IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 16, 2020, accepted September 25, 2020, date of publication October 1, 2020, date of current version October 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3028144

A Batch Process Monitoring Method Using Two-Dimensional Localized Dynamic Support Vector Data Description

XIAOHUI WANG[1,2], YANJIANG WANG[1], XIAOGANG DENG[1,3], (Member, IEEE), AND YUPING CAO[1]

[1]College of Control Science and Engineering, China University of Petroleum, Qingdao 266580, China
[2]College of Application Technology, Qingdao University, Qingdao 266071, China
[3]National Engineering Laboratory of Offshore Geophysical and Exploration Equipment, China University of Petroleum, Qingdao 266580, China

Corresponding author: Xiaogang Deng (dengxiaogang@upc.edu.cn)

This work was supported in part by the Major Scientific and Technological Projects of the China National Petroleum Corporation (CNPC) under Grant ZD2019-183-003, in part by the Fundamental Research Fund for the Central Universities under Grant 20CX02310A, and in part by the Opening Fund of National Engineering Laboratory of Offshore Geophysical and Exploration Equipment under Grant 20CX02310A.

ABSTRACT In order to mine the local behavior and dynamic characteristic of batch process data for effective process monitoring, a two-dimensional localized dynamic support vector data description (TLDSVDD) method is proposed in this article. The main contributions of the proposed method include three aspects. Firstly, considering that batch process variables may behave differently at each operation stage, a two-dimensional localization strategy is designed to mine the local behaviors of process data from the perspective of the variable dimension and the sample dimension. Secondly, for each local data segment, the slow feature analysis is applied to build the local dynamic sub-models, which can monitor the static and dynamic process changes simultaneously. Lastly, the model ensemble strategy based on Bayesian inference is employed and two holistic monitoring statistics are developed to indicate the process running status. The proposed method not only extracts the local process behaviors, but also determines whether the process fault belongs to the dynamic or static change. Finally, one case study on the simulated industrial batch process is carried out to exhibit the method performance.

INDEX TERMS Batch process monitoring, local information mining, slow feature analysis, support vector data description.

I. INTRODUCTION

Batch processes are extensively applied in the modern industry for the multi-variety, customized, and high value-added products. Some typical industrial batch systems include batch distillation, penicillin fermentation process, and semiconductor etch process [1]–[4]. Compared with traditional continuous processes, batch processes are more complicated because they are clearly with multi-stage/multi-phase operation, nonlinear variable relationship, dynamic and non-Gaussian data characteristic [5], [6]. The process complexity leads to high difficulty to ensure the continuity and safety of the process running. Therefore, process monitoring and fault detection technologies are particularly important to batch processes and have attracted widespread attention from scholars in the process control field [7], [8]. Due to the abundant data recorded by the advanced computer control systems, data-driven process monitoring methods are becoming the hot topic [9]–[12]. In past several decades, researchers have put forward many typical methods include multiway principal component analysis (MPCA) [13], multiway partial least squares (MPLS) [14], and multiway canonical variate/correlation analysis (MCVA) [15], [16]. However, these methods assume that the process data obey Gaussian distribution, which are not met strictly in the real applications. Recently, support vector data description (SVDD) has emerged as an effective tool for processing nonlinear and non-Gaussian data [17], [18]. As one typical one-class classification method, SVDD obtains the decision boundary with the minimum volume hyper-sphere for containing most of the training data. The samples outside of the hyper-sphere are viewed as the anomaly points. The SVDD's application to batch process monitoring is firstly discussed by Ge *et al.* [19], which firstly upfolds the multiway batch process data

The associate editor coordinating the review of this manuscript and approving it for publication was Alberto Cano.

and then builds the SVDD monitoring model. The SVDD with multiway data unfolding can be called multiway SVDD (MSVDD). Further, Ge and Song [20] put forward to a bagging SVDD method for batch process monitoring by applying the ensemble learning theory. In order to improve the monitoring accuracy, Wang *et al.* [21] proposed to construct the dynamic hypersphere SVDD model by mixing test samples and training samples. Zhang *et al.* [22] designed a robust SVDD method by improving the calculation of spherical radius. Subsequently, a lot of modified SVDD versions are presented by analyzing the specific batch process characteristics [23], [24].

If the entire batch data are taken to build a holistic SVDD monitoring model, the performance is usually unsatisfactory since it is difficult to depict the whole batch process precisely. Recently, localized SVDD modeling [20], [25], [26] has been one important topic in the SVDD-related process monitoring field. The localized SVDD methods mine the local process behavior and develop multiple local sub-models for elaborate monitoring. The localized SVDD methods are summarized into two classes: multi-phase methods and multi-block methods.

