Research Article

Comprehensive Monitoring of Complex Industrial Processes with Multiple Characteristics

Chenxing Xu [ID],[1] **Jiarula Yasenjiang** [ID],[1] **Pengfei Cui** [ID],[1] **Shengpeng Zhang** [ID],[2] **and Xin Zhang** [ID][2]

[1]*College of Mechanical Engineering, Xinjiang University, Urumqi, Xinjiang, China*
[2]*China National Petroleum Corporation, Western Drilling and Engineering Research Institute, Urumqi, Xinjiang, China*

Correspondence should be addressed to Jiarula Yasenjiang; dtcs118@163.com

Traditional onefold data-driven methods for fault detection in complex process industrial systems with high-dimensional, linear, nonlinear, Gaussian, and non-Gaussian coexistence often have less than satisfactory monitoring performance because only a single distribution of process variables is considered. To address this problem, a hybrid fault detection model based on PCA-KPCA-ICA-KICA-BI (Bayesian inference) is proposed, taking into account the advantages of principal component analysis (PCA), kernel principal component analysis (KPCA), independent component analysis (ICA), and kernel independent component analysis (KICA) in terms of dimensionality reduction and feature extraction. Foremost, this paper proposed a nonlinear evaluation method and divided the feature variables into Gaussian linear blocks, Gaussian nonlinear blocks, non-Gaussian linear blocks, and non-Gaussian nonlinear blocks by using the Jarque–Bera (JB) test and nonlinear discrimination method. Each division was monitored by the PCA-KPCA-ICA-KICA model, and finally the Bayesian fusion strategy proposed in this study is used to synthesize the detection results for each block. The hybrid model helps in evaluating variable features and bettering detection performance. Ultimately, the superiority of this hybrid model was verified through the Tennessee Eastman (TE) process and the Continuous Stirred Tank Reactor (CSTR) process, and the fault monitoring results showed an average accuracy of 85.91% for this hybrid model.

1. Introduction

People now set a higher threshold for the production quality, system performance, economical efficiency, and increasingly complicated grow process industry systems at the structure and automation level, amid the rapid advance of artificial intelligence and sensor detection technology. Reliability and safety of complex industrial processes now jostle for increasing attention and have to be safeguarded urgently. Industrial process fault detection shines in effectively improving product quality, safe operation, and continuous production. Condition monitoring can be performed by knowledge- and model-based or data-driven approaches [1–5]. The first two kinds face great challenges in dealing with complex process industrial system as they require a large number of model parameters and prior knowledge. Operation data can be recorded, transmitted, and stored; these features lay the foundation for data-driven fault detection techniques [6–8].

Data-driven fault detection methods have been intensively studied over the past years. Great progress was seen in the Multivariate Statistical Process Monitoring (MSPM). For example, principal component analysis (PCA), partial least squares (PLS), canonical correlation analysis (CCA), independent component analysis (ICA), and other methods have been widely studied and applied [3, 9–12]. MSPM dimensionally reduces high-dimensional process variables to low-dimensional space by projection under normal circumstances, thus outputting a process monitoring model. To address the issue of early fault detection, because the memory monitoring charts are sensitive to incipient anomalies in the process mean, the advantages of PCA and multivariate memory monitoring schemes are used for early abnormality detection, and the results showed that the scheme could detect early anomalies in multivariate data [13]. Given complex industrial process systems, varying basic MSPM improvement schemes have been proposed. For example, to reduce

computational cost and time while ensuring that important information is not lost or deleted, a dynamic simplification method which enhances the nonlinear process monitoring capability of KPLS is proposed [14]. In addition, an iterative robust kernel principal component analysis is brought forth to improve the robustness of the fault detection model by iteratively optimizing the function index and the kernel method [15]. The KPCA was revised to deal with nonlinear optimization problems. The newly improved approach greatly enhanced the fault detection performance of the original kernel principal component analysis [16]. Facts prove that not all variables of complex industrial processes completely obey the Gaussian distribution. ICA outperforms PLS as it saves the trouble of determining the input/output relationship of the system. Therefore, a new kernel independent component analysis method was driven by the aforementioned non-Gaussian nonlinear distribution. The proposed method effectively captures the nonlinear relationship in the process variables [17]. The above methods have improved the fault detection performance and model stability to a certain extent. However, the above methods all have certain assumptions, such as variables obeying a Gaussian or non-Gaussian distribution and linear or nonlinear correlation between variables. Such assumptions are unrealistic in complex industrial processes. Therefore, complex industrial processes require further research.

Given the characteristic that the process variables do not fully obey the Gaussian distribution, [18] an approach integrating ICA and PCA was proposed, which monitors non-Gaussian distributed variables and Gaussian distributed variables making use of ICA and PCA, respectively. Reference [19] proposed Gaussian and non-Gaussian dual subspace statistical process monitoring. Foremost, the D test identified the normality of the process variables, then divided the process variables into Gaussian and non-Gaussian subspaces, and finally classified the PCA and ICA models for fault detection in Gaussian and non-Gaussian subspaces, respectively. Reference [20] proposed a dynamic non-Gaussian mixture serial modeling method for industrial process monitoring, using the multivariate non-Gaussian evaluation method to divide the industrial process variables into Gaussian variable subspace and non-Gaussian variable subspace. Then, using DICA and DPCA, the two subspaces were monitored, and the monitoring performance was improved to a certain extent. Aiming at the characteristics of incomplete linear correlation of process variables, [21] proposed a new hybrid linear-nonlinear statistical model (SPCA) for nonlinear process monitoring, using PCA to extract the characteristics of linear features to monitor the linear subspace. Using KPCA can extract the characteristics of nonlinear features to monitor the nonlinear subspace, which makes better use of the underlying process to improve the monitoring performance. Reference [22] proposed a parallel PCA-KPCA (P-PCA-KPCA) model and monitoring scheme combining stochastic algorithm (RA) and genetic algorithm (GA). The GA-based optimization method was used to determine whether the parallel principal component analysis model (P-PCA) and parallel KPCA models (P-KPCA) contain variables, and the proposed method can

effectively handle nonlinear processes. In addition, some studies have adopted layering, subspace division, and Bayesian decision fusion mechanisms for process monitoring, and the monitoring performance has been improved to a certain extent [23–25].

Although the aforementioned research has been successful, the detection of faults in complex industrial processes where multiple features coexist remains a challenge. Complex industrial process variables always present high-dimensional, linear, nonlinear, Gaussian, and non-Gaussian coexistence. The proposed hybrid model provides a new direction for complex industrial process fault detection; however, the lack of in-depth research on nonlinear evaluation methods for process variables has led to the poor performance of existing hybrid models for complex industrial process fault detection with linear, nonlinear, Gaussian, and non-Gaussian feature coexistence. For example, the PCA-ICA model has good performance in industrial process monitoring where linear, Gaussian, and non-Gaussian features coexist, but there are always false positives and false negatives in nonlinear industrial processes. Therefore, it is important for theoretical research and engineering practice to carry out nonlinear evaluation of process variables and to build models for monitoring complex industrial processes.

In this study, a PCA-KPCA-ICA-KICA-BI mixed model is proposed for the coexistence of high-dimensional, linear, nonlinear, Gaussian, and non-Gaussian variables in complex industrial processes. First, a multivariate feature evaluation method is proposed. The original variables are classified using the Jarque–Bera test and the nonlinear discriminant method. This method calculates the kurtosis and skewness statistics of the variables. According to the normality test results, the process variables are divided into Gaussian blocks and non-Gaussian blocks; then we use Pearson correlation coefficient, maximum mutual information coefficient, and nonlinear evaluation function to divide Gaussian blocks and non-Gaussian blocks into Gaussian linear blocks, Gaussian nonlinear blocks, non-Gaussian linear blocks, and non-Gaussian blocks nonlinear blocks; then we apply the PCA-KPCA-ICA-KICA model to monitor the blocking process; and we finally propose a Bayesian inference fusion strategy to comprehensively decide the detection results of each block.

Section 2 briefly reviews PCA, ICA, kernel methods, and Jarque–Bera tests. Section 3 proposes nonlinear discriminant methods and Bayesian inference fusion strategies. Section 4 describes in detail the PCA-KPCA-ICA-KICA-BI method and its construction. Section 5 presents the application of PCA-KPCA-ICA-KICA-BI in the TE process and CSTR process and analyzes the corresponding monitoring performance. Section 6 concludes the paper.

2. Preliminaries

2.1. Principal Component Analysis and Kernel Principal Component Analysis. Principal component analysis constitutes one of the common data dimensionality reduction methods in linear correlation, Gaussian, and high-dimensional process monitoring. It maps industrial process data to

a low-dimensional space and then performs the analysis according to the data characteristics of the decomposed main subspace and residual subspace. For a given process dataset $X = [x_1, x_2, \ldots, x_m] \in R^{n \times m}$, n represents the number of samples and m represents the number of variables. After mapping, X is decomposed into

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

in which X is the normalized data matrix, $P \in R^{m \times k}$ is the loading matrix obtained by eigenvalue decomposition of the covariance matrix C, $T \in R^{n \times k}$ is the score matrix, E is the residual subspace, and k is the number of the principal components obtained from the cumulative percent variance (CPV):

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{m} \lambda_i} \times 100\% \geq 85\%, \tag{2}$$

in which λ_i is the eigenvalues of the covariance matrix C.

When a new sample $X_{new} = [x_1, x_2, x_3, \ldots, x_m]^T$ is obtained, according to the aforementioned decomposition model, the score vector t_{new} and residual subspace vector e of the new sample are as follows:

$$\left\{ t_{\text{new}} = x_{\text{new}}^T \widehat{P}, e = x_{\text{new}}^T \left(I - PP^T \right). \right. \tag{3}$$

In process monitoring, PCA algorithms typically adopt residual squared prediction error (SPE) and T^2 statistic to monitor operational status. T^2 statistic measures the PCA main subspace, and the SPE statistic considers the residual subspace. The loading matrix $\widehat{P}$ contains a lot of variance variation information. Since T^2 statistic is modeled based on a loading matrix $\widehat{P}$ with large singular values, it is sensitive to low singular value inaccuracies, which can be addressed by the SPE statistic [26, 27]. Similarly, the calculation formulas of the T^2 statistics and SPE statistics of the new sample x_{new} are as follows [26]:

$$\left\{ \begin{array}{l} T_{\text{PCA}}^2 = t_{\text{new}}^T \wedge^{-1} t_{\text{new}}, \\ \text{SPE}_{\text{PCA}} = e^T e, \end{array} \right. \tag{4}$$

in which $\wedge$ is the diagonal matrix containing the eigenvalues of the data variance matrix.

