ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Nonlinear Dynamic Process Monitoring Based on Ensemble Kernel Canonical Variate Analysis and Bayesian Inference

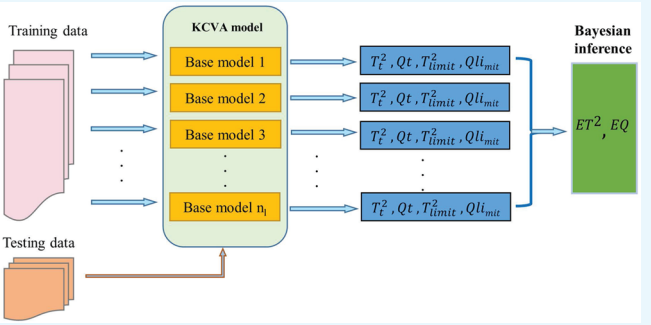
Xuemei Wang and Ping Wu*

Cite This: *ACS Omega* 2022, 7, 18904−18921

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: By considering autocorrelation among process data, canonical variate analysis (CVA) can noticeably enhance fault detection performance. To monitor nonlinear dynamic processes, a kernel CVA (KCVA) model was developed by performing CVA in the kernel space generated by kernel principal component analysis (KPCA). The Gaussian kernel is widely adopted in KPCA for nonlinear process monitoring. In Gaussian kernel-based process monitoring, a single learner is represented by a certain selected kernel bandwidth. However, the selection of kernel bandwidth plays a pivotal role in the performance of process monitoring. Usually, the kernel bandwidth is determined manually. In this paper, a novel ensemble kernel canonical variate analysis (EKCVA) method is developed by integrating ensemble learning and kernel canonical variate analysis. Compared to a single learner, the ensemble learning method usually achieves greatly superior generalization performance through the combination of multiple base learners. Inspired by the ensemble learning method, KCVA models are established by using different kernel bandwidths. Further, two widely used T^2 and Q monitoring statistics are constructed for each model. To improve process monitoring performance, these statistics are combined through Bayesian inference. A numerical example and two industrial benchmarks, the continuous stirred-tank reactor process and the Tennessee Eastman process, are carried out to demonstrate the superiority of the proposed method.



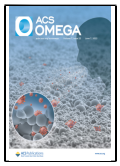
1. INTRODUCTION

Recent developments in the field of process monitoring have led to a renewed interest in ensuring operation safety in the process industry.[1−4] Model-based and data-driven methods are the two sorts of process monitoring strategies in general. Accurate process models, which are typically difficult to generate for complicated industrial processes, are required for model-based techniques. Data-driven methods, unlike model-based methods, rely solely on process data. A massive amount of process data can be captured, stored, and analyzed because of the rapid advancement of information, communication, and computing technology. As a result, data-driven process monitoring has triggered an amount of innovative inquiry.[5−8]

The multivariate analysis (MVA) approach has been widely used and explored among data-driven process monitoring methods for its simplicity and adaptability.[9,10] Representative MVA methods include principal component analysis (PCA), partial least-squares (PLS), and canonical variate analysis (CVA). In PCA, features are extracted from process data by maximizing the variance in the projection space. To address the shortcomings of the standard PCA, Rohoma et al.[11] suggested two methods for evaluating sparse principal components. These two approaches surpass PCA in threshold determination because they use resampling and measurement error covariance matrices to obtain the distribution of the loading components. Furthermore, Yu et al. suggested a nonlinear, robust, and spare PCA by using *Spearman's* and *Kendall's* correlation matrices.[12] PLS is usually employed in quality-related process monitoring. PLS concentrates on maximizing the covariance between input and output variables. However, many theoretical researchers have found that these two methods do not satisfy these conditions of temporal correlation and autocorrelation of most industrial data. To settle this issue, PCA and PLS have been extended to dynamic versions by using time-lagged process variables.[13,14] In addition, the construction of a graphical framework based on the PCA algorithm using a Bayesian network can also realize the dynamic connection between data.[15]

Different from PCA and PLS, as a variant of canonical correlation analysis (CCA), canonical variate analysis (CVA)

Received: March 28, 2022
Accepted: May 13, 2022
Published: May 24, 2022



attempts to extract the state-space from process variables. Hence, CVA is more suitable for monitoring dynamic processes. More recently, CVA-based process monitoring methods have attracted much attention in process monitoring.[16,17] Juricek et al. used a local approach to detect changes in the canonical variable coefficients which are extracted from the built CVA model.[18] Odiowei and Cao[19] constructed a minimum state-space model using CVA to derive the state variable and residual canonical variates, and employed kernel density estimations (KDE) to establish the upper control limits to deal with the nonlinearity and non-Gaussian issues for process monitoring. Jiang and Braatz[20] improved the interpretability of fault detection by using correlation features that provide a more efficient representation of process features and reduce redundancy to space compared to raw data. Lu et al.[21] employed a sparse CVA (SCVA) to generate sparse canonical variates to explore the major structures among process variables for process monitoring. A new metric to quantify the dissimilarity of past and future data was proposed by Pilario and Cao.[22] The new index is more reliable than the CVA T^2 and Q index for detecting incipient faults. Li et al.[23] pointed out that taking into account the residuals of the canonical variant produces a more reliable and interpretable contribution plot than contribution maps based on T^2 and Q statistics from the CVA for detecting faults that evolve slowly. Zheng and Zhao[24] combined CVA and slow feature analysis (SFA) to achieve the separation of process static and dynamic information, which better explains the process dynamics of the system.

However, the above CVA-based process monitoring methods would fail to detect faults efficiently for nonlinear processes owing to the assumption that the process variables are linear related. To address this issue, kernel methods have been widely adopted to extend linear MVA methods to nonlinear MVA methods.[25] Ciabatton et al. combined kernel canonical variate analysis (KCVA) and kernel density estimation (KDE) to develop a nonlinear monitoring method for detecting faults and occupant bad behaviors for a residential microgrid application.[26] Xiao et al. integrated locality preserving projections into KCVA to retain local manifold structure in kernel space for fault detection.[27] In the research of Shang et al.,[28] the exponential weighted moving average approach is used to update the covariance in the feature space after changing the operating conditions. The first-order perturbation theory is used to decrease the calculation of Hankel matrix singular value decomposition in high-dimensional eigenspace. To reduce the storage and computational expense, Yu et al. presented accelerated KCVA for fault isolation by applying full rank factorization incorporated in kernel matrix approximation.[29] A limitation of applying the KCVA to nonlinear dynamic systems is that a singular kernel matrix may be involved, particularly for industrial processes, since the process variables are often colinear. To solve this issue, Samuel and Cao[30] performed CVA in the kernel space created by KPCA. The singularity problem can be avoided owing to the compression of KPCA.

Only a single Gaussian kernel function is used in kernel-based process monitoring approaches such as KPCA and KCVA. The width parameters, on the other hand, are calculated empirically. Thus, the detection performance may be significantly degraded if an unfit Gaussian kernel function is selected while fault information is unavailable.[25] As a class of combinatorial optimization learning methods, the ensemble learning (EL) method can derive better performance by combing multiple simple models.[31] To enhance the process monitoring performance, Li and Yang, through the analysis of a nonlinear numerical example, pointed out that a single KPCA model can hardly meet the needs of the overall monitoring. Therefore, a new approach is proposed where multiple KPCA models are combined with a Bayesian inference strategy to obtain a new metric.[32] Cui et al.[33] further considered both global and local structure information on the original data and preserved the local information in the based KPCA models before performing ensemble learning.

Several nonlinear dynamic process monitoring approaches, such as DKPCA and KCVA, have recently been presented. The correlation relationship between process variables is investigated by KCVA. DKPCA, on the other hand, solely considers the variance information on process variables. The relationship between process variables is not taken into account. Yet, the kernel parameter is hard to determine. Moreover, only a single model is derived for one kernel parameter. Inspired by EDKPCA which combined ensemble learning and DKPCA, in this work, the KCVA is extended to ensemble KCVA by integrating the ensemble learning strategy into KCVA. Due to the merits of ensemble learning and KCVA, the dynamics and nonlinearity of complex industrial processes can be better presented for process monitoring. First, the KPCA models with different kernel parameters are established to extract the kernel principal components. Then, CVA is performed in each kernel principal component space, and T^2 and Q statistics are constructed. The corresponding thresholds are determined by KDE. For nonlinear dynamic process monitoring, a Bayesian combination strategy is employed to integrate these monitoring statistics to take advantage of these CVA models. Two widely employed case studies are presented to demonstrate the performance of the proposed EKCVA compared to other related methods.

The main contribution and novelty of this article lie in the following aspects:

- Ensemble learning approach is introduced to KCVA to improve the process monitoring performance. Through several base KPCA learners, the nonlinear characteristics of nonlinear dynamic processes are better described.
- Meanwhile, by taking the autocorrelation and cross-correlation into consideration, the proposed EKCVA method can present superior fault detection performance for nonlinear dynamic processes.
- A Bayesian inference strategy is adopted to define a single fault detection index to facilitate the fault detection logic.

The remainder of this work is organized as follows. Section 2 briefly recalls the KPCA and CVA methods. Section 3 describes in detail the proposed methodology and how the ensemble idea is applied. Section 4 tests the effectiveness of the proposed approach using a numerical example, the CSTR process, and the widely accepted industry benchmark TEP and analyzes the results. In the final section, we summarize the approach of this paper and provide an outlook on future work in the context of the experimental results.

2. BRIEF REVIEW OF KPCA AND CVA

2.1. KPCA. The general idea of KPCA is to perform PCA in a reproducing kernel Hilbert space. First, all original data samples are mapped to a high-dimensional space (called feature space) using a nonlinear mapping. Then, linear PCA is performed as a dimension reduction technique to extract features in this high-dimensional space.[34,35]

Given an input data matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n] \in \mathbb{R}^{n\times m}$ collected at n time points for m variables. A nonlinear function

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

$\Phi(\bullet)$ is employed to map the original data $\mathbf{X}$ into a high-dimensional feature space $\mathbb{H}$,

$$\mathbb{R}^m \xrightarrow{\Phi(\bullet)} \mathbb{H} \tag{1}$$

Assuming that $\mathbf{X}$ has been centralized, the covariance matrix of the high-dimensional feature space is constructed as

$$\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n}[\Phi(\mathbf{x}_i)][\Phi(\mathbf{x}_i)]^{\mathrm{T}} \tag{2}$$

Then, in the feature space $\mathbb{H}$, PCA is conducted by solving the eigenvalue problem,

$$\lambda\boldsymbol{\omega} = \mathbf{C}\boldsymbol{\omega} \tag{3}$$

Here, λ and $\boldsymbol{\omega}$ are the eigenvalue and eigenvector of the matrix $\mathbf{C}$, respectively. The eigenvector $\boldsymbol{\omega}$ can be represented through a linear combination of mapped data points with a series of coefficients $\beta_j (j = 1,\cdots,n)$,

$$\boldsymbol{\omega} = \sum_{j=1}^{n}\beta_j\Phi(\mathbf{x}_j) \tag{4}$$

By multiplying both sides of eq 3 by $\Phi(\mathbf{x}_k)^{\mathrm{T}}$ $(k = 1,\cdots,n)$, it yields,

$$\lambda\sum_{j=1}^{n}\beta_j < \Phi(\mathbf{x}_k), \Phi(\mathbf{x}_j) > = \frac{1}{n}\sum_{i=1}^{n}\alpha_j < \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) > < \Phi(\mathbf{x}_k), \Phi(\mathbf{x}_i)> \tag{5}$$

where $<\bullet, \bullet>$ represents the inner product operator. Generally, although the mapping $\Phi(\bullet)$ exists, it may not always be computationally accessible. To avoid such explicit mappings, the use of kernel functions has become popular. Thus, an $n \times n$ kernel matrix $\mathbf{K}_{ij} = <\Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j)>$ is defined. Further, the centering operation on kernel matrices is performed,

$$\mathbf{K}_c = \mathbf{K} - \mathbf{1}_n\mathbf{K} - \mathbf{K}\mathbf{1}_n + \mathbf{1}_n\mathbf{K}\mathbf{1}_n, \mathbf{1}_n = \frac{1}{n}\begin{bmatrix}1 & \cdots & 1\\ \vdots & \ddots & \vdots\\ 1 & \cdots & 1\end{bmatrix} \in \mathbb{R}^{n\times n} \tag{6}$$

The eigenvalue decomposition problem eq 5 is rewritten as,

$$n\lambda\boldsymbol{\beta} = \mathbf{K}_c\boldsymbol{\beta} \tag{7}$$

where $\boldsymbol{\beta} = [\beta_1, ..., \beta_n]^{\mathrm{T}}$ is the eigenvector of the kernel matrix $\mathbf{K}_c$ with corresponding eigenvalues $\lambda_1, ..., \lambda_n$.