Different from continuous processes with one steady state, batch processes usually involve multiple operation phases. Take the penicillin production as one example, it can be divided into five phases according to the microorganism growth procedure [25]. For different operation phases, the data relationships may be governed by different laws. Therefore, multi-phase SVDD methods are developed to monitor the batch process faults. The multi-phase SVDD method has two steps: phase division and SVDD modeling. The key point of multi-phase SVDD is the phase partition. Ge and Song [20] indicated that the phase partition can be implemented though the model mechanism or data clustering tools. Wang *et al.* [26] designed a multi-phase SVDD method by applying the multiscale fuzzy clustering and sequential phase partitioning. Some researchers made deep studies on the different clustering methods. Luo *et al.* [25] proposed a multi-phase monitoring method by using the K means clustering to divide the operation phases. In Peng *et al.*' work [27], the phase division is carried out by Gaussian mixture model (GMM). Tang and Li [28] defined a repeatability factor to partition the steady and transition stages. Chang *et al.* [29] utilized the affinity propagation (AP) clustering algorithm to distinguish the stages of batch process. The other related studies can be seen in literature [30], [31].

Multi-block analysis is another kind of localized SVDD modeling strategy. As the process faults may only trigger some local variables but not all the variables, one global SVDD model may overwhelm the real data variability. Considering that some faults only influence the correlated variables, Lv and Yan [32] designed a sub-space monitoring layer in the hierarchical SVDD model, which applies the mutual information and K means to perform the variable subspace division. Furthermore, Lv *et al.* [33] utilized the contribution array based on independent component analysis to obtain multiple variable subspaces, and then built the SVDD monitoring model for batch process. Hui and Zhao [34] adopted the mutual information to separate the related and independent variables and built greedy SVDD model to detect batch process faults. Wang *et al.* [35] divided variable blocks by Kullback-Leibler divergence according to statistical characteristics. Hierarchical clustering was applied to build multi-block models in Huang and Yan's work [36].

Although many multi-block and multi-phase SVDD methods have been discussed to deal with the local data behaviors of batch process, many unsolved issues still exist. Firstly, the present localized SVDD methods usually only focus on the one dimension. Multi-block methods are designed only for local variable analysis, while multi-phase methods are just used to deal with local sample characteristic. However, one-dimensional localization is not enough to describe the local data behavior sufficiently. Secondly, these methods only consider the local changes as the static fluctuations. However, batch processes often have strong dynamic property. The dynamic characteristic is the relationship between the historical data and future data. It reflects the information of process variables along the time. How to mine the local dynamic characteristic is also one important problem.

Based on the present studies and the investigated problems, a two-dimensional localized dynamic support vector data description (TLDSVDD) method is presented to improve the basic SVDD based batch process monitoring method. In this proposed method, process local behaviors are analyzed by a two-dimensional localization strategy, which involves both the variable and sample dimensions. Meanwhile, the dynamic data characteristic is extracted with slow feature analysis and two SVDD statistics including the static and dynamic statistics are constructed respectively for each local model. All the local models are integrated by the ensemble learning scheme of Bayesian inference.

II. PRELIMINARIES

A. MULTIWAY DATA UPFOLDING

Different from the traditional continuous processes with two-dimensional training data, batch processes usually have the training data expressed by the three-dimensional matrix, which is also called multiway data matrix. Given the batch process with J process variables and K measurement points for each batch, I batches are gathered to constitute the training data $\boldsymbol{X}(I \times J \times K)$. It is often difficult to train the monitoring model directly with multiway data matrix. Therefore, the common way is to firstly unfold the multiway data into two-dimensional matrix and then build the monitoring model.

There are different unfolding manners used in the present studies [11], [15]. A popular one is the batch-variable unfolding [15], which includes the batch unfolding and variable unfolding. At the first step, this manner unfolds the matrix along the batch direction and performs the z-score normalization. Then at the second step, the normalized data is re-arranged along the variable direction for monitoring models development.

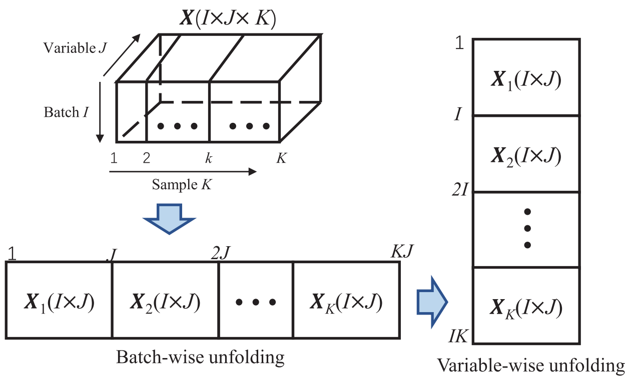

FIGURE 1. Batch-Variable unfolding demonstration.

