each window, only the middle parts of samples are detected for state. Denote the length of sample intervals at two ends of the window as h, which usually satisfies $h \leq H/3$. For each window, after the data are fitted, only the samples from $h+1$ to $H-h$ are detected for steady state. After one window is detected, it will be moved forward by $H-2h$ sample intervals. Fig. 2 gives the illustration about how this fitting and detecting procedures are carried out. In this figure, the data in "Window n" are fitted by the 2^{nd}-order polynomial. Then each sample in the middle "Data n" of this window is detected for steady state. After that, this window is forwarded by $H-2h$ steps to get "Window $n+1$".

C. THE STEADY-STATE DETECTION INDEX

In the ideal case, the curve for the steady-state data should be a horizontal line because variables remain unchanged. However, measured values always fluctuate due to the affection of equipment or other conditions in real industrial processes. For each PC series, polynomial function of degree two is fitted in the sliding windows. In each window, the graph after quadratic function fitting is a part of parabola. There are totally six groups of fitting curves, which are shown in Fig. 3. When the size of the fitting window is the same, the more sharply curved the graph appears, the smaller fluctuation the data changes. Meanwhile, the fitting curve should not show a general increasing or decreasing trend, which means that the symmetry axis is located in the fitting curve. Moreover, the steady-state data generally fluctuates less and should be close to the extreme value of the fitting curve. Hence, the first group of fitting curves are expected to appear in practice. The second and third group curves could also exist when there is a long period of data fluctuation. As for the fourth and fifth groups, though the whole data trend seems to be unsteady, there may be some steady-state data if there is short regulating time or measurement fluctuation. In the the last group, it is easily seen that the process data is unsteady.

Here, 3δ rule is used to determine steady-state samples. Assume the fitting curve function for the nth window of the dth PC data series is given by

$$\tilde{z}_d^n(t) = a_d^n t^2 + b_d^n t + c_d^n \tag{14}$$

where a_d^n, b_d^n, c_d^n are the quadratic coefficient, first-order coefficient and constant term, respectively. Then the fluctuation of the fitted value at sampling time t to the maximum or minimum value of this function is calculated as

$$s_d^n(t) = \left|\tilde{z}_d^n(t) - \left(4a_d^n c_d^n - (b_d^n)^2\right)/4a_d^n\right| \tag{15}$$

In order to identify steady and unsteady state points, the threshold for the fluctuation is defined as

$$\theta_{s_d^n} = w\delta_d^n + (1-w)\bar{\delta}_d \tag{16}$$

where δ_d^n is the standard deviation of the fitted data for the dth PC in the current nth window; $\bar{\delta}_d$ is the corresponding average value of standard deviations for all windows; w is the weight to control the trade-off between these two deviations. Meanwhile, to avoid misjudgment of the peak-valley data in the sixth group, standard deviation of current window is guaranteed less than a certain range.

$$\delta_d^n < \theta_{\delta_d^n} \tag{17}$$

Therefore, the steady detection index for sample at time t from the nth window of the dth PC series is determined as

$$\psi_d^n(t) = \begin{cases} 1 & s_d^n(t) < 3\theta_{s_d^n} \text{ and } \delta_d^n < \theta_{\delta_d^n} \\ 0 & \text{else} \end{cases} \tag{18}$$

Finally, the state of the whole process is detected by the synthetic evaluation of these principal components. As is mentioned before, different principal components provide different contribution of variance in representing the whole data. Thus, a novel synthetic evaluation index is designed as the normalized sum of the evaluation index for each component

$$\psi(t) = \sum_{d=1}^{D} CV_d \cdot \psi_d^n(t) \Big/ \sum_{d=1}^{D} CV_d \tag{19}$$

By predefining a threshold θ_ψ for the synthetic evaluation index, the process state can be classified into

steady or unsteady state as

$$SS(t) = \begin{cases} 1\ (steady\ \text{state}) & \psi(t) > \theta_\psi \\ 0\ (unsteady\ \text{state}) & \text{otherwise} \end{cases} \tag{20}$$

IV. INDUSTRIAL CASE STUDY

In this section, the feasibility and efficiency of the proposed steady-state detection method is illustrated in an industrial hydrocracking process.