T^2 and SPE control limits are calculated from normal operating data. Since the T^2 statistic is in line with the F distribution, given a certain degree of confidence α (usually 99%), the statistical control limits T_α^2 and SPE_α are as follows:

$$\left\{ \begin{array}{l} T_\alpha^T = \frac{k(m^2 - 1)}{m(m-k)} F(k, m-k, \alpha), \\ \\ \text{SPE}_\alpha = \theta_1 \left[\frac{\eta_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1} \right]^{(1/h_0)}, \end{array} \right. \tag{5}$$

in which the freedom of $F(k, m-k, \alpha)$ is $m-k$; θ_i is the sum of the residual eigenvalues to the power of 1, 2, and 3; and h_0 is the intermediate outcome variable, $h_0 = 1 - (2\theta_1 \theta_3 / 3\theta_2^2)$.

2.2. Independent Component Analysis. Unlike PCA, ICA can monitor the non-Gaussian processes by extracting independent non-Gaussian features from industrial process variables. The given process data vector $X = [x_1, x_2, \ldots, x_m] \in R^{1 \times m}$ can be represented as a linear combination of statistically independent d non-Gaussian sources $s = [s_1, s_2, s_3, \ldots, s_d]$, in which $d \leq m$. After normalization and preprocessing, the relationship between the original industrial process data and the independent components ICs is expressed as follows:

$$x = As + e, \tag{6}$$

in which $A \in R^{m \times d}$ is the mixture matrix and $e \in R^{m \times 1}$ is the residual matrix. The basic problem of ICA is to estimate the mixture matrix A and the independent component vector s. Because the FastICA algorithm is characterized in fast calculation speed, this study uses the FastICA method [28] to calculate the decomposition matrix W and reconstruct $\widehat{s}$ as follows:

$$\widehat{s} = Wx. \tag{7}$$

To eliminate the correlation between the process data and facilitate the calculation, normalization and whitening are required. This paper adopts the commonly used PCA whitening method to whiten the process data. Through singular value decomposition, the obtained covariance matrix and whitened data are as follows:

$$\begin{array}{l} C = E\left(xx^T \right) = U \wedge U^T, \\ a = \wedge^{-(1/2)} U^T, \\ z = Qx = QAs, \end{array} \tag{8}$$

in which E represents the expectation, C is the covariance matrix, Q is the whitening matrix, and z is the whitened data. Orthogonal matrix B can be calculated by the following formula:

$$B = QA. \tag{9}$$

From the above formula, $\widehat{s}$ can be expressed as follows:

$$\widehat{s} = B^T z = B^T Qx. \tag{10}$$

The relationship between the decomposition matrix W and B can be expressed by the following formula:

$$w = B^T Q. \tag{11}$$

When ICs is determined, process monitoring can be performed by establishing I^2 and statistical magnitude SPE. The specific process is as follows:

$$\left\{ I^2 = \widehat{s}^T \widehat{s}, \text{SPE} = e^T e. \right. \tag{12}$$

The kernel density estimation (KDE) is used to calculate the thresholds of the above two statistics [28].

2.3. Kernel Method. KPCA and KICA models map nonlinear process variables into high-dimensional space for processing through kernel methods. This section introduces the kernel method at length since the monitoring schemes of PCA and ICA models have been explicated in Sections 2.1 and 2.2.

Assume the process data $X = [x_1, x_2, \ldots, x_m] \in R^{n \times m}$ is a normalized 0-mean dataset. Through nonlinear mapping $\Phi(\cdot)$, the covariance matrix C^F of the mapped data in the feature space is as follows [28]:

$$C^F = \frac{1}{N} \sum_{j=1}^{N} \varnothing(x) \varnothing(x)^T. \tag{13}$$

C^F can be diagonalized by eigenvalue decomposition as follows:

$$\lambda v = C^F v, \tag{14}$$

in which λ represents the eigenvalue and satisfies $\lambda \geq 0$, and v represents the eigenvector. Formulas (13) and (14) suggest the following:

$$C^F v = \left(\frac{1}{N} \sum_{j=1}^{n} \varnothing(x_j) \varnothing(x_j)^T \right) v = \frac{1}{N} \sum \langle \varnothing(x_j), v \rangle \varnothing(x_j). \tag{15}$$

Since the eigenvectors of the covariance matrix are spanned by samples, there must be a coefficient α_i ($i = 1, 2, 3, \ldots, n$):

$$v = \sum_{j=1}^{n} \alpha_i \varnothing(x_i). \tag{16}$$

Formulas (15) and (16) suggest the following:

$$\lambda \sum_{i=1}^{n} \alpha_i \langle \varnothing(x_k), \varnothing(x_i) \rangle = \frac{1}{n} \sum_{i=1}^{n} \alpha_i \langle \varnothing(x_k), \sum_{j=1}^{n} \varnothing(x_i) \varnothing(x_j), \varnothing(x_i) \rangle. \tag{17}$$

The inner product of the reconstructed variables in the feature space can be described by a kernel function. This study adopts a Gaussian kernel function, which is defined as follows:

$$kx, y = \varnothing(x), \varnothing(y) = \exp\left(\frac{x - y^2}{2 \times \sigma^2} \right). \tag{18}$$

The kernel matrix k is $[k]_{ij} = k_{ij} = \langle \Phi(x_i), \Phi(x_j) \rangle$ [28], and (17) can be further simplified as follows:

$$\lambda n k \alpha = k^2 \alpha, \tag{19}$$

in which α and λ are the eigenvalues and eigenvectors of the kernel matrix k. According to the above formula, λ and v can be calculated.

2.4. Jarque–Bera Test. MSPM method usually performs parameter evaluation and feature extraction given that variables obey normal distribution. Although it is reasonable to assume that some industrial processes obey normality, this assumption is often questionable when faced with complex industrial processes. Therefore, the Jarque–Bera (JB) test for normality distribution of complex process data is proposed. In statistics, the Jarque–Bera test examines whether the sample data shows the skewness and kurtosis of a normal distribution and goodness.

Assume $X = [x_1, x_2, x_3, \ldots, x_n]$ to be a dataset of n independent random variables; the sample skewness and kurtosis of $x_1, x_2, x_3, \ldots, x_n$ are calculated as follows:

$$\sqrt{b_1} = \frac{(1/n) \sum_{i=1}^{n} (x_i - \overline{x})^3}{\left[(1/n) \sum_{i=1}^{n} (x_i - \overline{x})^2 \right]^{(3/2)}},$$

$$b_2 = \frac{(1/n) \sum_{i=1}^{n} (x_i - \overline{x})^4}{\left[(1/n) \sum_{i=1}^{n} (x_i - \overline{x})^2 \right]^2}, \tag{20}$$

in which $\overline{x} = (1/n) \sum_{i=1}^{n} x_i$; if the process variable follows a Gaussian distribution, the skewness and kurtosis of X are close to 0 and 3, respectively. The normality of the sample can be tested by the skewness and kurtosis deviation of the expected value. The formula for calculating the JB statistic is as follows:

$$\text{JB} = \frac{n}{6} \left[\sqrt{b_1}^2 + \left(\frac{(b_2 - 3)^2}{4} \right) \right]. \tag{21}$$

The threshold JB_α of the JB statistical magnitude is calculated from the significance level α and the number of samples n. If $\text{JB} < \text{JB}_\alpha$, this means that the process variable obeys the normality assumption.

3. Nonlinear Discriminant Method and Bayesian Inference Fusion Strategy

3.1. Nonlinear Discriminant Method. It is of great significance to divide the linearly correlated variables and the nonlinearly correlated variables into two subblocks for separate monitoring since the complex industrial process variables may have both linear and nonlinear correlations. This section proposes a nonlinear discriminant method for this problem, which is described at length as follows.

Mutual information (MI), a nonlinear evaluation method based on communication entropy theory, quantitatively describes the nonlinear correlation between two random variables [29, 30]. If there is a strong correlation between these two variables, their MI values will become relatively large. On the contrary, if the two variables are approximately independent, the MI value will become very small. The proposal of mutual information provides great convenience for characterizing the relationship between two variables [30], but the computational complexity of the joint probability in mutual information is relatively large. Therefore, this study selects the maximum mutual information coefficient (MIC) with low computational complexity to indicate the degree of correlation between variables.

Although the maximum mutual information coefficient has a strong ability to represent linear and nonlinear correlations, it cannot clearly indicate whether the variables are linear or nonlinear. The Pearson correlation coefficient can

just make up for the lack of MIC. It is obtained by dividing the covariance of the two variables by the product of the standard deviation. The Pearson correlation coefficient value is located in $[-1, 1]$: a value equal to 1 indicates that the two variables are perfectly linearly correlated, a value equal to -1 indicates that the two variables are perfectly linearly negatively correlated, and a value equal to 0 indicates that there is no linear relationship between the two variables. The calculation process of the MIC and Pearson correlation coefficient is as follows:

$$I(x, y) = \int p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \, dx dy,$$

$$\text{MIC}(x, y) = \max_{a \times b < B} \frac{I(x, y)}{\log_2 \min(a, b)},$$

$$P_{x,y} = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)}\sqrt{\text{var}(y)}}$$

$$= \frac{E(xy) - E(x)E(y)}{\sqrt{E(x^2) - E^2(x)}\sqrt{E(y^2) - E^2(y)}},$$

(22)

in which $\text{MIC}(x, y)$ represents the maximum mutual information coefficient of the two variables, and $P_{x,y}$ represents the Pearson correlation coefficient between the two variables. $E(\cdot)$ is the expected value, and $\text{var}(\cdot)$ is the variance value.

Although MIC and Pearson correlation coefficients shine in characterizing the correlation between variables, deficiencies are seen in determining whether the variables are linearly or nonlinearly correlated. Therefore, we introduce a nonlinear discriminant coefficient, which combines the strengths of MIC and Pearson's correlation coefficient to determine whether a variable is nonlinear. Here is the calculation:

$$\text{NLDV} = \left(1 - |P_{x,y}|\right)(\alpha + (1 - \alpha)\text{MIC}(x, y)), \quad (23)$$

in which NLDV represents the value of nonlinear discrimination, which represents the degree of nonlinear correlation of process variables. $\text{NLDV} \in [0, 1]$ can be inferred from (23). If $\text{NLDV} = 1$ represents a complete nonlinear correlation between variables and $\text{NLDV} = 0$ represents a completely linear relationship between variables, to illustrate other cases of the value of NLDV, a threshold of $\gamma \in [0.3, 0.5]$ is set. If $\text{NLDV} > \gamma$ means that the variables are nonlinearly correlated, this paper takes $\gamma = 0.4$,

α represents the weight coefficient, and $\alpha = (1/2)$ in this paper [30].