2.2. CVA. CVA is a technique for reducing dimensions that has been extensively researched in the field of process monitoring.[36,37] For the purpose of taking the autocorrelation into consideration, the past and future vectors $\mathbf{x}_{p,k}$ and $\mathbf{x}_{f,k}$ are constructed. With a time-lag number l, $\mathbf{x}_{p,k}$ and $\mathbf{x}_{f,k}$ are constructed as

$$\mathbf{x}_{p,k} = \begin{bmatrix}\mathbf{x}_{k-1}\\ \mathbf{x}_{k-2}\\ \vdots\\ \mathbf{x}_{k-l}\end{bmatrix} \in \mathbb{R}^{ml}, \quad \mathbf{x}_{f,k} = \begin{bmatrix}\mathbf{x}_{k}\\ \mathbf{x}_{k+1}\\ \vdots\\ \mathbf{x}_{k+l-1}\end{bmatrix} \in \mathbb{R}^{ml} \tag{8}$$

Then, the past and future Hankel matrices $\mathbf{X}_p$ and $\mathbf{X}_f$ are obtained,

$$\mathbf{X}_p = [\mathbf{x}_{p,l+1}, \mathbf{x}_{p,l+2},\cdots,\mathbf{x}_{p,l+N}] \in \mathbb{R}^{ml\times N}$$
$$\mathbf{X}_f = [\mathbf{x}_{f,l+1}, \mathbf{x}_{f,l+2},\cdots,\mathbf{x}_{f,l+N}] \in \mathbb{R}^{ml\times N} \tag{9}$$

where $N = n - 2l + 1$. The overall estimates of the sample-based covariance and cross-covariance matrices of the past and future vectors are calculated,

$$\begin{bmatrix}\Sigma_{pp} & \Sigma_{pf}\\ \Sigma_{fp} & \Sigma_{ff}\end{bmatrix} = \frac{1}{N-1}\begin{bmatrix}\mathbf{X}_p\mathbf{X}_p^{\mathrm{T}} & \mathbf{X}_p\mathbf{X}_f^{\mathrm{T}}\\ \mathbf{X}_f\mathbf{X}_p^{\mathrm{T}} & \mathbf{X}_f\mathbf{X}_f^{\mathrm{T}}\end{bmatrix} \tag{10}$$

To gain the best correlated linear combination of past and future observations, a singular value decomposition (SVD) is computed on the following Hankel matrix $\mathbf{H}$.[38]

$$\mathbf{H} = \Sigma_{ff}^{-1/2}\Sigma_{fp}\Sigma_{pp}^{-1/2} = \mathbf{U}\Lambda\mathbf{V}^{\mathrm{T}} \tag{11}$$

Here, $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular matrices of $\mathbf{H}$, respectively. $\Lambda = diag[\sigma_1, \sigma_2, ..., \sigma_q]$ is a diagonal matrix containing all singular values, and q is the rank of $\mathbf{H}$.

For dimensionality reduction, $\mathbf{U}_w \in \mathbb{R}^{ml\times w}$ and $\mathbf{V}_w \in \mathbb{R}^{ml\times w}$ contain the first w columns of $\mathbf{U}$ and $\mathbf{V}$, respectively. Thus, the weight matrices $\mathbf{J}$ and $\mathbf{L}$ are derived,

$$\mathbf{J} = \mathbf{V}_w^{\mathrm{T}}\Sigma_{pp}^{-1/2} \in \mathbb{R}^{w\times ml}$$
$$\mathbf{L} = (\mathbf{I} - \mathbf{V}_w\mathbf{V}_w^{\mathrm{T}})\Sigma_{pp}^{-1/2} \in \mathbb{R}^{ml\times ml} \tag{12}$$

Finally, the state variables and the residuals are derived from the past vector matrix, respectively.

$$\mathbf{Z} = \mathbf{J}\mathbf{X}_p \in \mathbb{R}^{w\times\mathbf{N}}$$
$$\mathbf{E} = \mathbf{L}\mathbf{X}_p \in \mathbb{R}^{ml\times\mathbf{N}} \tag{13}$$

3. PROPOSED EKCVA-BASED PROCESS MONITORING

3.1. EKCVA Model. To overcome the limitation that a singular kernel matrix may be involved, KPCA is performed before exploring the correlation relation in process variables. In KPCA, the Gaussian kernel function is widely employed for process monitoring which is established,

$$k(x, y) = \exp(-\|x - y\|^2/h) \tag{14}$$

where h is the kernel width. A specific h will only generate a single model. Although different methods for selecting an optimal h for process monitoring have been devised, the problem remains unsolved.[25] For the improvement of process monitoring performance, Li et al.[32] applied ensemble learning to the KPCA method. Inspired by this idea, we also adopt ensemble learning in the proposed EKCVA.

To do so, multiple Gaussian kernel functions are defined,

$$k^{(i)}(x, y) = \exp(-\|x - y\|^2/h_i) \tag{15}$$

where $h_i = 2^{i-1}cm\sigma^2$, and $i = 1, ..., n_l$ represents the ith submodel, and n_l stands for the ensemble size.

The training data set is denoted by $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2,..., \mathbf{x}_n] \in \mathbb{R}^{n\times m}$. For each submodel, the kernel matrix $\mathbf{K}^{(i)}$ is computed by $\mathbf{K}_{ij}^{(i)} = <\Phi(\mathbf{x}_i),\Phi(\mathbf{x}_j)>$. $\mathbf{K}_c^{(i)}$ denotes the centered kernel matrix of the ith submodel which is computed as eq 6. Now, performing KPCA in high-dimensional space is equivalent to resolving the eigenproblem of eq 16

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

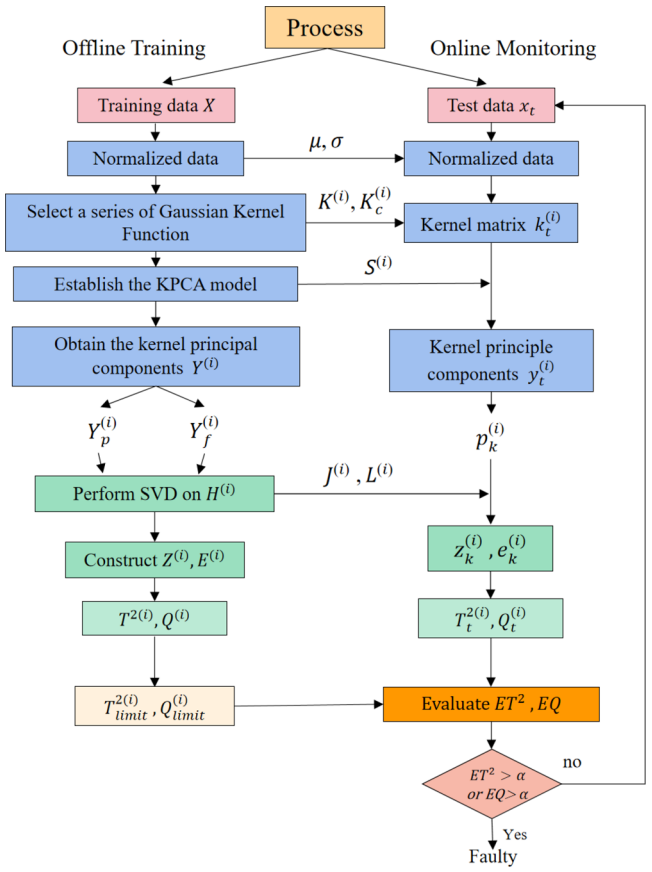


Figure 1. EKCVA model for nonlinear dynamic process monitoring.

$$n\lambda^{(i)}\boldsymbol{\beta}^{(i)} = \mathbf{K}_c^{(i)}\boldsymbol{\beta}^{(i)} \tag{16}$$

with $\lambda^{(i)} = diag(\lambda_1^{(i)}, ..., \lambda_n^{(i)})$ is the diagonal matrix of the corresponding eigenvalues and $\boldsymbol{\beta}^{(i)} = [\beta_1^{(i)}, ..., \beta_n^{(i)}]$ is the matrix containing the eigenvectors of the kernel matrix $\mathbf{K}_c^{(i)}$. Based on the solution of eigenvector problem eq 16, the original data is projected into kernel space. To avoid the singularity problem arising from the execution of the CVA algorithm in a high-dimensional space, the reduced Kernel matrix $\tilde{\mathbf{Y}}^{(i)}$ is computed as follows:

$$\tilde{\mathbf{Y}}^{(i)} = \mathbf{S}^{(i)^{\mathrm{T}}}\mathbf{K}_c^{(i)} \in \mathbb{R}^{r\times n} \tag{17}$$

where r is the number of the principal component and $\mathbf{S}^{(i)} = [\beta_1^{(i)}, ..., \beta_r^{(i)}]$.

$\tilde{\mathbf{Y}}^{(i)}$ can be considered as postprocessing data to build the CVA model in kernel space. Following this idea, the past and future data vectors $\tilde{\mathbf{y}}_{p,k}^{(i)}$ and $\tilde{\mathbf{y}}_{f,k}^{(i)}$ in kernel space are constructed by rearranging each row vector of $\tilde{\mathbf{Y}}^{(i)}$,

$$\tilde{\mathbf{y}}_{p,k}^{(i)} = \begin{bmatrix} \tilde{\mathbf{y}}_{k-1}^{(i)} \\ \tilde{\mathbf{y}}_{k-2}^{(i)} \\ \vdots \\ \tilde{\mathbf{y}}_{k-l}^{(i)} \end{bmatrix} \in \mathbb{R}^{rl}, \quad \tilde{\mathbf{y}}_{f,k}^{(i)} = \begin{bmatrix} \tilde{\mathbf{y}}_{k}^{(i)} \\ \tilde{\mathbf{y}}_{k+1}^{(i)} \\ \vdots \\ \tilde{\mathbf{y}}_{k+l-1}^{(i)} \end{bmatrix} \in \mathbb{R}^{rl} \tag{18}$$

Here, $\tilde{\mathbf{y}}_k^{(i)} \in \mathbb{R}^r$ is the row vector of $\tilde{\mathbf{Y}}^{(i)}$. Similarly, two Hankel matrices $\tilde{\mathbf{Y}}_p^{(i)}$ and $\tilde{\mathbf{Y}}_f^{(i)}$ are formed similar to eq 9.

Based on the constructed Hankel matrices $\tilde{\mathbf{Y}}_p^{(i)}$ and $\tilde{\mathbf{Y}}_f^{(i)}$, the transformation matrices $\tilde{\mathbf{J}}^{(i)}$ and $\tilde{\mathbf{L}}^{(i)}$ are estimated by performing SVD on the scaled Hankel matrix,

$$\tilde{\mathbf{H}}^{(i)} = \tilde{\Sigma}_{ff}^{(i)^{-1/2}}\tilde{\Sigma}_{fp}^{(i)}\tilde{\Sigma}_{pp}^{(i)^{-1/2}} = \tilde{\mathbf{U}}^{(i)}\tilde{\Lambda}^{(i)}\tilde{\mathbf{V}}^{(i)^{\mathrm{T}}} \tag{19}$$

where $\tilde{\Sigma}_{ff}^{(i)}$, $\tilde{\Sigma}_{fp}^{(i)}$ and $\tilde{\Sigma}_{pp}^{(i)}$ can be calculated in a similar way to eq 10. The transformation matrices $\tilde{\mathbf{J}}^{(i)}$ and $\tilde{\mathbf{L}}^{(i)}$ are calculated by retaining the first w singular vectors with dominant singular values,

$$\begin{aligned} \tilde{\mathbf{J}}^{(i)} &= \tilde{\mathbf{V}}_w^{(i)^{\mathrm{T}}}\tilde{\Sigma}_{pp}^{(i)^{-1/2}} \in \mathbb{R}^{w\times rl} \\ \tilde{\mathbf{L}}^{(i)} &= (\mathbf{I} - \tilde{\mathbf{V}}_w^{(i)}\tilde{\mathbf{V}}_w^{(i)^{\mathrm{T}}})\tilde{\Sigma}_{pp}^{(i)^{-1/2}} \in \mathbb{R}^{rl\times rl} \end{aligned} \tag{20}$$

Thus, the state variables and residual variables $\tilde{\mathbf{Z}}^{(i)}$ and $\tilde{\mathbf{E}}^{(i)}$ of the ith submodel are estimated,

$$\begin{aligned} \tilde{\mathbf{Z}}^{(i)} &= \tilde{\mathbf{J}}^{(i)}\tilde{\mathbf{Y}}_p^{(i)} \in \mathbb{R}^{w\times \mathbf{N}} \\ \tilde{\mathbf{E}}^{(i)} &= \tilde{\mathbf{L}}^{(i)}\tilde{\mathbf{Y}}_p^{(i)} \in \mathbb{R}^{rl\times \mathbf{N}} \end{aligned} \tag{21}$$

3.2. EKCVA-Based Process Monitoring. Offline modeling and online monitoring are two stages of the data-driven process monitoring strategy. The EKCVA model is developed utilizing the acquired samples under normal operating conditions in the offline modeling stage. With a chosen confidence level, the relevant upper control limits are determined. For online monitoring, the statistics of the newly collected samples are constructed and compared to the upper control limits. While the online monitoring statistic is over the corresponding upper control limit, an alarm is triggered to indicate the occurrence of a fault.

Two commonly used statistical metrics for fault detection are the $\mathbf{T}^2$ statistic and $\mathbf{Q}$ statistic.[19] For the ith submodel as expressed in eq 20, the monitoring $\mathbf{T}^2$ and $\mathbf{Q}$ statistics are established from $\tilde{\mathbf{Z}}^{(i)}$ and $\tilde{\mathbf{E}}^{(i)}$,

$$\begin{aligned} \mathbf{T}_k^{2(i)} &= \sum_{j=1}^{w} z_{j,k}^{(i)2} \\ \mathbf{Q}_k^{(i)} &= \sum_{j=1}^{rl} \varepsilon_{j,k}^{(i)2} \end{aligned} \tag{22}$$

where $z_{j,k}^{(i)}$ and $\varepsilon_{j,k}^{(i)}$ are the elements of the jth row and kth column of the $\tilde{\mathbf{Z}}^{(i)}$ and $\tilde{\mathbf{E}}^{(i)}$ matrices.