A. THE HYDROCRACKING PROCESS

The hydrocracking [27], [28] is an important part of the refining process, which aims to crack the high-boiling, high-molecular, low-quality heavy gas oils, heavy diesels or heavy distillates into more valuable low-boiling light distillates (like naphtha, diesel, kerosene, etc.), or base stock for lubricating oil manufacture. Two main kinds of reactions, hydrogenation and cracking reactions, are involved in this process. For the hydrogenation reactions, carbon-carbon double bonds are hydrogenated, which are highly exothermic and can liberate the heat for cracking reactions. While in the cracking reactions, carbon-carbon single bonds are cracked, which are slightly endothermic and provide olefins for the hydrogenation reactions. Since it can process a number of gas oils and produce valuable products with low sulphur content and high smoking point jet fuel, hydrocracking has been a very important refinery process that can adequately meet the requirements of green, clean and environmentally friendly fuels. Here, the proposed steady-state detection method is applied to an industrial hydrocracking process at a refinery from SINOPEC in China. The flowchart of this process is shown in Fig. 4, which mainly consists of the hydrogen compression, reaction, separation and fraction parts. First, the new hydrogen and recycled hydrogen are compressed and pre-heated to provide a continuous supply of hydrogen to the reaction part. Meanwhile, the raw oil materials are fully mixed and fed to reaction part. In the reaction part, the feeds of hydrogen and oil materials are combined to carry out the hydrogenation and cracking reactions. By a series of cooling, heating and heat exchanges, different products can be obtained after the separation and fraction section.

From the above description, there are numerous devices, reactions, manipulations and control strategies involved in this complex process. Hence, a large number of process parameters and indices need to be monitored and adjusted for real-time optimization, control and adjustment. Due to reasons like changes of raw material, process condition and product demand, the process should be optimized and controlled at different regions regularly. As this process is inherently nonstationary and parameter adjustment should be performed with nearly steady-state data, it is important to identify the steady-state region for effective and satisfactory control and optimization for this process. Also, there are a large number of variables being measured and collected in this process. Hence, it is computationally complex and impractical to calculate the steady-state index for every individual variable and combine the overall steady-state index. Moreover, the steady-state detection variables are usually strongly correlated with each other. To carry out steady-state detection effectively, it is necessary to apply dimension reduction to obtain the main trends of data information.

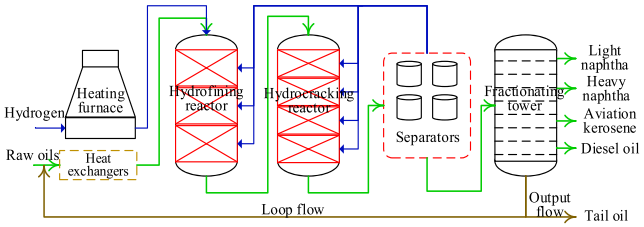


FIGURE 4. The flowchart of the hydrocracking process.

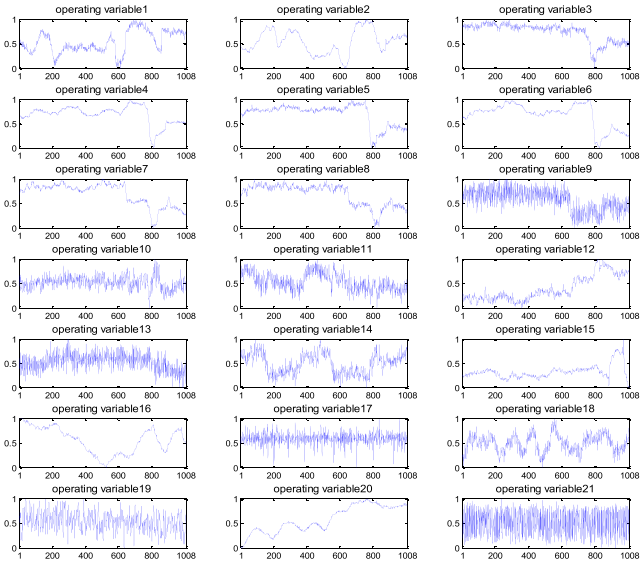


FIGURE 5. The trend information of the 21 operating variables.