3.2. Bayesian Inference Fusion Strategy.

Since the process variables are divided into multiple subspaces for separate monitoring, how to integrate the monitoring results becomes a problem. Reference [22] proposed a decision logic to determine the process running state. To improve the robustness of the model, this study adopts a Bayesian inference fusion strategy to determine the final monitoring results of the process. Bayesian inference is based on a probabilistic approach, which is similar to that in recent studies [31, 32]. Jarque–Bera and the nonlinear discriminant method divided the process dataset into Gaussian linear subspace, Gaussian nonlinear subspace, non-Gaussian linear subspace, and non-Gaussian nonlinear subspace. The corresponding monitoring statistics of each subspace are established. The failure probability of T^2 in a Gaussian linear subspace is defined as follows:

$$P_{T^2}(F|X_{\text{PCA}}) = \frac{P_{T^2}(X_{\text{PCA}}|F)P_{T^2 F}}{P_{T^2 X_{\text{PCA}}}},$$

$$P_{T^2 X_{\text{PCA}}} = P_{T^2}(X_{\text{PCA}}|N)P_{T^2 N} + P_{T^2}(X_{\text{PCA}}|F)P_{T^2 F},$$

(24)

in which N and F represent normal operating conditions and abnormal operating conditions, respectively. Given a confidence level α, $P_{T^2 F} = \alpha$, and $P_{T^2 N} = 1 - \alpha$. The calculation process of $P_{T^2}(X_{\text{PCA}}|N)$ and $P_{T^2}(X_{\text{PCA}}|F)$ is as follows:

$$P_{T^2}(X_{\text{PCA}}|N) = \exp\left(-\frac{T^2_{X_{\text{PCA}}}}{T^2_{\lim}}\right),$$

$$P_{T^2}(X_{\text{PCA}}|F) = \exp\left(-\frac{T^2_{\lim}}{T^2_{X_{\text{PCA}}}}\right),$$

(25)

in which $T^2_{\lim}$ is T^2 control limit.

Similarly, definition of the failure probability of the remaining three subspaces shares similarities with the calculation method of the Gaussian linear subspace given above. Combination of the statistical magnitude of T^2 and SPE of Gaussian linear subspace and Gaussian nonlinear subspace and that of non-Gaussian linear subspace and non-Gaussian nonlinear subspace is as follows:

$$\begin{cases} \text{BIC}_{T^2} = \dfrac{P^2_{T^2}(F|X_{\text{PCA}}) + P^2_{T^2}(F|X_{\text{KPCA}}) + P^2_{I^2}(F|X_{\text{ICA}}) + P^2_{I^2}(F|X_{\text{KICA}})}{P_{T^2}(F|X_{\text{PCA}}) + P_{T^2}(F|X_{\text{KPCA}}) + P_{I^2}(F|X_{\text{ICA}}) + P_{I^2}(F|X_{\text{KICA}})}, \\[3mm] \text{BIC}_{\text{SPE}} = \dfrac{P^2_{\text{SPE}}(F|X_{\text{PCA}}) + P^2_{\text{SPE}}(F|X_{\text{KPCA}}) + P^2_{\text{SPE}}(F|X_{\text{ICA}}) + PSPE(F|X_{\text{KICA}})}{P_{\text{SPE}}(F|X_{\text{PCA}}) + P_{\text{SPE}}(F|X_{\text{KPCA}}) + P_{\text{SPE}}(F|X_{\text{ICA}}) + P_{\text{SPE}}(F|X_{\text{KICA}})}. \end{cases}$$

(26)

Following weighted statistic formula is established according to the above formula to facilitate the monitoring:

$$
\begin{cases}
P^2 = P_{T^2}^2\left(F|X_{\mathrm{PCA}}\right) + P_{T^2}^2\left(F|X_{\mathrm{KPCA}}\right) + P_{I^2}^2\left(F|X_{\mathrm{ICA}}\right) + P_{I^2}^2\left(F|X_{\mathrm{KICA}}\right) + P_{\mathrm{SPE}}^2\left(F|X_{\mathrm{PCA}}\right) \\
\qquad + P_{\mathrm{SPE}}^2\left(F|X_{\mathrm{KPCA}}\right) + P_{\mathrm{SPE}}^2\left(F|X_{\mathrm{ICA}}\right) + P_{\mathrm{SPE}}^2\left(F|X_{\mathrm{KICA}}\right), \\[2mm]
P = P_{T^2}\left(F|X_{\mathrm{PCA}}\right) + P_{T^2}\left(F|X_{\mathrm{KPCA}}\right) + P_{I^2}\left(F|X_{\mathrm{ICA}}\right) + P_{I^2}\left(F|X_{\mathrm{KICA}}\right) + P_{\mathrm{SPE}}\left(F|X_{\mathrm{PCA}}\right) \\
\qquad + P_{\mathrm{SPE}}\left(F|X_{\mathrm{KPCA}}\right) + P_{\mathrm{SPE}}\left(F|X_{\mathrm{ICA}}\right) + P_{\mathrm{SPE}}\left(F|X_{\mathrm{KICA}}\right), \\[2mm]
\mathrm{BIC} = \dfrac{P^2}{P}.
\end{cases}
\tag{27}
$$

If the calculation result of the above formula exceeds the confidence level α, it indicates possible failures of the industrial process; otherwise, the industrial process is in a normal operating state.

4. Process Monitoring Program

The flowchart of PCA-KPCA-ICA-KICA-BI is shown in Figure 1, and the specific description is as follows.

4.1. Spatial Decomposition

Step 1 normalized the historical dataset X under normal operating conditions.

Step 2 calculated the value of each column of the dataset X.

Step 3 divided the dataset X into Gaussian subspace and non-Gaussian subspace by JB value and JB_α.

Step 4 divided the Gaussian subspace and the non-Gaussian subspace into Gaussian linear subspace, Gaussian nonlinear subspace, non-Gaussian linear subspace, and non-Gaussian nonlinear subspace through the nonlinear discrimination method proposed in Section 3.1.

4.2. Fault Detection

4.2.1. Offline Model. Step 1 adopted the historical dataset X under normal operating conditions for training; Step 2 performed space decomposition; Step 3 detected the Gaussian linear subspace, Gaussian nonlinear subspace, non-Gaussian linear subspace, and non-Gaussian nonlinear subspace making use of PCA, KPCA, ICA, and KICA models; Step 4 calculated the confidence limits for each subspace.

4.2.2. Online Monitoring

Step 1 collected the currently monitored industrial process system dataset X_{New} and normalized it.

Step 2 divided X_{New} into Gaussian linear subspaces, Gaussian nonlinear subspaces, non-Gaussian linear subspaces, and non-Gaussian nonlinear subspaces using spatial decomposition.

Step 3 calculated the monitoring statistics of each word space.

Step 4 learned the statistics in Step 3 and the confidence limits in Step 4 of the offline model. The final statistic BIC is calculated using the Bayesian inference fusion strategy of (27) in Section 3.2, where BIC > α indicates a fault, and the opposite indicates that no fault has occurred.

5. Case Study

To verify the effectiveness of the PCA-KPCA-ICA-KICA-BI hybrid model proposed in this study, two cases were adopted to verify the performance of the model.

5.1. Tennessee Eastman (TE) Chemical Process. The Tennessee Eastman (TE) chemical process described by Downs and Vogel consists of five main units: reactor, product condenser, vapor-liquid separator, recycle compressor, and product stripper. The reactor includes four reactive substances (A, C, D, and E) and an inert substance (B). After the reactants are fed into the reactor, products G and H and by-product F are generated through a series of chemical reactions. The product stream is condensed in a condenser and then separated by a vapor-liquid separator. The uncondensed product is sent back to the reactor by the centrifugal compressor for re-reaction, and the condensed product is sent to the stripper for stripping [33]. The flowchart of the TE chemical industry is shown in Figure 2. The TE process contains 41 measured variables (22 continuous process variables, 19 component measured variables) and 12 manipulated variables. The dataset adopted for this study was downloaded from https://web.mit.edu/braatzgroup/links.html. Table 1 presents all the variables included in the TE process, and the dataset contains 52 variables and 960
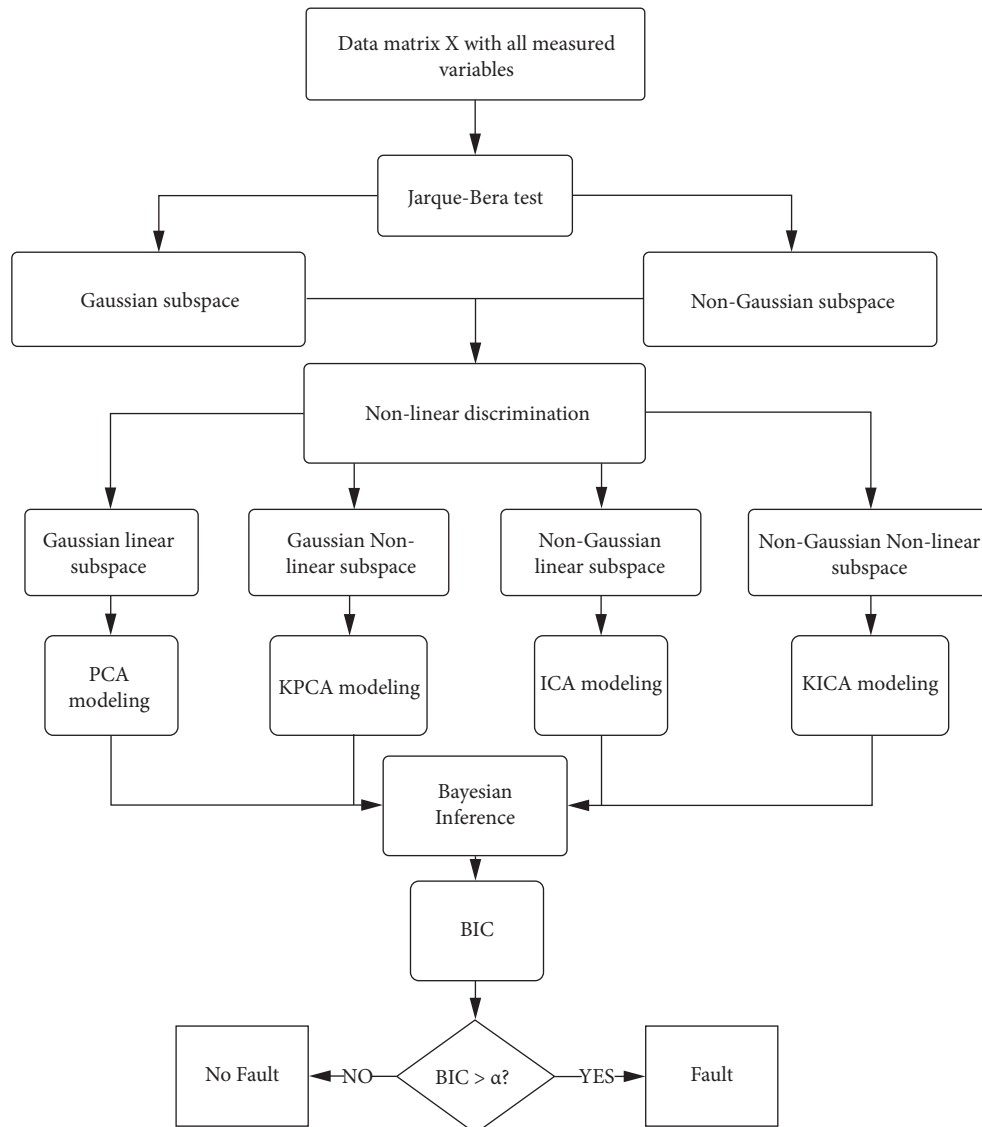
Figure 1: Flowchart for the PCA-KPCA-ICA-KICA-BI monitoring method.

samples. For the convenience of subsequent experimental analysis, Table 2 lists 21 different faults (faults are introduced after the 161st sample). In the following datasets of faults 11, 16, and 19, the PCA-KPCA-ICA-KICA-BI hybrid model proposed in this study is compared with the traditional single model to verify the effectiveness of the hybrid model.