Usually, the Gaussian distribution assumption would be violated in practice. Thus, it is inappropriate to calculate the upper control limits from F-distribution and χ square distribution for the established $\mathbf{T}^2$ and $\mathbf{Q}$ statistics. To determine the upper control limits, the widely used kernel density estimation (KDE) is employed in this study. In KDE, a smoothed peak function (kernel) is used to estimate the probability density functions (PDF) from the observed data points curve.[22,39,40] Thus, the Gaussian distribution assumption of process data is not necessary. Taking the $\mathbf{T}^2$ as an example, the PDF of the calculated N $\mathbf{T}^2(k)$, $k = 1, ..., N$ statistics is estimated as

$$p(\mathbf{T}^2) = \frac{1}{Nh}\sum_{j=1}^{N}\mathcal{K}\left(\frac{\mathbf{T}^2 - \mathbf{T}^2(k)}{d}\right) \tag{23}$$

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

where d is the kernel bandwidth, and $\mathcal{K}(\cdot)$ is the kernel function. A rough estimate of the optimal bandwidth is based on an approximation of the minimized mean squared error of integration.[19] Usually, the kernel function $\mathcal{K}(\cdot)$ is defined,

$$\mathcal{K}(g) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{g^2}{2}\right) \tag{24}$$

Then, for a given confidence level α, the upper control limits of $\mathbf{T}^{2(i)}$ and $\mathbf{Q}^{(i)}$ can be computed for the ith submodel by solving the following equations,

$$P(\mathbf{T}^{2(i)} < \mathbf{T}^{2(i)}_{\text{lim}it} = \int_{-\infty}^{\mathbf{T}^{2(i)}_{\text{limit}}} p(\mathbf{T}^{2(i)})\, \mathrm{d}\mathbf{T}^{2(i)} = \alpha$$

$$P(\mathbf{Q}^{(i)} < \mathbf{Q}^{(i)}_{\text{lim}it} = \int_{-\infty}^{\mathbf{Q}^{(i)}_{\text{limit}}} p(\mathbf{Q}^{(i)})\, \mathrm{d}\mathbf{Q}^{(i)} = \alpha \tag{25}$$

While a new sample $\mathbf{x}_t \in \mathbb{R}^m$ is collected, the centralized Gaussian matrix is calculated through a nonlinear mapping: $\mathbf{x}_t \rightarrow \phi(\mathbf{x}_t)$,

$$\mathbf{k}_t = \phi(\mathbf{x}_t)\Phi(\mathbf{X})^{\mathrm{T}}$$

$$\bar{\mathbf{k}}_c = \mathbf{k}_t - \mathbf{1}_t\mathbf{K} - \mathbf{k}_t\mathbf{1}_n + \mathbf{1}_t\mathbf{K}\mathbf{1}_n \tag{26}$$

where $\mathbf{1}_t = 1/n[1, ..., 1] \in \mathbb{R}^{1\times n}$. For the ith submodel, the projection of the test sample after performing KPCA is

$$\mathbf{y}_t^{(i)} = \mathbf{S}^{(i)}\bar{\mathbf{k}}_c^{(i)} \tag{27}$$

Then, two statistics $\mathbf{T}_t^{2(i)}$ $(i = 1, ..., n_l)$ and $\mathbf{Q}_t^{(i)}$ $(i = 1, ..., n_l)$ can be constructed for the ith KCVA submodel.

To take the monitoring results of all submodels into consideration and enhance the interpretation of the final monitoring results, the Bayesian inference strategy is used to convert the monitoring results into fault probabilities of each submodel. For this purpose, the fault probabilities of the two monitoring statistics are calculated,

$$P^{(i)}_{\mathbf{T}^2_t}(F|\mathbf{x}) = \frac{P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x}|F)P^{(i)}_{\mathbf{T}^2_t}(F)}{P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x})}$$

$$P^{(i)}_{\mathbf{Q}_t}(F|\mathbf{x}) = \frac{P^{(i)}_{\mathbf{Q}_t}(\mathbf{x}|F)P^{(i)}_{\mathbf{Q}_t}(F)}{P^{(i)}_{\mathbf{Q}_t}(\mathbf{x})} \tag{28}$$

$P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x})$ and $P^{(i)}_{\mathbf{Q}_t}(\mathbf{x})$ are computed,

$$P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x}) = P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x}|\mathcal{N})P^{(i)}_{\mathbf{T}^2_t}(\mathcal{N}) + P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x}|\mathcal{F})P^{(i)}_{\mathbf{T}^2_t}(\mathcal{F})$$

$$P^{(i)}_{\mathbf{Q}_t}(\mathbf{x}) = P^{(i)}_{\mathbf{Q}_t}(\mathbf{x}|\mathcal{N})P^{(i)}_{\mathbf{Q}_t}(\mathcal{N}) + P^{(i)}_{\mathbf{Q}_t}(\mathbf{x}|\mathcal{F})P^{(i)}_{\mathbf{Q}_t}(\mathcal{F}) \tag{29}$$

where $\mathcal{N}$ denotes the normal state and $\mathcal{F}$ denotes the fault state. The prior probability terms $P^{(i)}_{\mathbf{T}^2_t}(\mathcal{N})$, $P^{(i)}_{\mathbf{Q}_t}(\mathcal{N})$, $P^{(i)}_{\mathbf{T}^2_t}(\mathcal{F})$, and $P^{(i)}_{\mathbf{Q}_t}(\mathcal{F})$ are determined as $1-\alpha$. Meanwhile, $P^{(i)}_{\mathbf{T}^2_t}(N)$ and $P^{(i)}_{\mathbf{Q}_t}(\mathcal{N})$ are equal to α. The likelihood terms $P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x}|\mathcal{N})$, $P^{(i)}_{\mathbf{Q}_t}(\mathbf{x}|\mathcal{N})$, $P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x}|\mathcal{F})$, and $P^{(i)}_{\mathbf{Q}_t}(\mathbf{x}|\mathcal{F})$ are determined as,

$$P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x}|\mathcal{N}) = \exp(-\mathbf{T}_t^{2(i)}/\mathbf{T}^{2(i)}_{\text{limit}}) \tag{30}$$

$$P^{(i)}_{\mathbf{T}^2_t}(\mathbf{x}|\mathcal{F}) = \exp(-\mathbf{T}^{2(i)}_{\text{limit}}/\mathbf{T}_t^{2(i)}) \tag{31}$$

$$P^{(i)}_{\mathbf{Q}_t}(\mathbf{x}|\mathcal{N}) = \exp(-\mathbf{Q}_t^{(i)}/\mathbf{Q}^{(i)}_{\text{limit}}) \tag{32}$$

$$P^{(i)}_{\mathbf{Q}_t}(\mathbf{x}|\mathcal{F}) = \exp(-\mathbf{Q}^{(i)}_{\text{limit}}/\mathbf{Q}_t^{(i)}) \tag{33}$$

$P^{(i)}_{\mathbf{T}^2_t}(\mathcal{F}|\mathbf{x})$ and $P^{(i)}_{\mathbf{Q}_t}(\mathcal{F}|\mathbf{x})$ in eq 28 are easily obtained based on the above equations.

The weighted combination approach is used to fuse the calculated fault probabilities. First, as demonstrated in eq 34, the weights of the monitoring results for each submodel are determined. It can be found that the submodel with higher fault probability will make a bigger impact on the final monitoring result.

$$C^{(i)}_{\mathbf{T}^2_t} = \frac{P^{(i)}_{\mathbf{T}^2_t}(\mathcal{F}|\mathbf{x})}{\sum_{j=1}^{n_l} P^{(j)}_{\mathbf{T}^2_t}(\mathcal{F}|\mathbf{x})}$$

$$C^{(i)}_{\mathbf{Q}_t} = \frac{P^{(i)}_{\mathbf{Q}_t}(\mathcal{F}|\mathbf{x})}{\sum_{j=1}^{n_l} P^{(j)}_{\mathbf{Q}_t}(\mathcal{F}|\mathbf{x})} \tag{34}$$

where $C^{(1)}_{\mathbf{T}^2_t} + C^{(2)}_{\mathbf{T}^2_t} + \cdots + C^{n}_{\mathbf{T}^2_t l} = 1$ and $C^{(1)}_{\mathbf{Q}_t} + C^{(2)}_{\mathbf{Q}_t} + \cdots + C^{n_l}_{\mathbf{Q}_t} = 1$. Second, two new unified monitoring indicators are defined to fuse the monitoring results of the submodels:

$$ET^2 = \sum_{i=1}^{n_l} C^{(i)}_{\mathbf{T}^2_t} P^{(i)}_{\mathbf{T}^2_t}(\mathcal{F}|\mathbf{x})$$

$$EQ = \sum_{i=1}^{n_l} C^{(i)}_{\mathbf{Q}_t} P^{(i)}_{\mathbf{Q}_t}(\mathcal{F}|\mathbf{x}) \tag{35}$$

With the unified monitoring indicators, the fault detection logic is as follows,

$$\begin{cases} ET^2 < 1 - \alpha \text{ and } \mathrm{EQ} < 1 - \alpha \Rightarrow \text{fault} - \text{free} \\ \text{otherwise} \Rightarrow \text{faulty} \end{cases} \tag{36}$$

The details of each step of the process monitoring procedure are summarized as follows:

- Offline training
  - Collect and normalize the training samples $\mathbf{X}$ under normal condition.
  - Select n_l appropriate kernel width parameters $h_i = 2^{i-1}cm\sigma^2$, $i = 1, ..., n_l$, and perform KPCA algorithm to obtain the reduced kernel matrix $\tilde{\mathbf{Y}}^{(i)}$, $i = 1, ..., n_l$.
  - Construct past and future vectors from the reduced kernel component space using eq 18 for each submodel, and compute the cross-covariance and covariance matrices.
  - For the $i = 1, ..., n_l$ submodels, perform SVD on $\tilde{\mathbf{H}}^{(i)}$ to derive the projection matrices $\tilde{\mathbf{J}}^{(i)}$ and $\tilde{\mathbf{L}}^{(i)}$.
  - For the $i = 1, ..., n_l$ submodels, determine state variables $\tilde{\mathbf{Z}}^{(i)}$ and residuals $\tilde{\mathbf{E}}^{(i)}$, and compute the monitoring indices and their upper control limits using eq 22 and eq 25.
- Online monitoring
  - Collect and normalize the new sample $\mathbf{x}_t$.
  - Use eq 26 to construct the kernel matrix $\bar{\mathbf{k}}_t^{(i)}$, $i = 1, ..., n_l$ with normalized $\mathbf{x}_t$ for each submodel.
  - Calculate the score matrix $\mathbf{y}_t^{(i)}$ using eq 27, and construct the past data matrix at the point k.

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

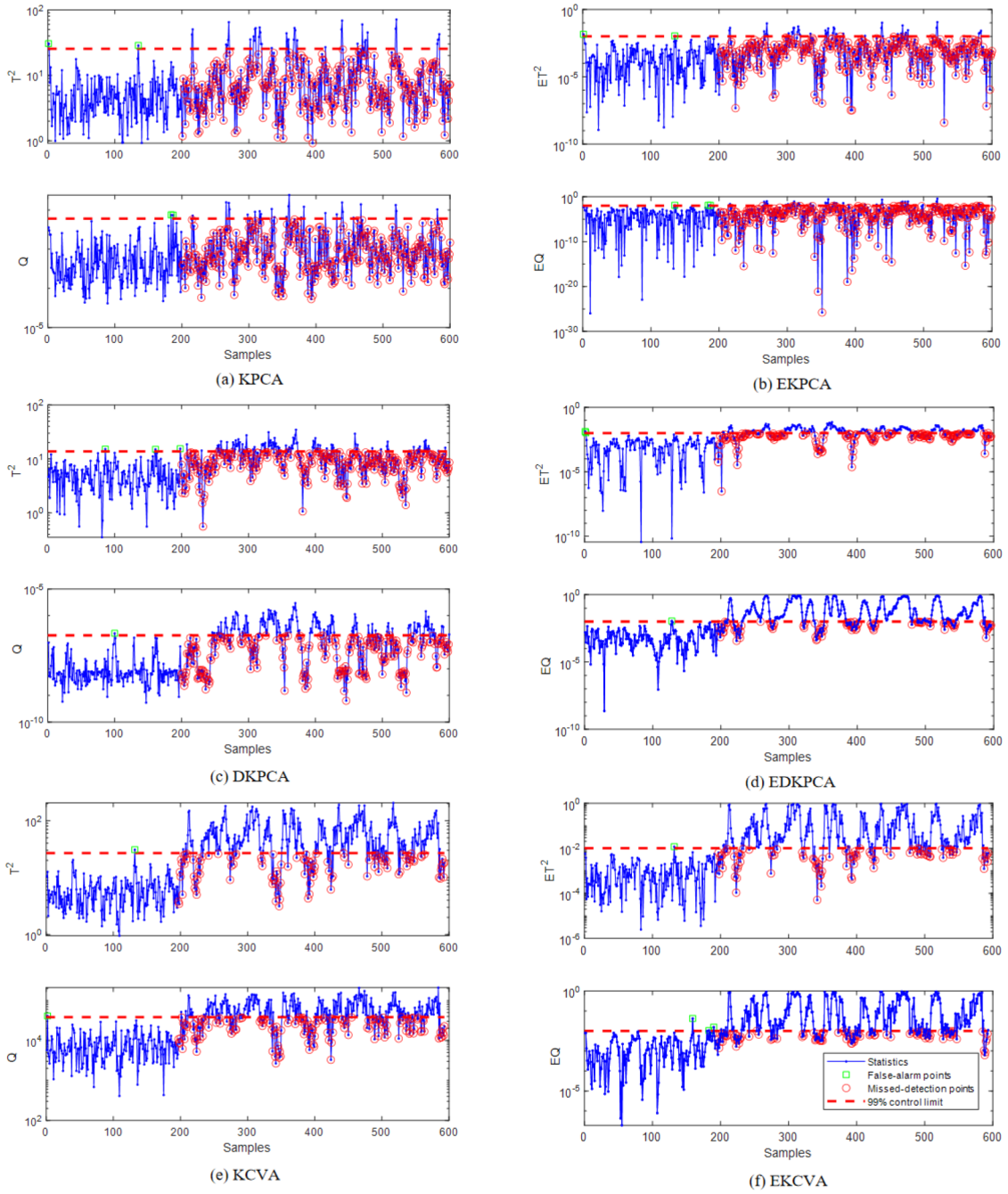


Figure 2. Monitoring diagrams of the dynamic fault (a) KPCA, (b) EKPCA, (c) DKPCA, (b) EDKPCA, (e) KCVA, and (f) EKCVA.