B. STEADY-STATE DETECTION RESULTS AND DISCUSSIONS

There are totally 21 critical variables selected and collected as the steady-state detection variables from the reactors, stripper, fractionation parts of the hydrocracking plant. The sampling frequency of each variable is 5 minutes per sample. The individual variable trends are shown in Fig. 5. It can be seen that the measured variables are contaminated by process noise. Moreover, there is information redundancy between variables. Hence, it is necessary carry out PCA to eliminate the noises and correlations before steady-state detection. The thresholds for initial individual contribution of variance and cumulative contribution of variance are set at 3.5% and 85%, respectively. Hence, there are totally nine principal components extracted to keep the main information of data. After applying PCA on data, we can use the first PC sequence to adaptively determine the size of the sliding window. By manually selecting a steady piece of data series from the first PC sequence, the strategy described in Section III.A is used to evaluate the relationship between the normalized standard

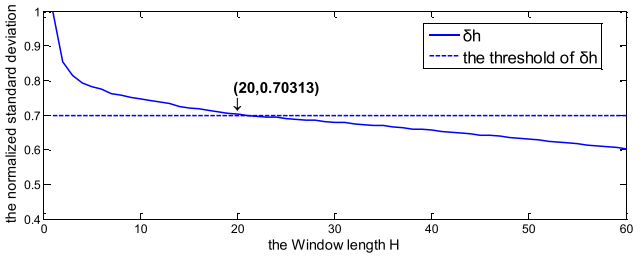


FIGURE 6. The relationship between the window length and the normalized standard deviation.

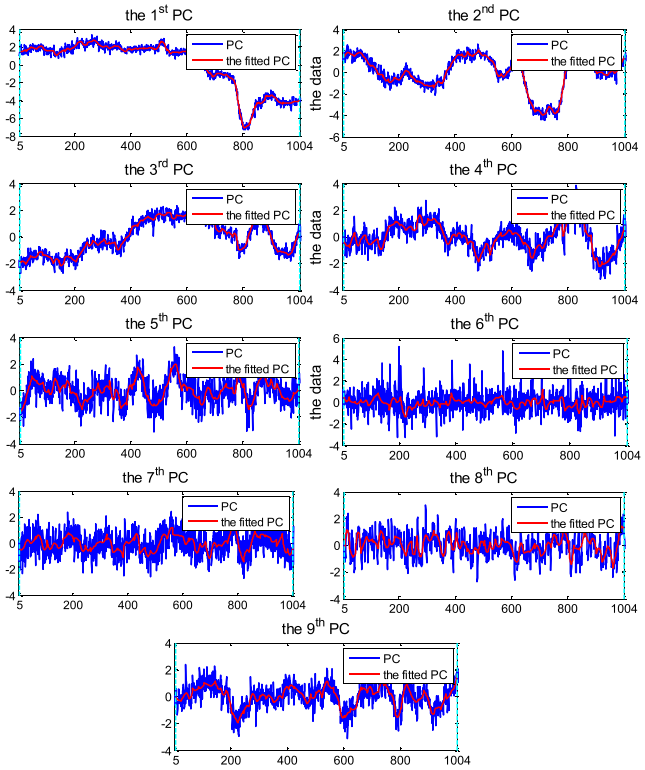


FIGURE 7. The fitted results for the data of PCs.

deviation δ_H and the window size, which is shown in Fig. 6. From this figure, it can be seen that δ_H decreases sharply when H is small and slightly when H increases to a certain extent. Hence, the dotted line represents that the threshold of the normalized standard deviation is 0.7. With this threshold, the window size is determined to be 20. Then, the discarded length of the detection is set as $h = H/5$ by trial and error.

After the window size is determined, each of the PC series is fitted with piecewise 2nd-order polynomial by sliding window strategy as described in Section III.B. The fitting results of the PC data are shown in Fig.7. It can be seen that the first PC can capture the main data information and the fitted curve reflects the smooth trend of this PC. With the increase of PC number, the corresponding PC occupies less data information than the former ones. To detect the steady-state samples for each PC, the trade-off weight w is determined to be 0.2. The threshold of the standard deviation $\theta_{\delta_d^n}$ is set as 0.5.