5.1.1. Subspace Division. Based on the offline modeling in Section 4.2, the Jarque–Bera test divided the variable space into Gaussian subspaces and non-Gaussian subspaces, and the nonlinear discriminant method classified the variable space into Gaussian linear subspaces, Gaussian nonlinear subspaces, non-Gaussian linear subspace, and non-Gaussian nonlinear subspace. The results are shown in Table 3.

5.1.2. TE Process Fault Detection. Fault 11 is caused by random changes to reactor cooling water inlet temperature. PCA, KPCA, ICA, and KICA models are monitored in four

subspaces. Figure 3 shows the fault detection results of fault 11 by traditional single PCA, KPCA, and ICA models and the hybrid model proposed in this study. As shown in Figures 3(a)–3(c), traditional single PCA, KPCA, and ICA models can detect fault 11; the SPE and T^2 fault detection accuracy of the PCA model are 0.755 and 0.529, respectively; the SPE and T^2 fault detection accuracy of the KPCA model are 0.205 and 0.327, respectively; and the SPE and I^2 fault detection accuracy of the ICA model are 0.229 and 0.372, respectively. As shown in Figure 3(d), the hybrid model proposed in this study can detect more fault samples than the single models, and the fault detection accuracy is 0.924, far exceeding the number of failure detection samples for a traditional single model.

To further verify the efficient performance of the PCA-KPCA-ICA-KICA hybrid model, the unknown fault 16 was selected for testing. The fault detection results shown in Figure 4(d) revealed that the fault detection accuracy of the hybrid model reached a very ideal result, 0.928. Furthermore, tests on a single model were carried out to compare and
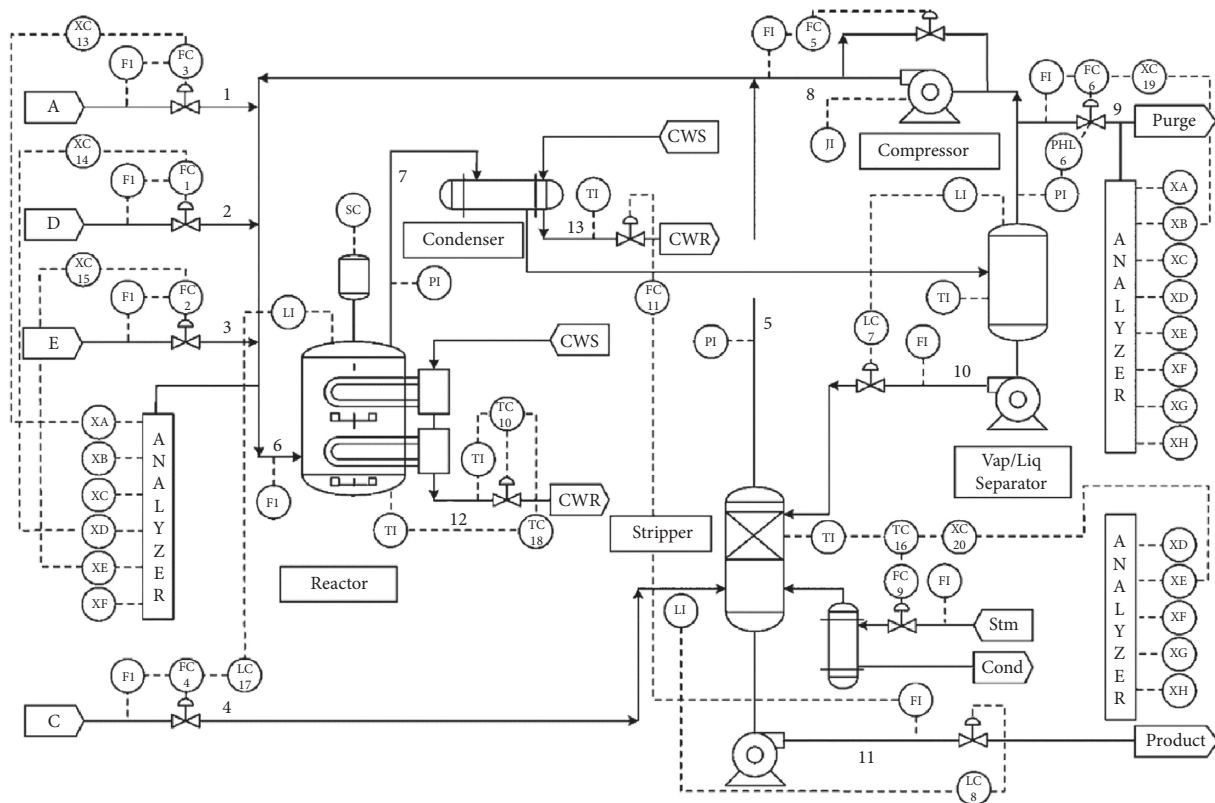
FIGURE 2: Control system of the TE process.

illustrate the superiority of this hybrid model. Figures 4(a)–4(c) show that the SPE and T^2 fault detection accuracy of the PCA model were 0.273 and 0.592, respectively; the SPE and T^2 fault detection accuracy of the KPCA model were 0.535 and 0.275, respectively; and the SPE and I^2 fault detection accuracy of the ICA model were 0.348 and 0.802, respectively.

The detection results of fault 11 and 16 in Figures 3 and 4 show that TE process variables have linear, nonlinear, Gaussian, and non-Gaussian distribution characteristics. The traditional single model only considers one distribution case and can only detect TE process faults caused by linearly distributed process variables, nonlinearly distributed process variables, or non-Gaussian distributed process variables, so the fault detection accuracy is relatively low. The hybrid model proposed in this study fully considers the multi-characteristic distribution of complex industrial process variables and divides the process variables into multiple subspaces using the JB test and the nonlinear evaluation method proposed in this study, and then a model with targeted distribution is used for each subspace for fault detection. Finally, the fault detection results are given using a Bayesian inference fusion strategy, and the results show that the proposed hybrid PCA-KPCA-ICA-KICA-BI model has higher fault detection accuracy for multifeatured complex industrial process fault detection.

The method proposed in this study considers the advantages of a single model and gives full play to the fault detection advantages of the single model using the JB test and nonlinear evaluation methods. Table 4 gives the TE

process 21 faults for the traditional PCA model, KPCA model, ICA model, PCA-ICA hybrid model, and PCA-KPCA-ICA-KICA-BI hybrid model proposed in this study. The proposed hybrid model and the existing PCA-ICA hybrid model have higher fault detection accuracy than the existing single PCA, KPCA, and ICA models, but the PCA-ICA hybrid model does not consider the distribution characteristics of the nonlinear process variables, and the decision logic of the hybrid model is insufficient, resulting in lower fault detection. The accuracy of the existing hybrid model is lower than that of the hybrid model proposed in this study. The existing single fault detection model, the PCA-ICA hybrid model, and the hybrid model proposed in this study have a fault detection accuracy close to 1 for faults 1, 2, 6, 7, 12, 14, 17, and 18 of the TE process; however, the fault detection accuracy of faults 3, 9, and 15 is not very satisfactory. This confirms that the magnitude and obviousness of faults can lead to a relatively large difference in fault detection accuracy. Therefore, to address the problem of small fault magnitude not being easily detected, due to the ability of exponentially weighted moving average (EWMA) control scheme to detect small changes, a strategy that combines the exponentially weighted sliding average control scheme with the PCA model is proposed to detect small faults in industrial processes [34]; how to detect initial faults or faults with a low signal-to-noise ratio will be a very interesting problem. Figures 3 and 4 show that a single model is inferior to the hybrid model in fault detection performance despite its ability to detect certain fault samples,

TABLE 1: Process variables in the TE process.

No.	Process measurements
1	A feed
2	D feed
3	E feed
4	Total feed
5	Recycle flow
6	Reactor feed rate
7	Reactor pressure
8	Reactor level
9	Reactor temperature
10	Purge rate
11	Product separator temperature
12	Product separator level
13	Product separator pressure
14	Product separator underflow
15	Stripper level
16	Stripper pressure
17	Stripper underflow
18	Stripper temperature
19	Stripper steam flow
20	Compressor work
21	Reactor cooling water outlet temperature
22	Separator cooling water outlet temperature
23	Composition A
24	Composition B
25	Composition C
26	Composition D
27	Composition E
28	Composition F
29	Composition A
30	Composition B
31	Composition C
32	Composition D
33	Composition E
34	Composition F
35	Composition G
36	Composition H
37	Composition D
38	Composition E
39	Composition F
40	Composition G
41	Composition H
42	D feed flow
43	C feed flow
44	A feed flow
45	Total feed flow valve
46	Compressor recycle valve
47	Purge valve
48	Separator pot liquid flow valve
49	Stripper liquid product flow valve
50	Stripper steam valve
51	Reactor cooling water flow
52	Condenser cooling water flow

TABLE 2: Faults of the TE process.