- For the *i*th submodel, calculate state and residual of $\mathbf{x}_t$ by multiplying the transformation matrices $\tilde{\mathbf{J}}^{(i)}$ and $\tilde{\mathbf{L}}^{(i)}$ to obtain the monitoring statistics ($\mathbf{T}_t^{2(i)}$ and $\mathbf{Q}_t^{(i)}$).
- Compute the fault probabilities for $i = 1, ..., n_l$ submodels by eq 28−eq 33.
- Obtain the monitoring indices ET^2 and EQ using eq 34 and eq 35.
- Determine the process status based on the fault detection logic eq 36.

The scheme of the proposed EKCVA-based process monitoring method is depicted in Figure 1.

3.3. Parameters Selection. The parameters selection is critical in the proposed EKCVA-based process monitoring algorithm. The number of time lag l, the number of kernel principal elements (KPCs) r, the number of canonical variables (CVs) w, the kernel width parameter h in KPCA, and the ensemble size n_l should be appropriately selected. The determination of these parameters is discussed as follows:

- number of time lag l: l is determined from the autocorrelation function of the measured values. Since the autocorrelation of each variable is different, we used the autocorrelation of the sum of all variables in this work. This allows for a better description of the dynamic characteristics of the process.[19]
- number of KPCs r: To determine how many KPCs should be retained, the cumulative percent variance (CPV) is

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

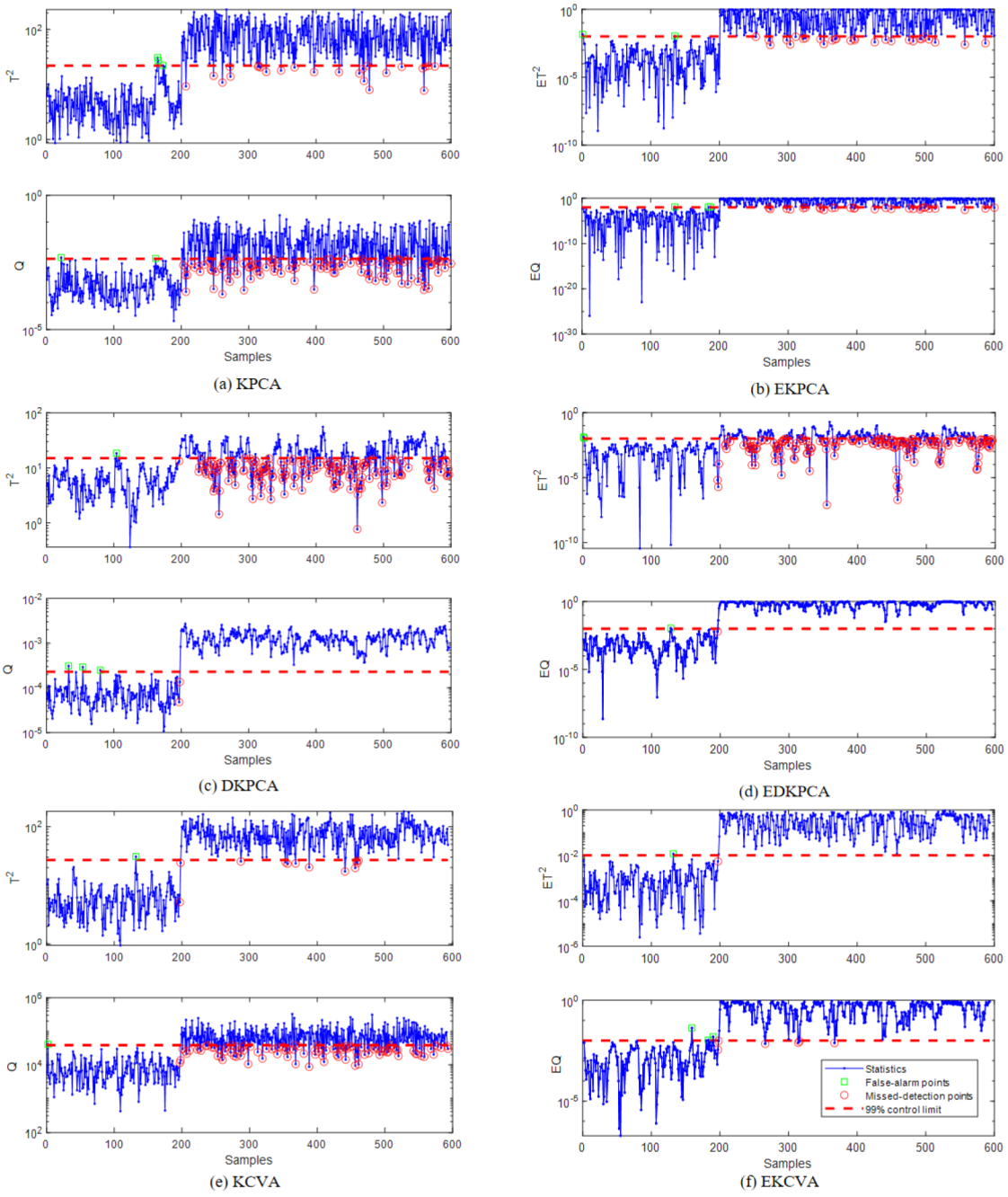


Figure 3. Monitoring diagrams of the nonlinear fault (a) KPCA, (b) EKPCA, (c) DKPCA, (b) EDKPCA, (e) KCVA, and (f) EKCVA.

Table 1. Numerical Example: Monitoring Performance by Five Methods for Each Fault

	KPCA		EKPCA		DKPCA		EDKPCA		KCVA		EKCVA	
fault	T^2	Q	T^2	Q	T^2	Q	T^2	Q	T^2	Q	T^2	Q
1	9.75[a]	6.25	10.75	11.25	25.75	41.5	56.25	71.50	70.75	59.75	75.72	75.00
	1.00[b]	1.00	1.00	1.50	1.53	0.51	1.02	0.51	0.51	0.51	0.51	1.53
2	93.75	69.25	92.25	93.25	54.25	99.50	55.00	99.75	96.25	72.75	99.75	98.50
	1.50	1.00	1.50	1.00	0.51	1.53	1.02	0.51	0.51	0.51	0.51	1.53

[a]First row: FDRs. [b]Second row: FARs.

often calculated.[41,42] The number of KPCs is selected while the CPV is over a predefined value.

- number of CVs w: Similar to that in ref 43, the number of CVs is determined by finding the point where the "knee" appears in the singular value curve.
- kernel width parameter h: According to literature studies, the kernel width parameter is usually determined empirically such as $h = cm\sigma^2$, where c is a constant. m and σ^2 are the dimension and variance of the process data,

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

respectively. If the data are normalized to $\sigma^2 = 1$, then $h = cm$.[32,44]

- ensemble size n_l: Obviously, the selection of n_l is crucial for ensemble learning based process monitoring. If n_l is set to a small value, the process characteristics can not be fully explored. The performance of process monitoring will be degraded. On the other hand, a larger value of n_l implies higher computational complexity, which is not recommended, particularly for kernel-based methods. By considering the trade-off between fault detection performance and computational complexity, the number of ensemble size is determined by cross-validation in a similar way as in ref 33.

4. CASE STUDIES

In this section, the proposed EKCVA-based process monitoring method is evaluated through a numerical example, a closed-loop CSTR process, and an industry benchmark of the Tennessee Eastman process. For comparison study, KPCA, EKPCA,[32] dynamic KPCA (DKPCA),[45] ensemble DKPCA, and KCVA are employed.

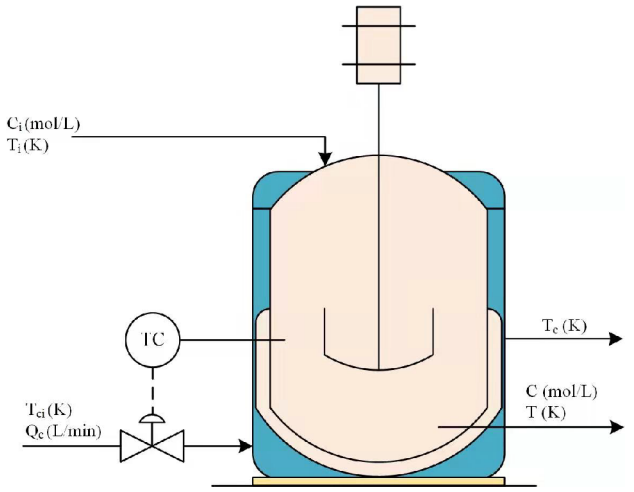


Figure 4. Schematic of the closed-loop CSTR.

Table 2. Constant Values in the CSTR Model

parameter	description	value	units
Q	inlet flow rate	100.0	L/min
V	tank volume	150.0	L
V_c	jacket volume	10.0	L
ΔH_r	heat of reaction	-2.0×10^5	cal/mol
UA	heat transfer coefficient	7.0×10^5	cal/min/K
k_0	pre-exponential factor to k	7.2×10^{10}	min^{-1}
R	activation energy	1.0×10^4	K
ρ, ρ_c	fluid density	1000	g/L
C_p, C_{pc}	fluid heat capacity	1.0	cal/g/K

4.1. Numerical Example. The nonlinear state space model of the numerical example is described as follows,[46]

$$\begin{cases} \mathbf{z}_k = \mathbf{A}\mathbf{z}_{k-1} + \mathbf{c} + \boldsymbol{\xi}_k \\ \mathbf{x}_k = \mathbf{P}\mathbf{z}_k + f(\mathbf{u}_k) + \boldsymbol{\varepsilon}_k \end{cases} \tag{37}$$

Table 3. Fault Description in CSTR Model

fault no.	description	value	type
1	$f = f_0 e^{-\varrho t}$	5×10^{-4}	multiplicative
2	$g = g_0 e^{-\varrho t}$	1×10^{-3}	multiplicative
3	Faults 1 and 2 occur simultaneously		multiplicative
4	$C_i = C_i + \varrho t$	1×10^{-3}	additive
5	$T_i = T_i + \varrho t$	0.05	additive
6	$T_{ci} = T_{ci} + \varrho t$	0.05	additive
7	$C = C + \varrho t$	1×10^{-3}	additive
8	$T = T + \varrho t$	0.05	additive
9	$T_c = T_c + \varrho t$	0.05	additive
10	$Q_c = Q_c + \varrho t$	−0.1	additive

$$\mathbf{A} = \begin{bmatrix} 0.5205 & 0.1022 & 0.0599 \\ 0.5367 & -0.0139 & 0.4159 \\ 0.0412 & 0.6054 & 0.3874 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0.5205 \\ 0.5367 \\ 0.0412 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} 0.4316 & 0.1723 & -0.0574 \\ 0.1202 & -0.1463 & 0.5348 \\ 0.2483 & 0.1982 & 0.4797 \\ 0.1151 & 0.1557 & 0.3739 \\ 0.2258 & 0.5461 & -0.0424 \end{bmatrix}$$

The state space vector $\mathbf{z}_k \in \mathbb{R}^3$ used to describe the dynamic relationship between multivariate time series is which is generated from a vector autoregressive VAR(1) process. $f(\mathbf{u}) = 0.05[u_1, u_2, 5u_1 - 2u_2, u_1^2 - 3u_2, -u_1^3 + 3u_2^2]^T$ is a nonlinear mapping function. The input $\mathbf{u} = [u_1, u_2]$ is a vector uniformly distributed over the interval $[0, 2]$. $\boldsymbol{\xi}_k \in \mathbb{R}^3 \sim \mathbf{N}(0, 1^2)$ and $\boldsymbol{\varepsilon}_k \in \mathbb{R}^5 \sim \mathbf{N}(0, 0.1^2)$ denote the noises.

Two types of faults are considered including the change of dynamics and nonlinearity. All faults are injected from the 201st sampling point. The faults are described as follows:

- Fault 1: $\mathbf{z}_k := \mathbf{z}_k + [0, 2, 0]^T$ for $k > 200$
- Fault 2: $\mathbf{u}_k := \mathbf{u}_k + [3, 0]^T$ for $k > 200$

500 samples were generated under normal conditions for the training model. Two faulty test data sets were generated, which all consist of 600 samples.