Hence, we can evaluate the steady-state index for each sampling instant for different PCs. Then, the process

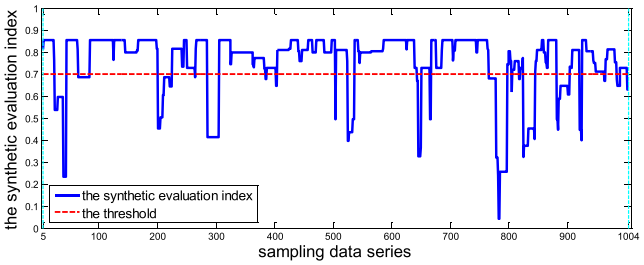


FIGURE 8. The synthetic evaluation number of the operating data.

TABLE 2. The detection results of the four methods on different datasets.

Method	1	2	3	4	5	6
PCA -IPF	94.4%	91%	93.4%	88.6%	89.2%	100%
PCA -IPF 1	83%	81%	86.6%	89%	85.4%	98.6%
PCA -IPF 2	44.8%	43.4%	28.2%	36.6%	48.3%	56%
R-statistical	81%	78.8%	94%	91.6%	66.2%	22.2%

steady-state results are determined by the synthetic weighted sum of individual steady-state index for each PC. The detection results are shown in Fig. 8. Here, the threshold for the synthetic index is set as 70%, which indicates that the sample points above the red dot line are detected as steady state data while the others are unsteady state.

For performance comparison, we have further evaluated the proposed detection method on 6 different datasets with three other approaches, which are the R-statistical method, PCA-IPF1 (The window is sliding forward with step one), PCA-IPF2 (with discriminant criterion from [23] and [29]). The data changes frequently in datasets 3 and 4, while datasets 5 and 6 have a smooth data trend. The detection accuracy is shown in Table 2 for the four methods on the six datasets. As PCA- IPF1 cannot extract the trend accurately at the edge of each window, its detection accuracy is lower than PCA-IPF. Moreover, for PCA-IPF and PCA-IPF2, IAF-PCA can provide the detection results for each sampling time, while IAF-PCA2 can only give the overall detection result for the whole window, in which there may be both steady and unsteady points. Hence, its detection accuracy is much lower than PCA-IPF and PCA-IPF1. The R-statistical method gives no indication about how close the process is to the steady state because the detection result is only obtained by comparing the changes in two points. Therefore, the R-statistical method performs a little better in test 3 and test 4, in which the process data changes frequently. However, for the other four datasets, R-statistical method can only provide much lower accuracy of steady-state samples than PCA-IPF.

V. CONCLUSION

In this paper, the limitations of traditional steady-state detection methods are mainly focused, which usually ignore the correlations of variables and cannot provide accurate point SSD result for process sampling instants. Therefore, PCA is utilized to process data for main feature extraction in order to eliminate data correlations, redundancy and noises. Then, SSD can be carried out on the selected principal components,

which can represent the main trends of process data. As the first principal component usually carries the most data information, it is used to determine the size of sliding window. After that, the 2nd-order polynomial is fitted for each component in the sliding windows. The fluctuation is calculated between the fitted value at each sampling time and the maximum or minimum value of the fitted function. Also, the threshold is adaptively determined by the fitting function for all data series. By comparing the fluctuation and the threshold, the state can be determined at different sampling instants for each principal component. At last, the final process state is calculated by weighted sum of each principal component.

REFERENCES

[1] B. Jiang, F. Yang, W. Wang, and D. Huang, "Simultaneous identification of bidirectional path models based on process data," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 2, pp. 666–679, Apr. 2015.

[2] X. Yuan, Y. Wang, C. Yang, Z. Ge, Z. Song, and W. Gui, "Weighted linear dynamic system for feature representation and soft sensor application in nonlinear dynamic industrial processes," *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1508–1517, Feb. 2018.

[3] J. J. Shynk and N. J. Bershad, "Steady-state analysis of a single-layer perceptron based on a system identification model with bias terms," *IEEE Trans. Circuits Syst.*, vol. 38, no. 9, pp. 1030–1042, Sep. 1991.