No.	Process variable	Type
1	A/C feed ratio, B composition constant	Step
2	B composition, A/C ratio constant	Step
3	D feed temperature	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	A feed loss	Step
7	C header pressure loss-reduced availability	Step
8	A, B, C feed composition	Random variation
9	D feed temperature	Random variation
10	C feed temperature	Random variation
11	Reactor cooling water inlet temperature	Random variation
12	Condenser cooling water inlet temperature	Random variation
13	Reaction kinetics	Slow drift
14	Reactor cooling water valve	Sticking
15	Condenser cooling water valve	Sticking
16	Unknown	Unknown
17	Unknown	Unknown
18	Unknown	Unknown
19	Unknown	Unknown
20	Unknown	Unknown
21	Valve position constant	Constant position

indicating its deficiency, that is, only considering one distribution of variables. Simply put, the hybrid model shines in dealing with the increasingly complex industrial processes.

TE process failure simulation experiments demonstrate the considerable advantages of the hybrid model proposed in this study. To further validate the performance of the model, the continuous stirred reactor process is used in Section 5.2.

5.2. CSTR Case. Figure 5 shows the simple flowchart of the CSTR process consisting of 9 variables. Here is the vector form of the dataset:

$$X = \left[x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8 x_9\right]^T = \left[T_C T_0 C_{AA} C_{AS} F_S F_C C_A T F_A\right]^T, \tag{28}$$

where T_C denotes the cooling water temperature, T_0 denotes the inlet temperature, C_{AA} and C_{AS} denote the inlet concentrations, F_S denotes the solvent flow, F_C denotes the cooling water flow, C_A denotes the outlet concentration, T denotes the temperature, and F_A denotes the reactant flow [22, 33].

In this study, using the same simulation conditions as Yoon and Macgregor [33], 960 normal operating CSTR data samples were collected through the established CSTR process simulation platform as the training set, and another 160 normal samples and 800 fault samples were collected as the test set. The test set was used to verify the fault detection performance of the hybrid model proposed in this study.

5.2.1. Subspace Division. In view of the offline modeling in Section 4.2, first, Jarque–Bera test divided the CSTR process variable space into Gaussian subspaces and non-Gaussian subspaces, and the nonlinear discriminant method classified the variable space into Gaussian linear subspaces, Gaussian

TABLE 3: Variable division in each subspace.

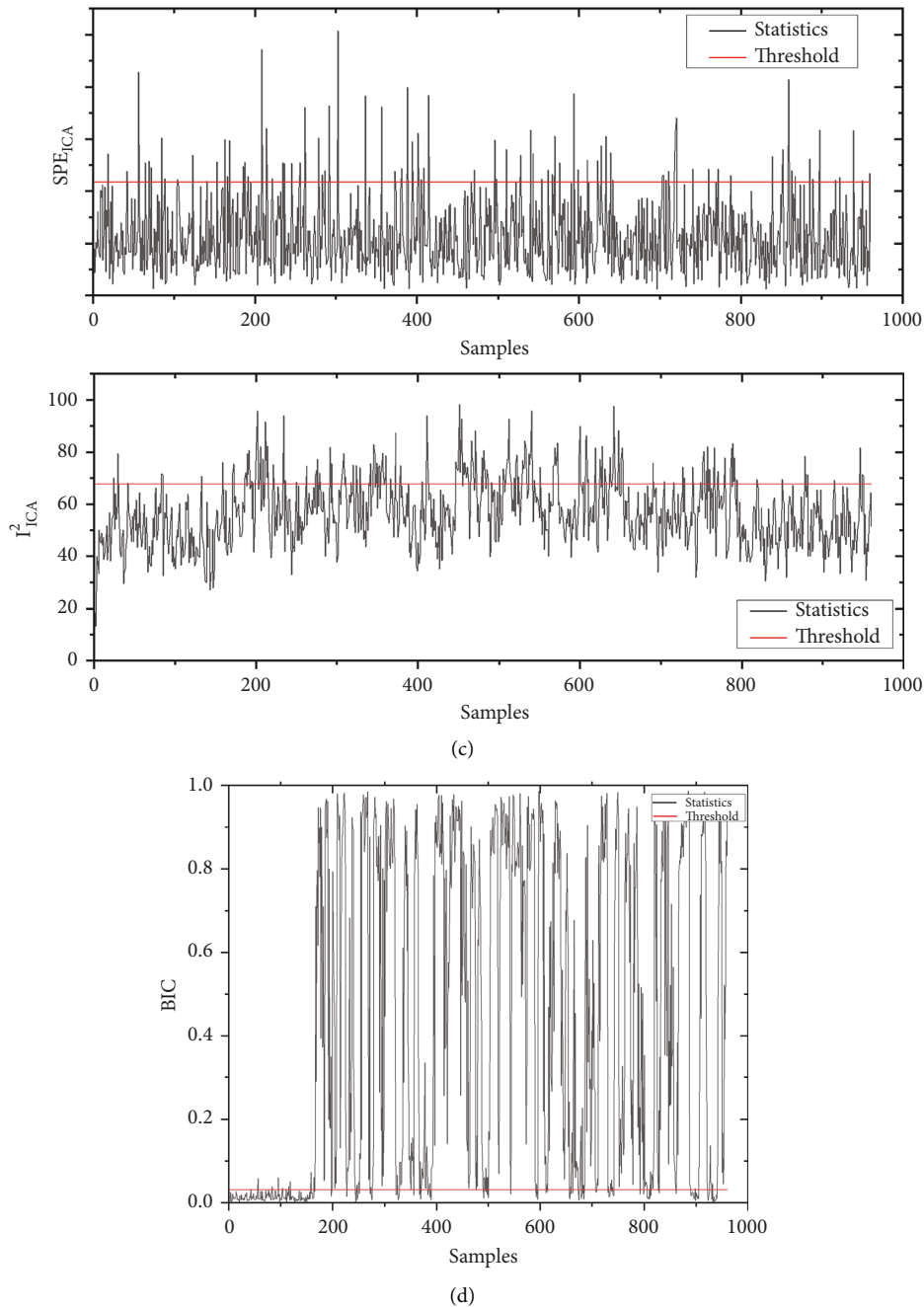| Subspace	Variable no.
Gaussian linear subspace	1, 9, 11, 12, 15, 17, 20, 22, 29, 44, 48, 49, 51, 52
Gaussian nonlinear subspace	4, 5, 6, 8, 14, 21, 23, 26, 27, 24, 25, 28, 30, 35, 37, 38, 41, 42, 43, 45
Non-Gaussian linear subspace	7, 13, 16, 18, 19, 50
Non-Gaussian nonlinear subspace	2, 3, 10, 31, 32, 33, 34, 36, 39, 40, 46, 47
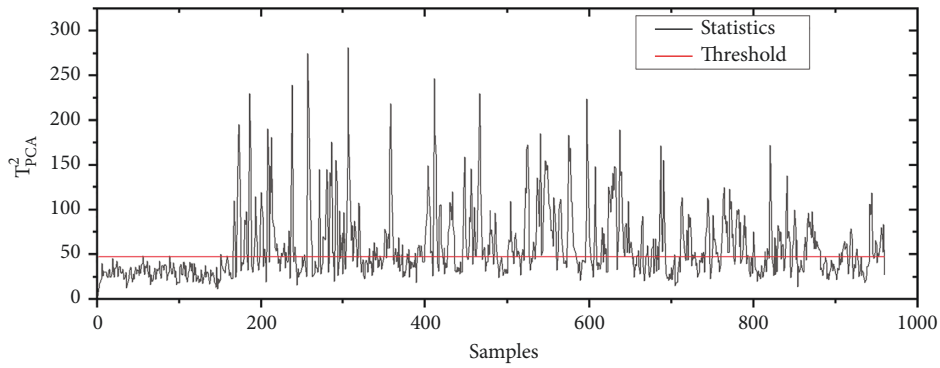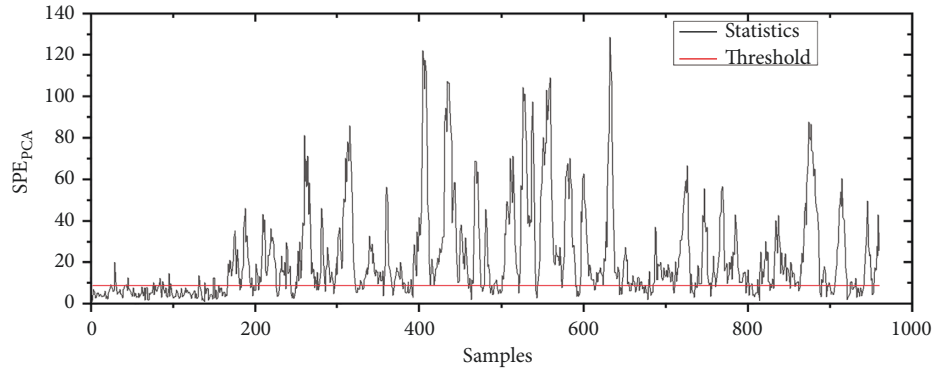
(a)
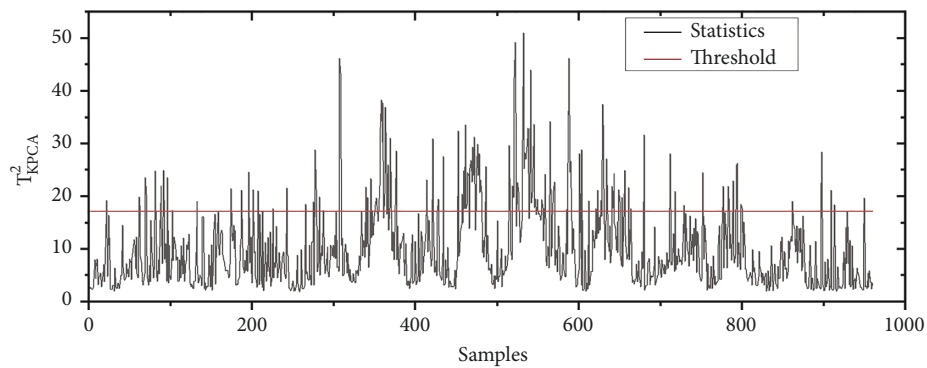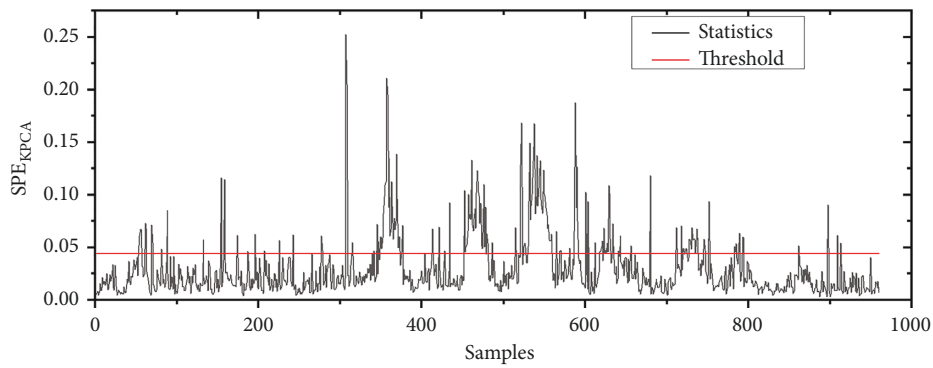
(b)

FIGURE 3: Continued.