For fair comparison, the number of time lag l is set as 4 for all methods; $h = 200$, $r = 5$ are selected for the KPCA, DKPCA, and KCVA models, and $w = 6$ for the KCVA method. For EKPCA, EDKPCA, and EKCVA, the ensemble size is selected as $n_l = 8$, and the kernel parameters $h_i = 5{\cdot}5{\cdot}2^{(i-1)}$ $(i = 1, ..., n_l)$. Other parameters of submodels are selected as KPCA, DKPCA, and KCVA. The confidence level α is 0.99

The results obtained from process monitoring of Fault 1 are set out in Figure 2. As can be seen from Figure 2a,b, the KPCA and EKPCA fail to detect the abnormality from both T^2 and Q statistics. The reason is that the variation of the state space vector is covered due to the process dynamics as shown in eq 37. On the contrary, DKPCA, EDKPCA, and KCVA methods can model the dynamic characteristics of the process by augmenting the process data. Thus, the monitoring performance is significantly improved as displayed in Figure 2c−f. Besides, the good generalization capability inherited from ensemble learning can further improve the model efficiency. Thus, EDKPCA and EKCVA can provide better monitoring performance than DKPCA and KCVA, respectively. Since the nonlinear dynamic behavior can be modeled more accurately by the

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

Table 4. Comparison Results of FDRs for CSTR Process

	KPCA (%)		EKPCA (%)		DKPCA (%)		EDKPCA (%)		KCVA (%)		EKCVA (%)	
fault no.	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ
1	75.50	44.40	75.80	54.40	59.54	83.33	60.04	84.14	84.26	87.69	84.36	87.89
2	83.10	87.90	83.10	94.10	68.27	87.75	68.88	89.16	85.17	88.70	85.37	88.90
3	78.50	88.30	78.60	94.70	69.78	85.54	69.98	86.75	82.14	85.37	82.34	85.57
4	94.80	70.30	94.80	74.70	92.57	72.79	92.57	72.49	91.32	90.21	91.32	90.31
5	87.30	72.10	87.30	73.10	78.11	80.02	78.71	77.91	79.52	75.68	79.72	75.78
6	85.80	81.20	85.80	91.90	77.91	80.62	78.31	80.52	78.30	79.41	78.61	79.72
7	99.10	94.90	99.10	97.40	95.78	99.40	95.78	99.50	99.39	99.60	99.50	99.60
8	91.00	96.70	91.00	98.50	75.30	96.39	75.30	96.79	95.06	96.67	95.26	96.87
9	93.20	96.90	93.20	98.50	83.63	96.39	83.73	96.89	95.76	96.87	95.76	96.87
10	79.70	81.00	79.70	92.60	68.88	78.71	69.28	80.62	77.60	81.84	78.00	81.84
Average	86.82	73.27	86.84	86.99	76.98	86.12	77.26	86.48	86.85	88.20	87.02	88.32

Table 5. Comparison Results of the Average FARs for CSTR Process

KPCA (%)		EKPCA (%)		DKPCA (%)		EDKPCA (%)		KCVA (%)		EKCVA (%)	
T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ
1.00	0	1.00	1.15	0.15	0.30	0.20	0.40	0.40	0.85	0.40	0.90

Table 6. Comparison Results of DDs (hours) for CSTR Process

	KPCA (%)		EKPCA (%)		DKPCA (%)		EDKPCA (%)		KCVA (%)		EKCVA (%)	
fault no.	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ
1	3.50	11.62	3.50	5.02	4.92	2.72	4.92	2.65	2.68	2.20	2.68	2.20
2	3.13	2.25	3.13	1.20	3.52	1.97	3.52	1.93	2.57	1.97	2.57	1.97
3	2.67	1.87	2.67	0.97	4.60	2.43	4.60	2.27	2.70	2.62	2.70	2.62
4	0.95	5.23	0.95	4.67	1.32	4.60	1.32	4.60	1.37	1.82	1.37	1.82
5	2.42	4.67	2.42	4.58	3.70	3.33	3.70	3.35	3.33	4.08	3.33	4.08
6	2.33	3.07	2.33	1.55	3.72	3.35	3.50	3.03	3.63	3.63	0.30	3.43
7	0.18	0.92	0.18	0.50	0.72	0.12	0.72	0.10	0.12	0.08	0.10	0.08
8	1.80	0.73	1.80	0.35	4.32	0.62	4.32	0.57	0.83	0.58	0.80	0.58
9	1.40	0.60	1.40	0.28	2.87	0.57	2.87	0.53	0.75	0.53	0.75	0.53
10	0.67	3.38	0.67	1.18	5.42	0.60	4.73	0.60	0.60	0.60	0.60	0.60
Average	1.91	3.43	1.91	2.03	3.51	2.03	3.42	1.96	1.85	1.81	1.52	1.79

EKCVA model, it can also be observed that the proposed EKCVA method can acquire more excellent performance over DKPCA, EDKPCA, and KCVA.

Fault 2 represents the change of nonlinearity. The monitoring result of Fault 2 is plotted in Figure 3. It can be seen that Fault 2 can be detected by most methods. Compared to the KPCA based method, the KCVA based method can provide a better performance, since the autocorrelation is taken into consideration. Through ensemble learning, the nonlinear feature is more efficiently extracted than single model approaches. Thus, EKCVA-based process monitoring has the best performance among the comparative methods.

To further carry out the comparison of monitoring performance, two indexes are used. One is the fault false alarm rate (FAR), and the other one is the fault detection rate (FDR). The definitions of FAR and FDR are as follows:

$$\mathrm{FDR} = \frac{\text{no. of samples } (J > J_{\mathrm{th}} | f \neq 0)}{\text{total sample } (f \neq 0)} \times 100$$

$$\mathrm{FAR} = \frac{\text{no. of samples } (J > J_{\mathrm{th}} | f = 0)}{\text{total sample } (f = 0)} \times 100 \tag{38}$$

where J and J_{th} denote the test statistic and corresponding threshold. $f \neq 0$ and $f = 0$ denote fault and fault-free, respectively.

Table 1 lists the FDRs and FARs of the comparative methods. It is noticed that the EDKPCA Q statistic can provide better performance than KPCA, EKPCA, and DKPCA for Fault 1 and Fault 2. Thus, it will be helpful for process monitoring by using ensemble learning and considering the process dynamics. On the other hand, it can be found that the EKCVA T^2 statistic can offer the best performance in terms of FAR and FDR. For Fault 1, the FAR of EKCVA T^2 statistic is 0.51% and FDR is 75.72%. For Fault 2, the FAR of EKCVA T^2 statistic is 0.51% and FDR is 99.75%.

4.2. CSTR Process. A CSTR, also known as a completely mixed flow reactor, is a tank reactor with stirring paddles. The physical and chemical changes in the substance in the system are part of the reaction process. Temperature, pressure, and flow rate are some of the parameters that characterize its system attributes. Researchers have used various forms of CSTRs for process monitoring in a variety of studies.[22,39,47] A simplified diagram of the three-state closed-loop CSTR process is depicted in Figure 4. The following equations primarily describe the mechanism of the CSTR process:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{Q}{V}(T_i - T) - f\frac{(\Delta H_r)kC}{\rho C_p} - g\frac{UA}{\rho C_p V}(T - T_c) + \nu_1$$

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

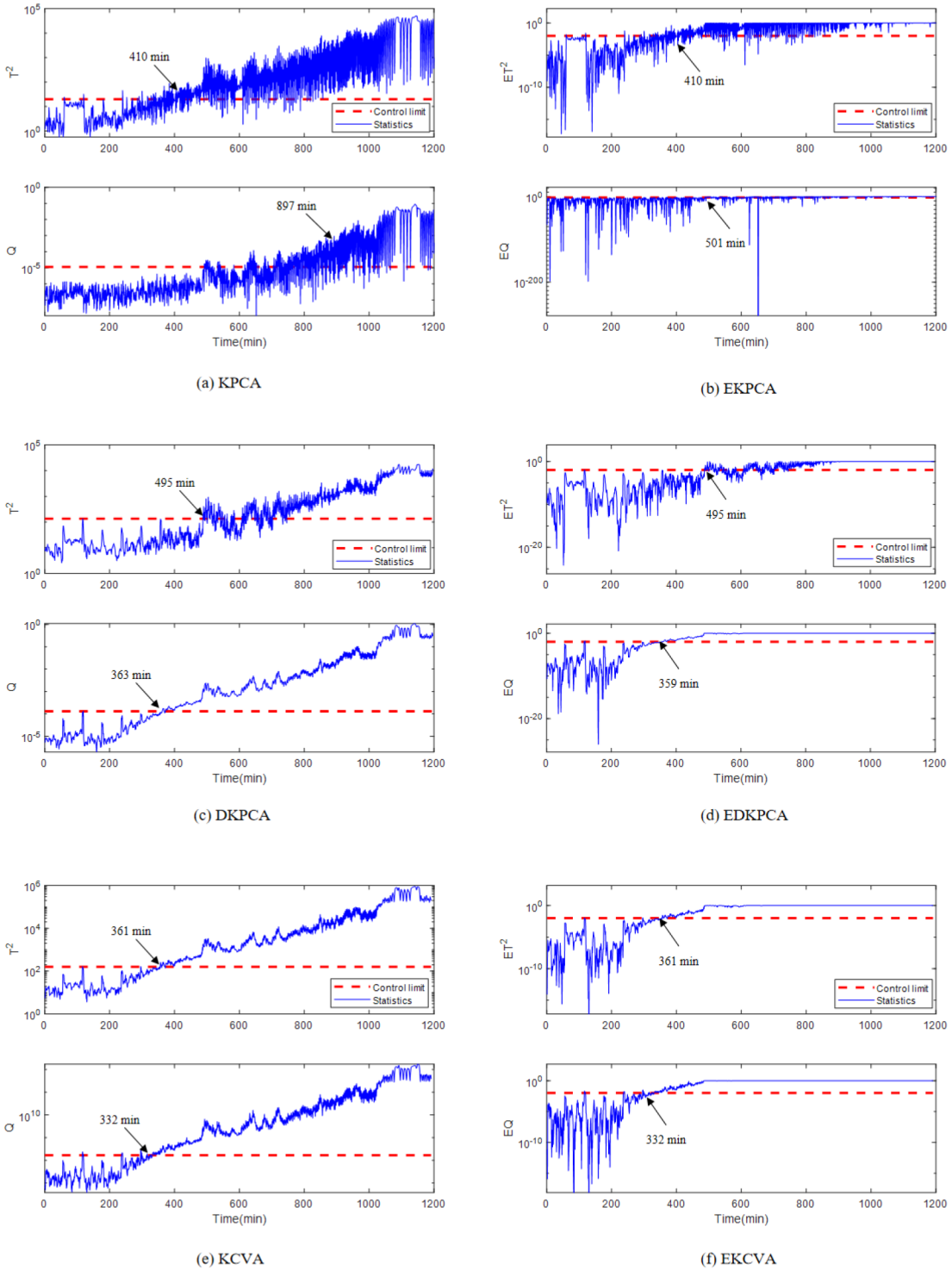


Figure 5. Monitoring charts in CSTR process of Fault 1: (a) KPCA, (b) EKPCA, (c) DKPCA, (b) EDKPCA, (e) KCVA, (f) EKCVA.

$$\frac{dC}{dt} = \frac{Q}{V}(C_i - C) - fkC + v_2$$

$$\frac{dT_c}{dt} = \frac{Q_c}{V_c}(T_{ci} - T_c) + g\frac{UA}{\rho_c C_{pc} V_c}(T - T_c) + v_3$$

There are some constant values in the model, and these parameters are listed in Table 2. The process variables of the model $\mathbf{X} = [T_i, C_i, T_{ci}, T, C, T_c, Q_c]^T$. Furthermore, v_i, $i = 1, 2, 3$ are Gaussian white noise. For all variables, the sampling interval is 1 min. $k = k_0 \exp^{-RT}$ is an Arrhenius-type rate.

In this example, we simulate 10 types of faults in the CSTR process. For Faults 1−3, f and g are equal to 1.00 in normal operation. By attenuating their values to zero, the decay of catalyst and heat transfer pollution can be simulated, respectively. Faults 4−10 are sensor drifts. The details of these faults are described in Table 3. The training data set contains 1200 data points generated under normal operating conditions; 1200 samples are collected for each fault. The time point for fault

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

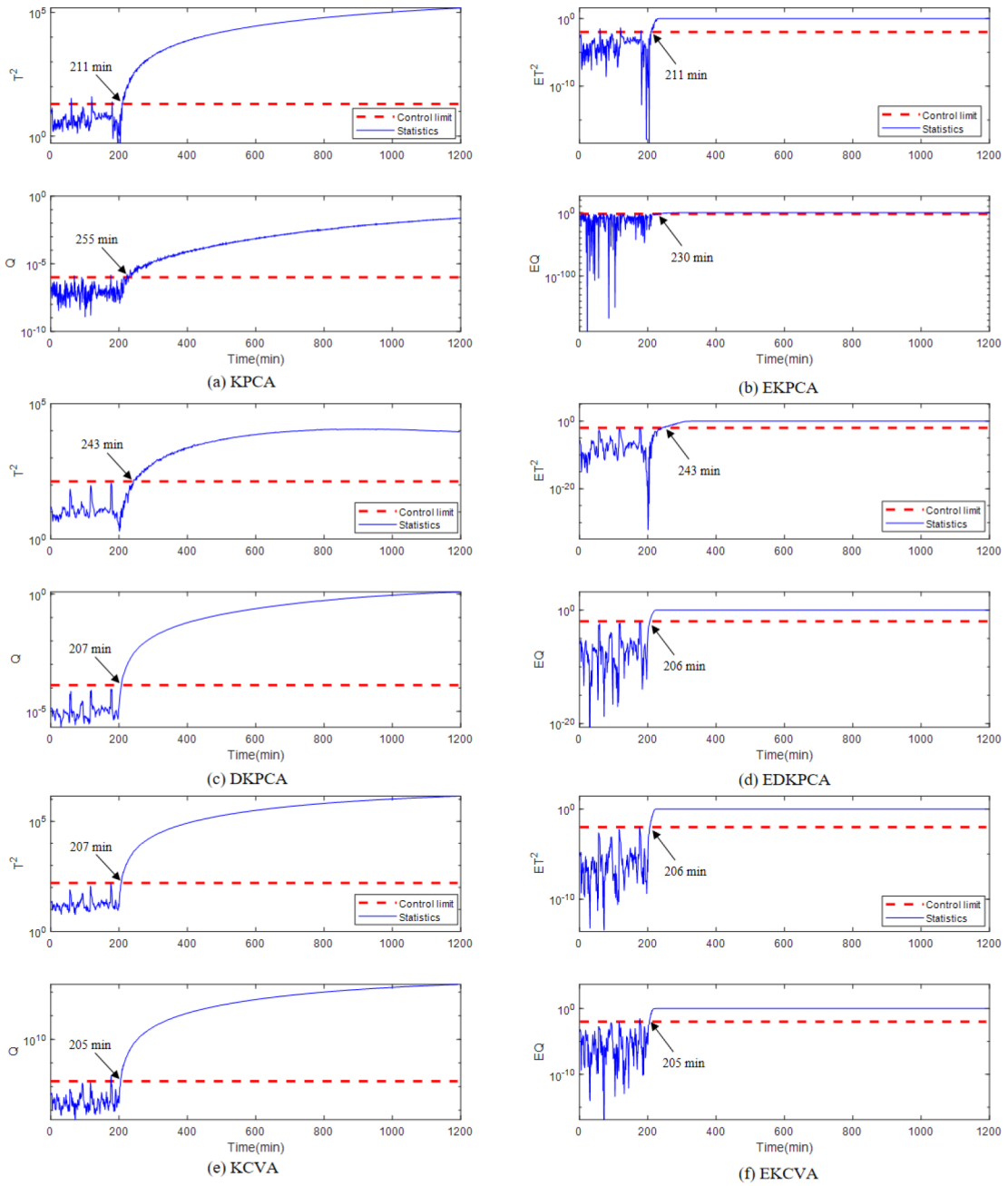


Figure 6. Monitoring charts in CSTR process of Fault 7: (a) KPCA, (b) EKPCA, (c) DKPCA, (b) EDKPCA, (e) KCVA, (f) EKCVA.

import is the 200th minute. The data acquisition for this experiment can be downloaded from https://www.mathworks.com/matlabcentral/fileexchange/66189-feedback-controlled-cstr-process-for-fault-simulation. As it can be seen from the model expression, the CSTR is a typical nonlinear dynamic process.