[4] X. Yuan, B. Huang, Y. Wang, C. Yang, and W. Gui, "Deep learning based feature representation and its application for soft sensor modeling with variable-wise weighted SAE," *IEEE. Trans. Ind. Informat.*, to be published, doi: 10.1109/TII.2018.2809730.

[5] A. G. Marchetti, A. Ferramosca, and A. H. González, "Steady-state target optimization designs for integrating real-time optimization and model predictive control," *J. Process Control*, vol. 24, no. 1, pp. 129–145, 2014.

[6] X. Yuan, Z. Ge, B. Huang, Z. Song, and Y. Wang, "Semisupervised JITL framework for nonlinear industrial soft sensing based on locally semisupervised weighted PCR," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 532–541, Apr. 2017.

[7] M. Schladt and B. Hu, "Soft sensors based on nonlinear steady-state data reconciliation in the process industry," *Chem. Eng. Process., Process Intensification*, vol. 46, no. 11, pp. 1107–1115, 2007.

[8] S. A. Bhat and D. N. Saraf, "Steady-state identification, gross error detection, and data reconciliation for industrial process units," *Ind. Eng. Chem. Res.*, vol. 43, no. 15, pp. 4323–4336, 2004.

[9] M. Korbel, S. Bellec, T. Jiang, and P. Stuart, "Steady state identification for on-line data reconciliation based on wavelet transform and filtering," *Comput. Chem. Eng.*, vol. 63, pp. 206–218, Apr. 2014.

[10] X. Yuan, Z. Ge, B. Huang, and Z. Song, "A probabilistic just-in-time learning framework for soft sensor development with missing data," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 3, pp. 1124–1132, May 2017.

[11] M. Kim, S. H. Yoon, P. A. Domanski, and W. V. Payne, "Design of a steady-state detector for fault detection and diagnosis of a residential air conditioner," *Int. J. Refrig.*, vol. 31, no. 5, pp. 790–799, 2008.

[12] X. Yuan, Z. Ge, and Z. Song, "Locally weighted kernel principal component regression model for soft sensing of nonlinear time-variant processes," *Ind. Eng. Chem. Res.*, vol. 53, no. 35, pp. 13736–13749, 2014.

[13] J. Liu, M. Gao, Y. Lv, and T. Yang, "Overview on the steady-state detection methods of process operating data," *Chin. J. Sci. Instrum.*, vol. 34, no. 8, pp. 1739–1748, 2013.

[14] P. Prabhakar and A. Kumar, "Performance evaluation of voltage stability index to assess steady state voltage collapse," in *Proc. 6th IEEE Power India Int. Conf.*, Dec. 2014, pp. 1–6.

[15] R. Dorr, F. Kratz, J. Ragot, F. Loisy, and J.-L. Germain, "Detection, isolation, and identification of sensor faults in nuclear power plants," *IEEE Trans. Control Syst. Technol.*, vol. 5, no. 1, pp. 42–60, Jan. 1996.

[16] S. Narasimhan, R. S. H. Mah, A. C. Tamhane, J. W. Woodward, and J. C. Hale, "A composite statistical test for detecting changes of steady states," *AIChE J.*, vol. 32, no. 9, pp. 1409–1418, 1986.

[17] S. Narasimhan, S. K. Chen, and R. S. H. Mah, "Detecting changes of steady states using the mathematical theory of evidence," *AIChE J.*, vol. 33, no. 11, pp. 1930–1932, 1987.

[18] A. Maseleno and G. Hardaker, "Malaria detection using mathematical theory of evidence," *Songklanakarin J. Sci. Technol.*, vol. 38, no. 3, pp. 257–263, 2016.

[19] R. R. Rhinehart, "A novel method for automated identification of steady-state," in *Proc. Amer. Control Conf.*, vol. 6. Jun. 1995, pp. 4065–4066.

[20] F. Flehmig, R. V. Watzdorf, and W. Marquardt, "Identification of trends in process measurements using the wavelet transform," *Comput. Chem. Eng.*, vol. 22, no. 12, pp. S491–S496, 1998.