Figure 3: Monitoring results of fault 11 in the TE process: (a) PCA; (b) KPCA; (c) ICA; (d) PCA-KPCA-ICA-KICA-BI.

nonlinear subspaces, non-Gaussian linear subspace, and non-Gaussian nonlinear subspace, as shown in Table 5.

5.2.2. CSTR Process Fault Detection. Figure 6 shows the fault detection results of the CSTR process. Figures 6(a)–6(d) show that the hybrid model proposed in this paper integrates the distribution characteristics of the process variables and uses the BI method to make a comprehensive decision on the fault detection results with significant advantages, where the fault detection rate of the PCA-KPCA-ICA-KICA-BI hybrid model is 1 and is lower than the FPR of the single model.

Since the CSTR process is a typical nonlinear process [22, 33], the PCA model and the ICA model cannot handle the nonlinear distributed characteristic variables well, resulting in lower fault detection rates, with 0.83 for the PCA model and 0.08 for the ICA model. KPCA can map the nonlinear data to a high-dimensional space to become linear, so the KPCA model has good applicability in nonlinear process fault detection, which is verified in Figure 6(b). Comparison of Figures 6(b) and 6(d) shows that the fault detection rate of the KPCA model is comparable to that of the hybrid model proposed in this paper, which indicates that full consideration of the distribution of process
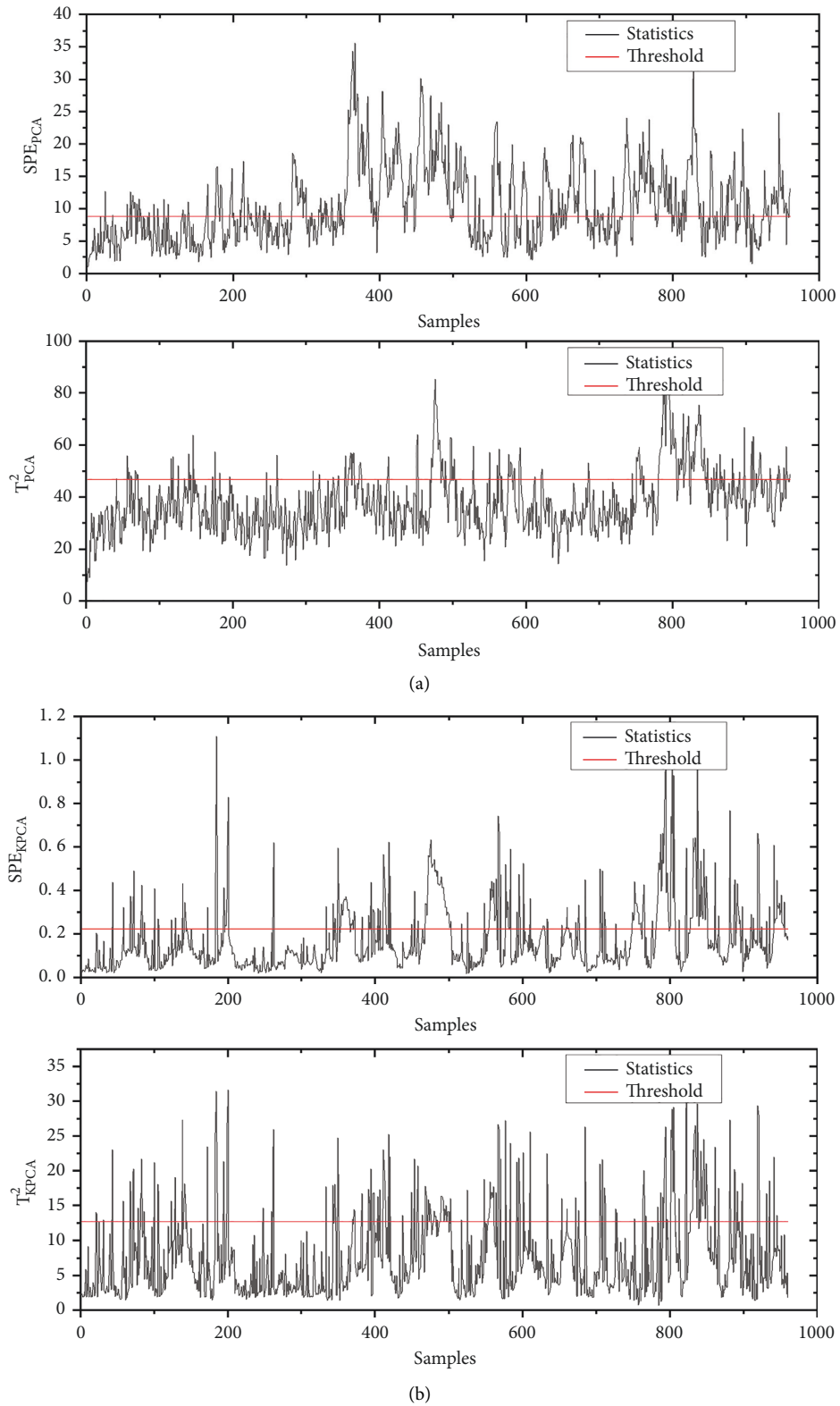
(a)

(b)
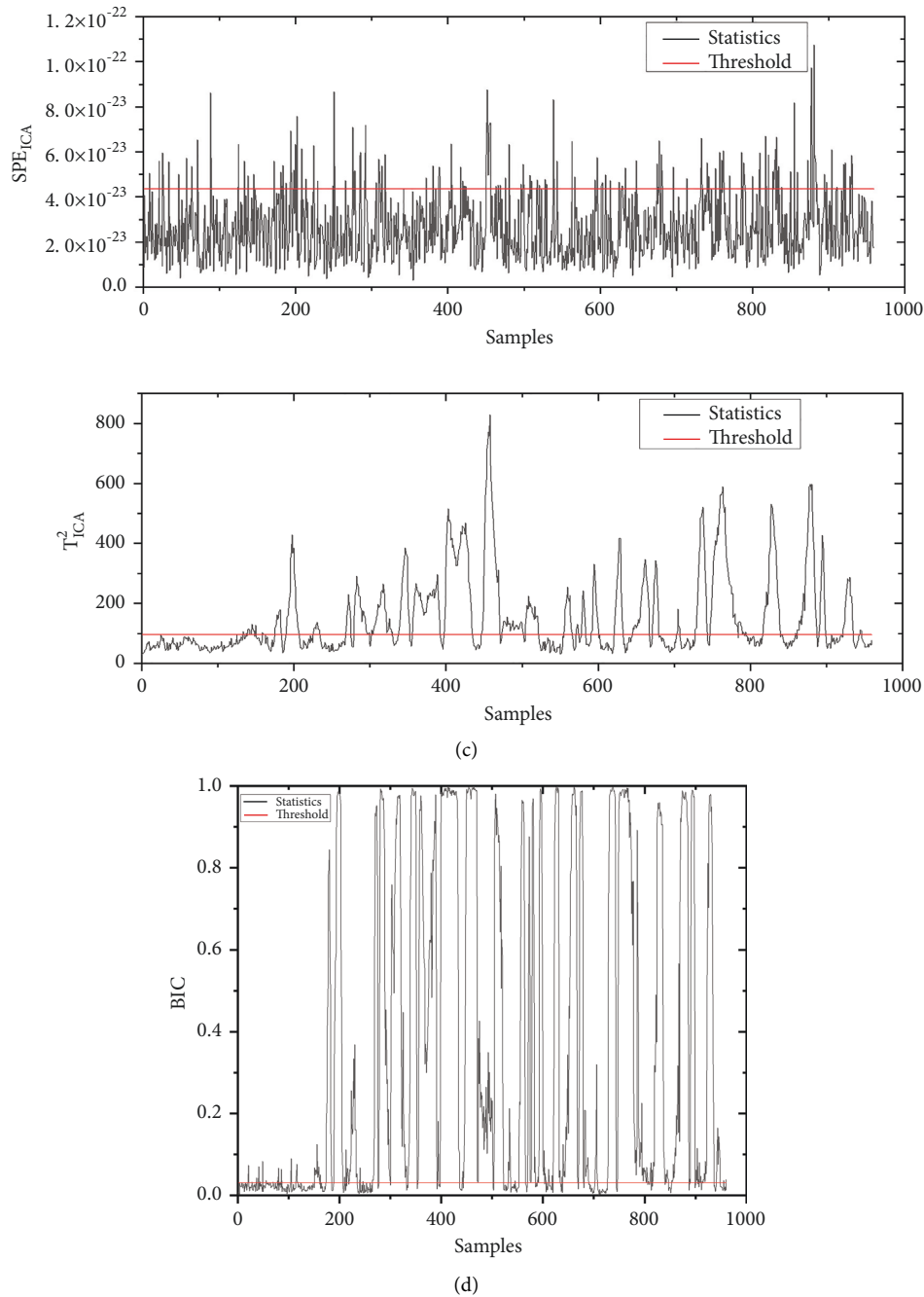
Figure 4: Continued.

(c)

(d)

Figure 4: Monitoring results of fault 16 in the TE process: (a) PCA; (b) KPCA; (c) ICA; (d) PCA-KPCA-ICA-KICA-BI.

variables and the use of a targeted fault detection model is helpful in improving the fault detection accuracy. The fault detection results of the CSTR process show that the hybrid PCA-KPCA-ICA-KICA-BI model can handle complex multi-characteristic process variables well and improve the fault detection accuracy of complex industrial processes.

TABLE 4: Detection rates of 21 faults.