According to the parameter selection procedure, $h = 7{\cdot}20{\cdot}2^{(i-1)}$ $(i = 1, ..., n_l)$, $l = 5$, $r = 6$, $w = 5$, $n_l = 10$ are set for the EKCVA model. For the EDKPCA model, $l = 5$, $r = 6$, $n_l = 10$, $h = 7{\cdot}20{\cdot}2{\cdot}2^{(i-1)}$ $(i = 1, ..., n_l)$. For EKPCA model, $r = 6$, $n_l = 10$, $h = 7{\cdot}20{\cdot}2^{(i-1)}$ $(i = 1, ..., n_l)$. KCVA, DKPCA, and KPCA are submodels of EKCVA, EDKPCA, and EKPCA, respectively, with kernel parameters of $h = 7{\cdot}20$. The confidence level α is 0.99 for determining the control limits. To evaluate the monitoring performance, three indexes including DD, FAR, and FDR are employed. Moreover, the detection time is determined as the first time after five consecutive alarms have been triggered to make the comparison more fair. DDs computed in the experiment are in units of hours.

The monitoring results are listed in Tables 4, 5, and 6. The EKCVA has the highest average FDRs (i.e., 87.02% and 88.32% for the T^2 and Q statistics, respectively). It is also noticed that the average FDRs of the ensemble models are both higher than those of the individual models. EKCVA T^2 and Q statistics are considered to be the most sensitive to all faults, where the DDs are 1.52 and 1.79 h, respectively. Compared to EDKPCA and KCVA, the FARs of EKCVA T^2 and Q statistics are slightly higher. However, the FAR values are at the same level as EKPCA. Nevertheless, the EKCVA T^2 and Q statistics provide the best monitoring performance among these comparative methods.

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

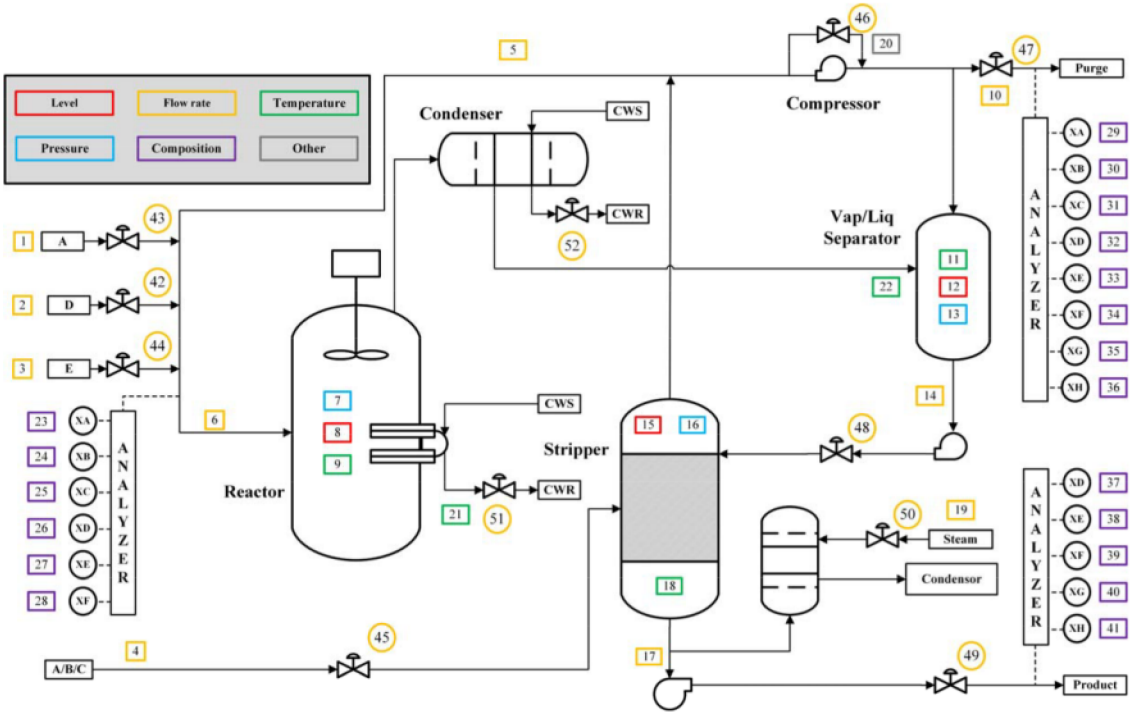


Figure 7. Diagram of Tennessee Eastman benchmark process.

Table 7. Disturbance Description for TE Process

fault no.	fault description	type
1	B composition constant, A/C feed ratio in Steam 4	step change
2	A/C ratio constant, B composition in Steam 4	step change
3	D feed temperature in Steam 2	step change
4	reactor cooling water inlet temperature	step change
5	condenser cooling water inlet temperature	step change
6	a feed loss in Steam 1	step change
7	C header pressure loss in Steam 4	step change
8	A, B, C feed composition in Steam 4	random variation
9	D feed temperature in Steam 2	random variation
10	C feed temperature in Steam 4	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
16–20	unknown faults	unknown
21	stable valve in Stream 4	constant position

The results obtained from the preliminary analysis of Fault 1 and Fault 7 can be compared in Figure 5 and Figure 6, respectively. For Fault 1, it is a catalyst degradation fault. Variations in process variables are not immediately apparent at first. The difference between normal and abnormal conditions will progressively become significant after a few hours. As shown in Figure 5, the DDs of all methods are relatively large, but the EKCVA model has the smallest DDs. Additionally, EKCVA achieves the best monitoring results with the highest FDR. Fault 7 is the sensor drifts in the concentration of product reactants with a small magnitude of failure. Since KPCA and EKPCA models ignore the dynamics of process data, it can be found that there are larger DDs as displayed in Figure 6. Nevertheless, the proposed EKCVA method has superior monitoring performance than other methods in terms of FDR, FAR, and DD.

4.3. TEP. The industrial benchmark of the Tennessee Eastman process (TEP) was developed by Eastman Chemical Company based on a practical chemical process.[48] The flow diagram of TEP is exhibited in Figure 7. In TEP, the process data have time-varying, strongly coupled, and nonlinear characteristics, thus it is widely used in the comparison study of process monitoring and fault detection.[46,49−51]

In the TEP benchmark, 21 faults were predefined. The specific description of 21 faults is listed in Table 7. For process monitoring, 22 continuous variables and 11 manipulated variables were selected. The TEP data set can be found on the Web site (http://web.mit.edu/braatzgroup/links.html). A total of 500 observations under normal operating conditions were collected as training data. The testing data were acquired under a 48 h run of the simulation, with faults introduced at the eighth hour. And a total of 960 observations were collected.

In this case study, the kernel width of KPCA model is empirically determined as $h = 10m = 330$, $r = 30$.[52] $h_i = 5m{\cdot}2^{i-1}$ $(i = 1, ..., 11)$ is selected for EKPCA model proposed Li and Yang. For EDKPCA model, $l = 5$, $r = 30$, and $h_i = 5m{\cdot}2^{i-1}$ $(i = 1, ..., 8)$. For DKPCA model, $l = 5$, $r = 30$, and $h = 10m$. For KCVA model, $h = 10m$, $r = 30$, $l = 5$, and $w = 26$. For EKCVA, $r = 37$, $w = 24$, $l = 5$, $h_i = 50m{\cdot}2^{i-1}$ $(i = 1, ..., 11)$. The confidence level α is 0.99.

Table 8 and Table 9 summarize the monitoring results for all 21 faults. The proposed EKCVA model can provide superior performance over other models for most faults such as faults 4, 8, 10, and 11 as shown in Table 8. The average FDRs of ET^2 and EQ are 83.74% and 81.26%, respectively, whereas the average FDRs of other methods are below 80%. Although the FAR of EQ (i.e., 4.946%) is higher than other Q statistics, the FAR of ET^2(i.e., 5.554%) is at the same level as other T^2 statistics. The numbers of detection delays (DD) are provided in Table 10. From this table, we can see that the average DD of the EKCVA EQ statistic is the smallest among the monitoring statistics, and

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

Table 8. Comparison Results of FDRs for TE Process

	KPCA (%)		EKPCA (%)		DKPCA (%)		EDKPCA (%)		KCVA (%)		EKCVA (%)	
fault no.	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ
1	99.88	99.25	100	99.25	100	99.5	100	99.50	99.75	99.25	98.87	99.62
2	98.63	98.13	98.75	98.25	98.50	98.25	99.50	98.25	98.62	98.24	98.99	98.25
3	17.75	8.13	18.25	6.50	3.37	3.00	4.25	6.88	9.80	6.28	17.74	15.60
4	100	22.25	100	27.63	100	77.00	100	77.13	98.74	83.17	99.87	99.25
5	36.75	29.13	38.00	100	28.88	100	29.38	100	33.67	27.26	100	100
6	100	100	100	100	100	100	100	100	99.87	99.87	100	100
7	100	99.50	100	100	100	100	100	100	100	99.25	100	100
8	99.38	97.25	99.38	98.25	98.00	97.88	98.13	97.88	99.37	96.48	99.75	99.37
9	14.88	4.75	14.37	5.87	3.50	2.38	4.4	5.63	9.92	6.78	18.24	16.10
10	58.75	60.88	88.62	73.38	85.75	59.50	85.50	91.25	58.42	36.06	93.71	91.57
11	81.88	46.00	71.88	66.75	80.00	62.13	80.63	65.00	71.23	55.65	86.63	78.56
12	99.75	98.38	99.75	97.75	99.38	99.75	99.50	99.87	99.62	96.98	99.75	99.87
13	95.38	94.25	95.50	94.25	95.25	94.38	95.25	94.25	95.10	93.97	95.60	95.85
14	100	99.63	100	100	100	100	100	100	99.87	99.87	99.87	99.87
15	24.38	14.50	25.37	13.88	4.50	5.63	4.38	15.25	18.22	8.79	29.31	25.28
16	49.25	53.13	94.13	78.75	90.25	54.38	90.50	85.87	44.54	21.61	95.22	91.57
17	96.63	81.13	96.88	83.50	97.13	87.50	97.13	89.00	95.85	86.55	95.72	94.84
18	91.00	89.38	91.25	89.38	90.38	89.38	90.50	85.87	90.20	89.82	91.44	91.19
19	48.50	4.13	84.50	21.75	92.50	4.63	93.88	22.75	16.08	12.31	98.74	91.55
20	70.88	49.25	91.87	47.38	80.75	61.88	81.00	87.50	58.54	36.18	90.19	76.76
21	51.50	38.25	56.87	32.37	53.38	41.13	55.00	51.25	42.96	24.37	58.11	42.64
Average	73.10	61.31	79.81	69.99	74.22	68.49	76.62	75.09	68.60	60.89	83.74	81.26

Table 9. Comparison Results of the Average FARs for the TE Process

KPCA (%)		EKPCA (%)		DKPCA (%)		EDKPCA (%)		KCVA (%)		EKCVA (%)	
T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ
5.800	1.875	5.750	1.750	5.030	1.625	5.233	2.201	3.503	1.557	5.554	4.946

Table 10. DDs (Samples) For All Faults with Different Methods

	KPCA (%)		EKPCA (%)		DKPCA (%)		EDKPCA (%)		KCVA (%)		EKCVA (%)	
fault no.	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ	T^2	Q	ET^2	EQ
1	2	7	1	7	1	5	1	5	3	5	2	4
2	13	15	9	15	13	15	13	15	11	11	9	13
3	5	15	5	15	15	15	15	15	13	16	5	3
4	1	1	1	1	1	1	1	1	1	2	1	1
5	1	1	1	1	1	1	1	1	2	3	1	1
6	1	1	1	1	1	1	1	1	2	2	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1
8	1	21	1	1	16	15	11	15	3	4	3	5
9	1	1	1	6	1	1	1	1	3	93	3	5
10	8	16	6	16	6	16	6	16	13	13	11	6
11	6	7	6	7	6	7	6	5	7	9	6	6
12	3	3	3	3	4	3	3	2	3	4	1	1
13	38	47	26	47	38	46	38	47	40	5	1	4
14	1	2	1	2	1	1	1	1	2	2	2	2
15	86	92	92	92	238	92	66	51	397	3	79	2
16	245	30	1	10	7	17	7	10	26	26	6	9
17	20	28	2	26	20	24	20	21	2	4	21	21
18	18	20	15	61	18	87	18	61	62	62	18	4
19	1	11	2	2	2	11	2	9	3	3	2	2
20	40	75	63	70	63	73	63	75	81	81	67	48
21	281	257	40	257	251	257	244	257	259	167	208	52
avg	37	31	13	31	34	33	25	29	44	24	21	9

the average DD of EKCVA ET^2 statistic is 21, which is smaller than KPCA T^2, DKPCA T^2, EDKPCA ET^2, and KCVA T^2. Nevertheless, the proposed EKCVA has superior monitoring performance than the other methods in terms of FDR, FAR, and DD.