[21] T. Jiang, B. Chen, X. He, and P. Stuart, "Application of steady-state detection method based on wavelet transform," *Comput. Chem. Eng.*, vol. 27, no. 4, pp. 569–578, 2003.

[22] J. Wu, Y. Chen, S. Zhou, and X. Li, "Online steady-state detection for process control using multiple change-point models and particle filters," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 688–700, Apr. 2016.

[23] K.-C. Fu, I.-K. Dai, and T.-J. Wu, "Method of adaptive steady-state detection based on polynomial filtering," *Control Instrum. Chem. Ind.*, vol. 33, no. 5, p. 18, 2006.

[24] A. Marco and J.-J. Martı, "Polynomial least squares fitting in the Bernstein basis," *Linear Algebra Appl.*, vol. 433, no. 7, pp. 1254–1264, 2010.

[25] B. R. Bakshi, "Multiscale PCA with application to multivariate statistical process monitoring," *AIChE J.*, vol. 44, no. 7, pp. 1596–1610, 1998.

[26] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, nos. 1–3, pp. 37–52, 1987.

[27] J. Ancheyta, S. Sánchez, and M. A. Rodríguez, "Kinetic modeling of hydrocracking of heavy oil fractions: A review," *Catalysis Today*, vol. 109, nos. 1–4, pp. 76–92, 2005.

[28] H. K. And and G. F. Froment, "Mechanistic kinetic modeling of the hydrocracking of complex feedstocks, such as vacuum gas oils," *Ind. Eng. Chem. Res.*, vol. 46, no. 18, pp. 5881–5897, 2006.

[29] S. Xie, C. Yang, Y. Xie, and X. Wang, "The steady state detection based on outliers identification for sodium aluminate solution evaporation process," in *Proc. Chin. Autom. Congr.*, Nov. 2015, pp. 281–285.

YALIN WANG (M'17) received the B.Eng. degree in industrial electrical automation and the Ph.D. degree in control science and engineering from Central South University, Changsha, China, in 1995 and 2001, respectively. She visited the University of Alberta, Canada, from 2011 to 2012. She is currently a Full Professor with the School of Information Science and Engineering, Central South University. Her research interests include the modeling, optimization, and control of complex industrial processes, pattern recognition, and machine learning.

KENAN SUN received the B.Eng. degree in measurement and control technology and instrument from Central South University, Changsha, China, in 2015, where he is currently pursuing the M.S. degree with the School of Information Science and Engineering. His research interests include process control and optimization, machine learning, and data mining.

XIAOFENG YUAN (M'17) received the B.Eng. and Ph.D. degrees from the Department of Control Science and Engineering, Zhejiang University, Hangzhou, China, in 2011 and 2016, respectively.

He was a Visiting Scholar with the Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB, Canada, from 2014 to 2015. He is currently an Associate Professor with the School of Information Science and Engineering, Central South University. His research interests include big data and deep learning, artificial intelligence and machine learning, and data-driven modeling for industrial processes.

YUE CAO (S'18) received the B.Eng. degree in automation from Central South University, Changsha, China, in 2014, where he is currently pursuing the Ph.D. degree with the School of Information Science and Engineering. His research interests include process monitoring, fault diagnosis, and machine learning.

LING LI received the B.Eng. degree in automation from Xiangtan University in 2010 and the M.S. degree in control theory and control engineering from the Nanjing University of Science and Technology in 2013. She is currently pursuing the Ph.D. degree with the School of Information Science and Engineering, Central South University, Changsha, China. Her research interests include process performance assessment, machine learning, and data mining.

HEIKKI N. KOIVO (S'67–M'71–SM'86) received the B.S.E.E. degree from Purdue University, West Lafayette, IN, USA, and the M.S. degree in electrical engineering and the Ph.D. degree in control sciences from the University of Minnesota, Minneapolis, MN, USA. He has served in various academic positions at the University of Toronto, Toronto, ON, Canada, and at the Tampere University of Technology, Tampere, Finland. Since 1995, he has been a Professor in control engineering with the Helsinki University of Technology, Helsinki, Finland. He is a Fellow of Finnish Academy of Technology.