Fault no.	PCA		KPCA		ICA		PCA-ICA		Proposed
	T^2	SPE	T^2	SPE	I^2	SPE	TI^2_{cs}	SPE_{cs}	BIC
1	0.992	0.991	0.997	0.993	0.997	0.997	0.994	0.992	0.995
2	0.985	0.979	0.985	0.985	0.986	0.986	0.986	0.978	0.988
3	0.009	0.007	0.027	0.070	0.162	0.176	0.17	0.212	0.336
4	0.493	0.941	0.975	0.977	0.775	0.873	0.401	0.991	1
5	0.275	0.294	0.378	0.496	1	1	0.352	0.425	0.724
6	0.992	1	0.998	1	1	1	0.992	0.998	1
7	1	1	1	1	1	1	1	1	1
8	0.976	0.976	0.983	0.984	0.969	0.974	0.977	0.974	0.978
9	0.137	0.182	0.187	0.225	0.112	0.132	0.175	0.2	0.282
10	0.334	0.341	0.501	0.697	0.749	0.762	0.594	0.702	0.895
11	0.592	0.796	0.327	0.205	0.372	0.229	0.403	0.772	0.924
12	0.975	0.982	0.983	0.982	0.997	0.997	0.992	0.965	0.995
13	0.944	0.952	0.959	0.959	0.954	0.955	0.953	0.95	0.962
14	1	0.988	0.999	0.996	1	0.999	0.973	0.997	1
15	0.171	0.247	0.228	0.289	0.041	0.152	0.174	0.224	0.302
16	0.592	0.273	0.275	0.535	0.802	0.348	0.682	0.734	0.928
17	0.752	0.893	0.937	0.942	0.862	0.894	0.785	0.952	0.979
18	0.896	0.914	0.913	0.917	0.893	0.896	0.91	0.918	0.930
19	0.175	0.293	0.179	0.394	0.598	0.374	0.397	0.654	0.984
20	0.335	0.598	0.547	0.635	0.758	0.733	0.736	0.742	0.872
21	0.482	0.599	0.489	0.562	0.695	0.647	0.682	0.704	0.825

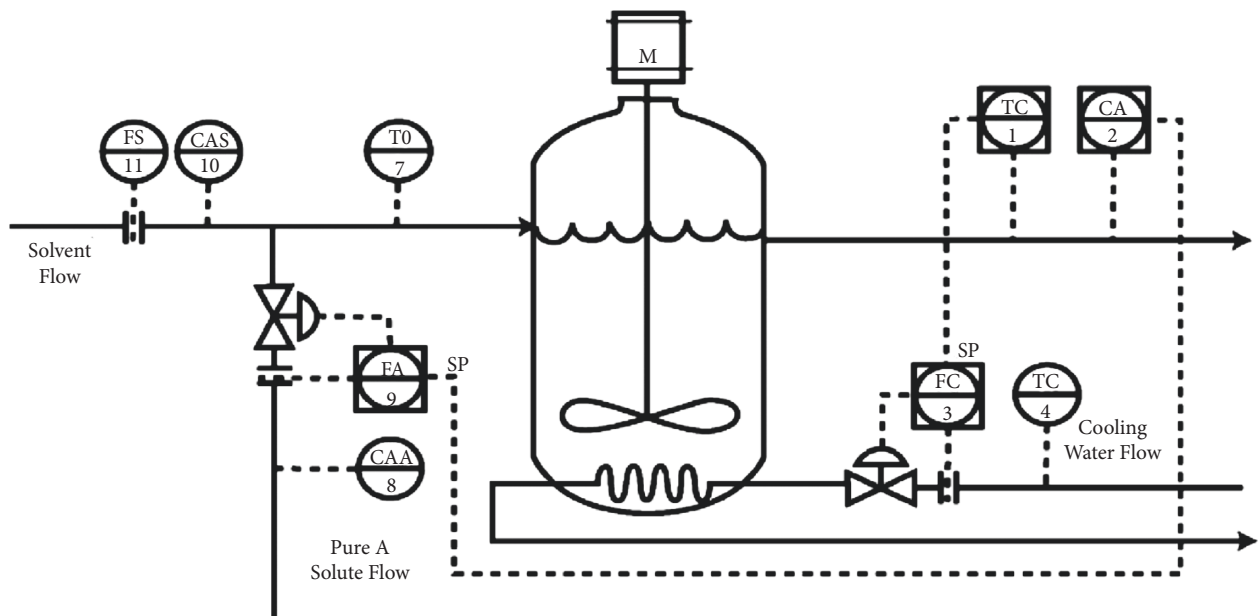FIGURE 5: Simplified schematic of the CSTR [22, 33].

TABLE 5: CSTR process variable division in each subspace.

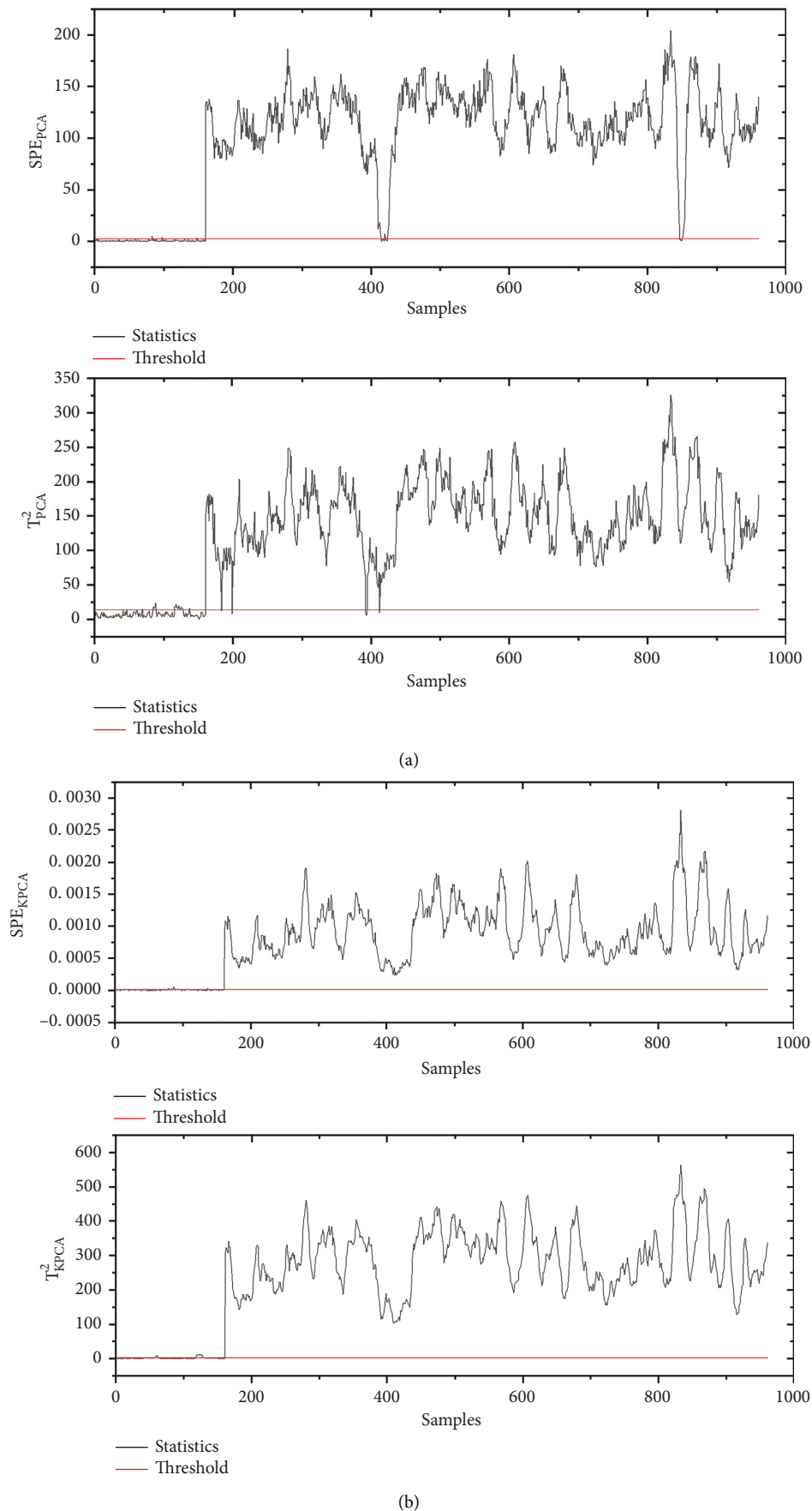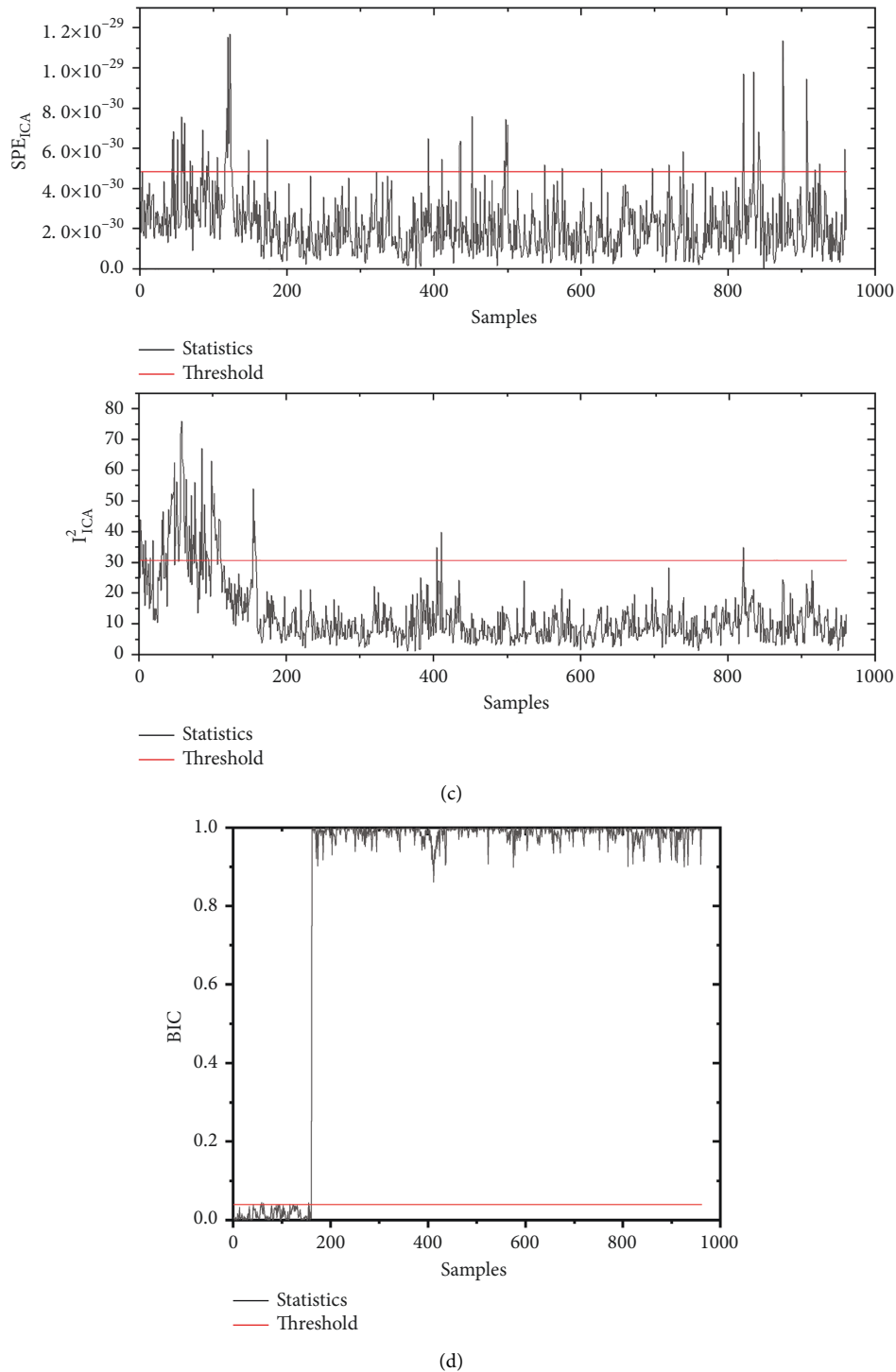| Subspace	Variable no.
Gaussian linear subspace	1, 2, 8
Gaussian nonlinear subspace	3, 4, 5, 6, 7
Non-Gaussian linear subspace	9
Non-Gaussian nonlinear subspace	

(a)

(b)

FIGURE 6: Continued.