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

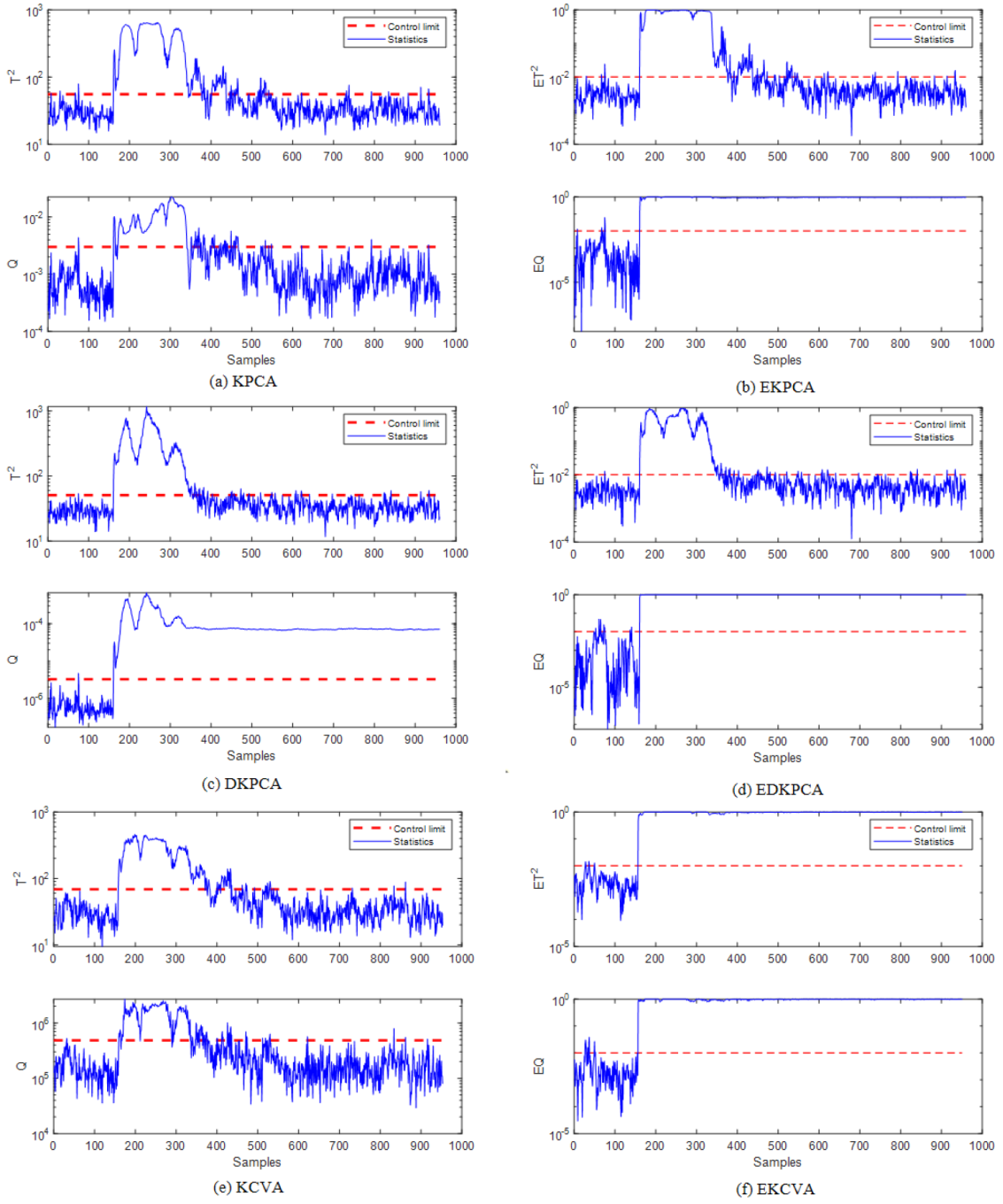


Figure 8. Monitoring results in TEP of Fault 5: (a) KPCA, (b) EKPCA, (c) DKPCA, (b) EDKPCA, (e) KCVA, (f) EKCVA.

Faults 5, 10, and 19 are employed for comparison. Fault 5 involves a step change in condenser cooling water inlet temperature. This fault does not cause a continuous change in the variable as the controller will compensate for the effects of the fault. However, this type of fault can also disrupt the normal operating process. The monitoring results are plotted in Figure 8. It can be found that the monitoring statistics of KPCA, DKPCA, and KCVA have a similar trend where the monitoring statistics will degrade after 10 h, although the step change still was imposed. On the other hand, through ensemble learning, more models are established by EKPCA, EDKPCA, and EKCVA methods. Thus, they can capture this change accurately. The monitoring statistics of EKPCA, EDKPCA, and EKCVA can detect fault 5 more sensitively as shown in Figure 8.

Fault 10 is a randomly varying fault, and the controller cannot fully compensate for the variation of this fault, and the dynamics of the process change under this fault condition. The monitoring results for this fault are shown in Figure 9. The monitoring charts for KPCA, DKPCA, and KCVA do not indicate the occurrence of all faults. Similar to fault 5, EKCVA still achieves the best monitoring results compared with both methods, EKPCA and EDKPCA, where almost all fault samples are detected. EKCVA ET^2 and EQ statistics reach 93.71% and 91.57% of the FDRs, respectively.

The fault information of fault 19 is unknown. From Figure 10, it is obvious that the Q statistic of the EKCVA model has improved significantly over the Q statistic of the other five compared methods, reaching a fault detection rate of 99.55%.

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

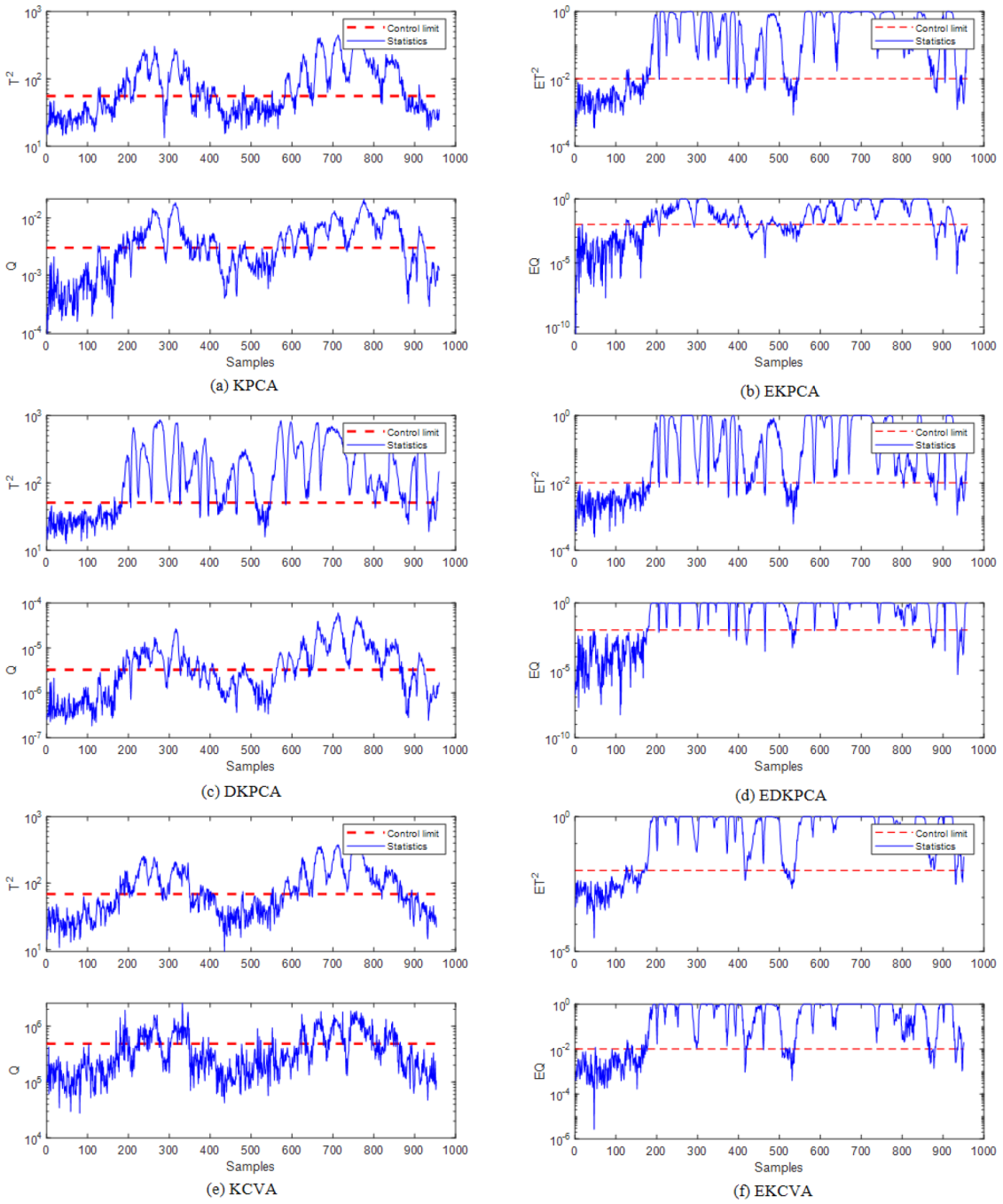


Figure 9. Monitoring results in TEP of fault 10: (a) KPCA, (b) EKPCA, (c) DKPCA, (b) EDKPCA, (e) KCVA, (f) EKCVA.

This indicates that the system deviates from the set operational level. In addition, EDKPCA outperforms EKPCA due to the capture of process dynamics. However, the EKCVA method still has the best performance.

5. CONCLUSION

The findings reported here shed new light on a data-driven dynamic nonlinear process monitoring approach. In the proposed EKCVA method, an ensemble learning strategy is introduced into KCVA. With the aid of ensemble learning, the capability of extracting nonlinearity features of KCVA is enhanced. Additionally, to give a final comprehensive statistic, Bayesian inference is applied to reorganize the monitoring results derived by KCVA which achieves comprehensive and accurate monitoring of process changes. To evaluate the proposed EKCVA method, a numerical example, the industrial TEP benchmark, and the CSTR process are carried out. Experimental results show that the EKCVA method outperforms other methods in terms of detection rate and detection time.

Although the superiority of the presented method has been verified, there are three issues that deserve further investigation. First, only the Gaussian kernel function is adopted in the model establishment. As a local kernel function, the Gaussian kernel function has good interpolation ability. The linear kernel is often used as the global kernel. Therefore, to further improve the feature extraction performance, mixed kernels that contain local and global kernel functions would be studied. Second, this study

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

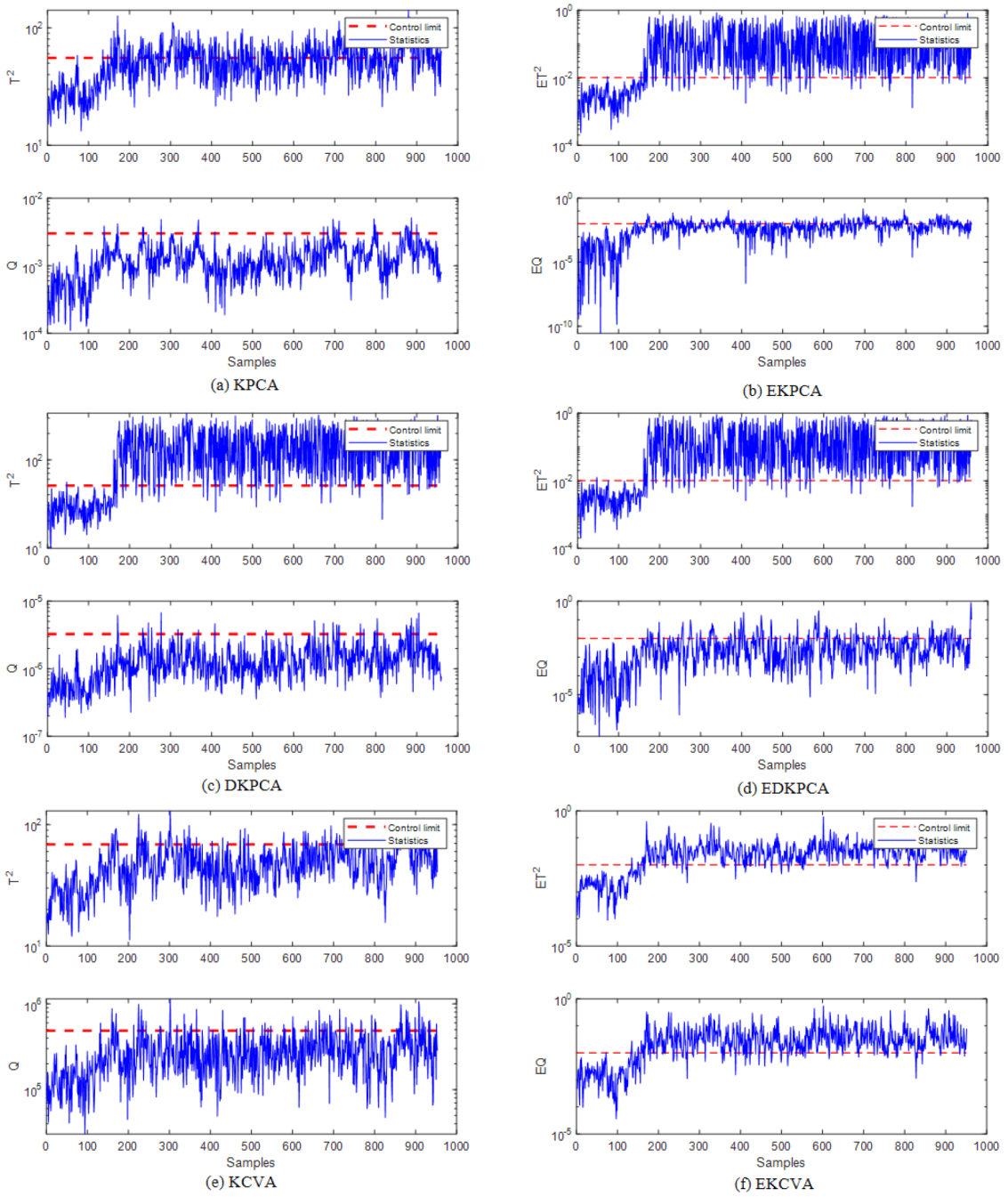


Figure 10. Monitoring results in TEP of fault 19: (a) KPCA, (b) EKPCA, (c) DKPCA, (b) EDKPCA, (e) KCVA, (f) EKCVA.

mainly focuses on the fault detection issue for nonlinear dynamic processes. Fault identification is crucial to determine the root cause of faults. Reconstruction based fault identification has been widely studied in kernel-based process monitoring methods. However, since there is more than one model used in the EKCVA method, the decision making strategy should be another topic to investigate in future work for fault identification.

AUTHOR INFORMATION

Corresponding Author

Ping Wu − *School of Mechanical Engineering & Automation, Zhejiang Sci-Tech University, Hangzhou 310018, P. R. China*; orcid.org/0000-0002-2729-9669; Email: pingwu@zstu.edu.cn

Author

Xuemei Wang − *School of Mechanical Engineering & Automation, Zhejiang Sci-Tech University, Hangzhou 310018, P. R. China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c01892

Notes

The authors declare no competing financial interest.

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

■ ACKNOWLEDGMENTS

This work was supported in part by grants from the National Natural Science Foundation of China under Grant 61703371, Social Development Project of Zhejiang Provincial Public Technology Research (No. LGF19F030004 and LGG21F030015), and in part by Fundamental Research Funds of Zhejiang Sci-Tech University (No: 2021Q024), and in part by Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China, Grant No. ICT2021B45.

■ REFERENCES

(1) Jiang, Q.; Yan, X.; Huang, B. Review and perspectives of data-driven distributed monitoring for industrial plant-wide processes. *Ind. Eng. Chem. Res.* **2019**, *58*, 12899−12912.
(2) Park, Y.-J.; Fan, S.-K. S.; Hsu, C.-Y. A Review on Fault Detection and Process Diagnostics in Industrial Processes. *Processes* **2020**, *8*, 1123.
(3) Rebello, S.; Yu, H.; Ma, L. An integrated approach for real-time hazard mitigation in complex industrial processes. *RELIAB ENG SYST SAFE* **2019**, *188*, 297−309.
(4) Kopbayev, A.; Khan, F.; Yang, M.; Halim, S. Z. Fault detection and diagnosis to enhance safety in digitalized process system. *Comput. Chem. Eng.* **2022**, *158*, 107609.
(5) Ge, Z. Review on data-driven modeling and monitoring for plant-wide industrial processes. *Chemom. Intell. Lab. Syst.* **2017**, *171*, 16−25.
(6) Qin, S. J. Survey on data-driven industrial process monitoring and diagnosis. *Annu. Rev. Control* **2012**, *36*, 220−234.
(7) Amin, M. T.; Khan, F.; Ahmed, S.; Imtiaz, S. A novel data-driven methodology for fault detection and dynamic risk assessment. *Can. J. Chem. Eng.* **2020**, *98*, 2397−2416.
(8) Amin, M. T.; Khan, F.; Ahmed, S.; Imtiaz, S. A data-driven Bayesian network learning method for process fault diagnosis. *PROCESS SAF ENVIRON* **2021**, *150*, 110−122.
(9) Yin, S.; Ding, S. X.; Xie, X.; Luo, H. A review on basic data-driven approaches for industrial process monitoring. *IEEE Trans. Ind. Electron.* **2014**, *61*, 6418−6428.
(10) Yu, H.; Khan, F.; Garaniya, V. A probabilistic multivariate method for fault diagnosis of industrial processes. *Chem. Eng. Res. Des.* **2015**, *104*, 306−318.
(11) Rahoma, A.; Imtiaz, S.; Ahmed, S. Sparse principal component analysis using bootstrap method. *Chem. Eng. Sci.* **2021**, *246*, 116890.
(12) Yu, H.; Khan, F.; Garaniya, V. A sparse PCA for nonlinear fault diagnosis and robust feature discovery of industrial processes. *AlChE J.* **2016**, *62*, 1494−1513.
(13) Ku, W.; Storer, R. H.; Georgakis, C. Disturbance detection and isolation by dynamic principal component analysis. *CHEMOMETR INTELL LAB* **1995**, *30*, 179−196.
(14) Komulainen, T.; Sourander, M.; Jämsä-Jounela, S.-L. An online application of dynamic PLS to a dearomatization process. *Comput. Chem. Eng.* **2004**, *28*, 2611−2619.
(15) Adedigba, S. A.; Khan, F.; Yang, M. Dynamic failure analysis of process systems using principal component analysis and Bayesian network. *Ind. Eng. Chem. Res.* **2017**, *56*, 2094−2106.
(16) Ruiz-Cárcel, C.; Cao, Y.; Mba, D.; Lao, L.; Samuel, R. Statistical process monitoring of a multiphase flow facility. *Control Eng. Pract.* **2015**, *42*, 74−88.
(17) Russell, E. L.; Chiang, L. H.; Braatz, R. D. Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *CHEMOMETR INTELL LAB* **2000**, *51*, 81−93.
(18) Juricek, B. C.; Seborg, D. E.; Larimore, W. E. Fault detection using canonical variate analysis. *Ind. Eng. Chem. Res.* **2004**, *43*, 458−474.
(19) Odiowei, P.-E. P.; Cao, Y. Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. *IEEE Trans. Ind. Inf.* **2010**, *6*, 36−45.
(20) Jiang, B.; Braatz, R. D. Fault detection of process correlation structure using canonical variate analysis-based correlation features. *J. Process Control* **2017**, *58*, 131−138.
(21) Lu, Q.; Jiang, B.; Gopaluni, R. B.; Loewen, P. D.; Braatz, R. D. Sparse canonical variate analysis approach for process monitoring. *J. Process Control* **2018**, *71*, 90−102.
(22) Pilario, K. E. S.; Cao, Y. Canonical variate dissimilarity analysis for process incipient fault detection. *IEEE Trans. Ind. Inf.* **2018**, *14*, 5308−5315.
(23) Li, X.; Yang, X.; Yang, Y.; Bennett, I.; Collop, A.; Mba, D. Canonical variate residuals-based contribution map for slowly evolving faults. *J. Process Control* **2019**, *76*, 87−97.
(24) Zheng, J.; Zhao, C. Enhanced canonical variate analysis with slow feature for dynamic process status analytics. *J. Process Control* **2020**, *95*, 10−31.
(25) Pilario, K. E.; Shafiee, M.; Cao, Y.; Lao, L.; Yang, S.-H. A Review of Kernel Methods for Feature Extraction in Nonlinear Process Monitoring. *Processes* **2020**, *8*, 24.
(26) Ciabattoni, L.; Comodi, G.; Ferracuti, F.; Fonti, A.; Giantomassi, A.; Longhi, S. Multi-apartment residential microgrid monitoring system based on kernel canonical variate analysis. *Neurocomputing* **2015**, *170*, 306−317.
(27) Xiao, S. Locality Kernel Canonical Variate Analysis for Fault Detection. *J. Phys. Conf Ser.* **2019**, *1284*, 012003.
(28) Shang, L.; Liu, J.; Zhang, Y. Efficient recursive kernel canonical variate analysis for monitoring nonlinear time-varying processes. *Can. J. Chem. Eng.* **2018**, *96*, 205−214.
(29) Yu, J.; Wang, K.; Ye, L.; Song, Z. Accelerated Kernel Canonical Correlation Analysis with Fault Relevance for Nonlinear Process Fault Isolation. *Ind. Eng. Chem. Res.* **2019**, *58*, 18280−18291.
(30) Samuel, R. T.; Cao, Y. Kernel canonical variate analysis for nonlinear dynamic process monitoring. *IFAC-PapersOnLine* **2015**, *48*, 605−610.
(31) Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip Rev. Data Min Knowl Discovery* **2018**, *8*, e1249.
(32) Li, N.; Yang, Y. Ensemble kernel principal component analysis for improved nonlinear process monitoring. *Ind. Eng. Chem. Res.* **2015**, *54*, 318−329.
(33) Cui, P.; Zhan, C.; Yang, Y. Improved nonlinear process monitoring based on ensemble KPCA with local structure analysis. *Chem. Eng. Res. Des.* **2019**, *142*, 355−368.
(34) Choi, S. W.; Lee, C.; Lee, J.-M.; Park, J. H.; Lee, I.-B. Fault detection and identification of nonlinear processes based on kernel PCA. *CHEMOMETR INTELL LAB* **2005**, *75*, 55−67.
(35) Nguyen, V. H.; Golinval, J.-C. Fault detection based on kernel principal component analysis. *Eng. Struct.* **2010**, *32*, 3683−3691.
(36) Wu, P.; Lou, S.; Zhang, X.; He, J.; Liu, Y.; Gao, J. Data-Driven Fault Diagnosis Using Deep Canonical Variate Analysis and Fisher Discriminant Analysis. *IEEE Trans. Ind. Inf* **2021**, *17*, 3324−3334.
(37) Jiang, B.; Huang, D.; Zhu, X.; Yang, F.; Braatz, R. D. Canonical variate analysis-based contributions for fault identification. *J. Process Control* **2015**, *26*, 17−25.
(38) Jiang, Q.; Yan, X. Locally weighted canonical correlation analysis for nonlinear process monitoring. *Ind. Eng. Chem. Res.* **2018**, *57*, 13783−13792.
(39) Wu, P.; Ferrari, R. M. G.; Liu, Y.; van Wingerden, J.-W. Data-Driven Incipient Fault Detection via Canonical Variate Dissimilarity and Mixed Kernel Principal Component Analysis. *IEEE Trans. Ind. Inf.* **2021**, *17*, 5380−5390.
(40) Wu, P.; Lou, S.; Zhang, X.; He, J.; Gao, J. Novel Quality-Relevant Process Monitoring based on Dynamic Locally Linear Embedding Concurrent Canonical Correlation Analysis. *Ind. Eng. Chem. Res.* **2020**, *59*, 21439−21457.
(41) Lahdhiri, H.; Taouali, O.; Elaissi, I.; Jaffel, I.; Harakat, M. F.; Messaoud, H. A new fault detection index based on Mahalanobis distance and kernel method. *J. Adv. Manuf. Technol.* **2017**, *91*, 2799−2809.

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921

(42) Amin, M. T.; Imtiaz, S.; Khan, F. Process system fault detection and diagnosis using a hybrid technique. *Chem. Eng. Sci.* **2018**, *189*, 191−211.
(43) Negiz, A.; Çlinar, A. Statistical monitoring of multivariable dynamic processes with state-space models. *AlChE J.* **1997**, *43*, 2002−2020.
(44) Samuel, R. T.; Cao, Y. Fault detection in a multivariate process based on kernel PCA and kernel density estimation. *2014 20th Int. Conf. Automat. Comput.* **2014**, 146−151.
(45) Choi, S. W.; Lee, I.-B. Nonlinear dynamic process monitoring based on dynamic kernel PCA. *Chem. Eng. Sci.* **2004**, *59*, 5897−5908.
(46) Guo, L.; Wu, P.; Lou, S.; Gao, J.; Liu, Y. A multi-feature extraction technique based on principal component analysis for nonlinear dynamic process monitoring. *J. Process Control* **2020**, *85*, 159−172.
(47) Bhadriraju, B.; Kwon, J. S.-I.; Khan, F. Risk-based fault prediction of chemical processes using operable adaptive sparse identification of systems (OASIS). *Comput. Chem. Eng.* **2021**, *152*, 107378.
(48) Downs, J. J.; Vogel, E. F. A plant-wide industrial process control problem. *Comput. Chem. Eng.* **1993**, *17*, 245−255.
(49) Harkat, M.-F.; Mansouri, M.; Nounou, M.; Nounou, H. Fault detection of uncertain nonlinear process using interval-valued data-driven approach. *Chem. Eng. Sci.* **2019**, *205*, 36−45.
(50) Deng, X.; Tian, X.; Chen, S.; Harris, C. J. Deep learning based nonlinear principal component analysis for industrial process fault detection. *2017 Int. Joint Conf. Neural Networks* **2017**, 1237−1243.
(51) Gharahbagheri, H.; Imtiaz, S.; Khan, F. Combination of KPCA and causality analysis for root cause diagnosis of industrial process fault. *Can. J. Chem. Eng.* **2017**, *95*, 1497−1509.
(52) Lee, J.-M.; Yoo, C.; Choi, S. W.; Vanrolleghem, P. A.; Lee, I.-B. Nonlinear process monitoring using kernel principal component analysis. *Chem. Eng. Sci.* **2004**, *59*, 223−234.

https://doi.org/10.1021/acsomega.2c01892
ACS Omega 2022, 7, 18904−18921