FIGURE 6: Monitoring results of faults in the CSTR process: (a) PCA; (b) KPCA; (c) ICA; (d) PCA-KPCA-ICA-KICA-BI.

6. Conclusion

A hybrid fault detection method based on principal component analysis, kernel principal component analysis, independent component analysis, kernel independent component analysis, and Bayesian inference (PCA-KPCA-ICA-KICA-BI) is proposed for complex industrial processes. First, the Jarque–Bera test and the nonlinear discriminant method divided the feature variables into Gaussian linear blocks, Gaussian nonlinear blocks, non-Gaussian linear blocks, and non-Gaussian nonlinear blocks. Each division was monitored by the PCA-KPCA-ICA-KICA-BI model, and the corresponding results were given comprehensively by the Bayesian inference. Finally, the efficient performance is verified by the Tennessee Eastman (TE) and Continuous Stirred Tank Reactor (CSTR) processes. The traditional

single model has a fault detection accuracy of 0.68, and the hybrid PCA-ICA model has a detection accuracy of 0.72, which means that more faults are considered normal, while the hybrid model proposed in this paper has a fault detection accuracy of 0.86, which is a significant improvement.

Complex industrial processes are characterized by strong coupling of variables, and any small fault can cause huge safety hazards, which makes it difficult to collect data on actual industrial process faults. Therefore, this study used TE and CSTR processes with complex characteristics, such as high-dimensional, linear, nonlinear, Gaussian, and non-Gaussian features, to verify the fault detection performance of the hybrid model proposed in this paper. The analysis of the fault detection results shows that the PCA-KPCA-ICA-KICA-BI hybrid model can fully consider the variable distribution characteristics of TE and CSTR processes and has a higher fault detection accuracy than the traditional single models and the PCA-ICA hybrid model. The application of the hybrid model proposed in this study to the TE process and the CSTR process shows that the hybrid model has good prospects for securing complex industrial processes. To further optimize and improve the PCA-KPCA-ICA-KICA-BI hybrid model and to enhance the applicability of the model, further research directions could be suggested in the following areas:

(1) Initial failures with insignificant changes in complex industrial process data could be detected.

(2) Compound failures could be detected.

(3) Most current fault detection models require a large number of data samples. However, in the pursuit of a safe and sound industrial process, intriguing is how to realize fault detection based on a small number of samples.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] L. H. Chiang, E. L. Russell, and R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*, Springer, London, UK, 2001.

[2] R. Isermann, "Model-based fault detection and diagnosis: status and applications," *IFAC Proceedings Volumes*, vol. 37, no. 6, pp. 49–60, 2004.

[3] T. Kourti and J. F. Macgregor, "Process analysis, monitoring and diagnosis, using multivariate projection methods," *Chemometrics and Intelligent Laboratory Systems*, vol. 28, pp. 3–21, 1995.

[4] X. Dai and Z. Gao, "From model, signal to knowledge: a data-driven perspective of fault detection and diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 9, pp. 2226–2238, 2013.

[5] P. Zhou, W. Li, W. Hong, M. Li, and T. Chai, "Robust online sequential RVFLNs for data modeling of dynamic time-varying systems with application of an ironmaking blast furnace," *IEEE Transactions on Cybernetics*, vol. 99, pp. 1–13, 2019.

[6] R. Fazai, M. Mansouri, K. Abodayeh, H. Nounou, and M. Nounou, "Online reduced kernel PLS combined with GLRT for fault detection in chemical systems," *Process Safety and Environmental Protection*, vol. 128, pp. 228–243, 2019.

[7] V. Venkatasubramanian, R. Rengaswamy, and S. N. Kavuri, "A review of process fault detection and diagnosis," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 313–326, 2003.

[8] Y. Meng, Q. Lan, J. Qin, S. Yu, H. Pang, and K. Zheng, "Data-driven soft sensor modeling based on twin support vector regression for cane sugar crystallization," *Journal of Food Engineering*, vol. 241, pp. 159–165, 2019.

[9] J. F. MacGregor, C. Jaeckle, C. Kiparissides, and M. Koutoudi, "Process monitoring and diagnosis by multiblock PLS methods," *AICHE Journal*, vol. 40, pp. 826–838, 1994.

[10] R. Muradore and P. Fiorini, "A PLS-based statistical approach for fault detection and isolation of robotic manipulators," *IEEE Transactions on Industrial Electronics*, vol. 59, pp. 3167–3175, 2012.

[11] M. Kano, S. Tanaka, S. Hasebe, I. Hashimoto, and H. Ohno, "Monitoring independent components for fault detection," *AICHE Journal*, vol. 49, no. 4, pp. 969–976, 2003.

[12] Z. Chen, X. D. Steven, Z. Kai, Z. Li, and Z. Hu, "Canonical correlation analysis-based fault detection methods with application to alumina evaporation process," *Control Engineering Practice*, vol. 46, pp. 51–58, 2016.

[13] F. Harrou, M. Madakyaru, Y. Sun, and S. Khadraoui, "Improved detection of incipient anomalies via multivariate memory monitoring charts: application to an air flow heating system," *Applied Thermal Engineering*, vol. 109, pp. 65–74, 2016.

[14] M. Said and O. Taouali, "Improved dynamic optimized kernel partial least squares for nonlinear process fault detection," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6677944, 16 pages, 2021.

[15] M. Kallas, G. Mourot, K. Anani, J. Ragot, and D. Maquin, "Fault detection and estimation using kernel principal component analysis," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1025–1030, 2017.

[16] J. Huang and X. Yan, "Quality relevant and independent two block monitoring based on mutual information and KPCA," *IEEE Transactions on Industrial Electronics*, vol. 64, pp. 6518–6527, 2017.

[17] J.-M. Lee, S. J. Qin, and I.-B. Lee, "Fault detection of nonlinear processes using kernel independent component analysis," *Canadian Journal of Chemical Engineering*, vol. 85, pp. 526–536, 2008.

[18] X. Yuan, S.-Q. Shen, Y.-L. He, and Q.-X. Zhu, "A novel hybrid method integrating ICA-PCA with relevant vector machine for multivariate process monitoring," *IEEE Transactions on Control Systems Technology*, vol. 27, pp. 1–8, 2018.

[19] J. Huang and X. Yan, "Gaussian and non-gaussian double subspace statistical process monitoring based on principal

component analysis and independent component analysis," *Industrial & Engineering Chemistry Research*, vol. 54, 2014.

[20] S. Li, X. Zhou, H. Shi, and F. Pan, "Dynamic non-gaussian hybrid serial modeling for industrial process monitoring," *Chemometrics and Intelligent Laboratory Systems*, vol. 216, Article ID 104371, 2021.

[21] X. Deng, X. Tian, S. Chen, and C. J. Harris, "Nonlinear process fault diagnosis based on serial principal component analysis," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, pp. 560–572, 2018.

[22] Q. Jiang and X. Yan, "Parallel PCA-KPCA for nonlinear process monitoring," *Control Engineering Practice*, vol. 80, pp. 17–25, 2018.

[23] W. Li, C. Zhao, and F. Gao, "Linearity evaluation and variable subset partition based hierarchical process modeling and monitoring," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2683–2692, 2018.

[24] J. Zhu, Z. Ge, Z. Song, L. Zhou, and G. Chen, "Large-scale plant-wide process modeling and hierarchical monitoring: a distributed bayesian network approach," *Journal of Process Control*, vol. 65, pp. 91–106, 2018.

[25] J. Zhu, Z. Ge, and Z. Song, "Distributed parallel PCA for modeling and monitoring of large-scale plant-wide processes with big data," *IEEE Transactions on Industrial Informatics*, vol. 99, 2017.

[26] M. Adil, M. Abid, A. Q. Khan, and G. Mustafa, "Comparison of PCA and FDA for monitoring of coupled liquid tank system," in *Proceedings of the: 2016 13th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, Islamabad, Pakistan, January 2016.

[27] S. Roweis, "Em algorithms for pca and spca," *Advances in Neural Information Processing Systems*, vol. 10, 1999.

[28] J. M. Lee, C. Yoo, and I. B. Lee, "Statistical process monitoring with independent component analysis," *Journal of Process Control*, vol. 14, no. 5, pp. 467–485, 2004.

[29] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Physical Review*, vol. 69, no. 6, Article ID 066138, 2004.

[30] S. Zhang, F. Wang, L. Zhao, S. Wang, and Y. Chang, "A novel strategy of the data characteristics test for selecting a process monitoring method automatically," *Industrial & Engineering Chemistry Research*, vol. 55, pp. 1642–1654, 2016.

[31] Z. Ge, M. Zhang, and Z. Song, "Nonlinear process monitoring based on linear subspace and Bayesian inference," *Journal of Process Control*, vol. 20, pp. 676–688, 2010.

[32] H. Jian, K. E. Okan, and X. Yan, "Slow feature analysis based on online feature reordering and feature selection for dynamic chemical process monitoring—ScienceDirect," *Chemometrics and Intelligent Laboratory Systems*, vol. 169, pp. 1–11, 2017.

[33] S. Yoon and J. F. Macgregor, "Fault diagnosis with multivariate statistical models part I: using steady state fault signatures," *Journal of Process Control*, vol. 11, pp. 387–400, 2001.

[34] F. Harrou, M. Nounou, and H. Nounou, "A statistical fault detection strategy using PCA based EWMA control schemes," in *Proceedings of the 2013 9th Asian Control Conference (ASCC)*, Istanbul, Turkey, June 2013.