B. SUPPORT VECTOR DATA DESCRIPTION

SVDD is an effective method for data description and processing. Its key is to find the smallest volume hyper-sphere containing the described objects in high dimensional space. The normal and abnormal samples are distinguished by spherical decision boundary constructed by support vectors.

Given a dataset $\{\boldsymbol{x_i}, i = 1, 2 \cdots l\}$, SVDD is to solve the following optimization problem [17]

$$\min_{R,a,\xi} \quad R^2 + C\sum_{i=1}^{l} \xi_i, \tag{1}$$

$$s.t. \quad \|\Phi(\boldsymbol{x_i}) - \boldsymbol{a}\|^2 \le R^2 + \xi_i\,, \quad \xi_i \ge 0, \quad i = 1, 2 \cdots l, \tag{2}$$

where $\boldsymbol{a}$ is the center, $\Phi(\cdot)$ is the nonlinear mapping function, R is the radius, ξ_i is the slack value of each sample, C is a penalty constant [17]. The Lagrange multiplier α_i is introduced to construct the Lagrange function, which leads to the following constraints as

$$\sum_{i=1}^{l}\alpha_i = 1, \quad \boldsymbol{a} = \sum_{i=1}^{l}\alpha_i\Phi(\boldsymbol{x}_i), \quad 0 \le \alpha_i \le C. \tag{3}$$

The inner product computation of nonlinear function results in the kernel expression as $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \Phi(\boldsymbol{x}_i)\cdot\Phi(\boldsymbol{x}_j)$. Based on the kernel trick, the SVDD objective function is transformed as [17]

$$\max\ L = \sum_{i=1}^{l}\alpha_i K(\boldsymbol{x}_i, \boldsymbol{x}_i) - \sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j K(\boldsymbol{x}_i, \boldsymbol{x}_j), \tag{4}$$

$$s.t. \quad 0 \le \alpha_i \le C, \quad \sum_{i=1}^{l}\alpha_i = 1. \tag{5}$$

When $\alpha_i > 0$ is satisfied, the corresponding sample $\boldsymbol{x}_i^*$ is Support Vector (SV). The hyper-sphere radius is the distance between the center $\boldsymbol{a}$ and any SV $\boldsymbol{x}_i^*$, as shown in the following formula [17]

$$R = \sqrt{1 - 2\sum_{i=1}^{l}\alpha_i K(\boldsymbol{x}_i^*, \boldsymbol{x}_i) + \sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j K(\boldsymbol{x}_i, \boldsymbol{x}_j)}. \tag{6}$$

To determine whether new sample $\boldsymbol{x}_{new}$ is inside the sphere, the distance from new sample $\boldsymbol{x}_{new}$ to center $\boldsymbol{a}$ must be calculated. This distance D_{ist} is defined as the anomaly detection indicator expressed by [17]

$$D_{ist} = \sqrt{1 - 2\sum_{i=1}^{l}\alpha_i K(\boldsymbol{x}_{new}, \boldsymbol{x}_i) + \sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j K(\boldsymbol{x}_i, \boldsymbol{x}_j)}. \tag{7}$$

If the hyper-sphere radius is set as the confidence limit, the statistical meaning is not clear. Therefore, we adopt the kernel density estimation to determine the 99% confidence limit D_{limit} for the distance statistic D_{ist}. If $D_{ist} > D_{limit}$, the new sample $\boldsymbol{x}_{new}$ is considered as the fault sample.

The basic multiway SVDD (MSVDD) method builds a batch process monitoring model by integrating multiway data unfolding and SVDD modeling. It takes multiple batches of data as a whole and establishes a single SVDD monitoring model. This kind of overall modeling strategy may miss some faults which only affect the local data behaviors. Furthermore, SVDD is intrinsically one static modeling technique and can not indicate the process dynamic changes. These shortcomings of MSVDD lead to the limited monitoring performance on the batch process faults.

III. A TWO-DIMENSIONAL LOCALIZED SVDD METHOD WITH DYNAMIC CHARACTERTIC ANALYSIS

In this section, we firstly clarify the motivation for our method and give its basic framework. Then, the technical details, including the two-dimensional localization technique, dynamic characteristic analysis based on slow feature analysis, and the multiple sub-model ensemble strategy, are explained one by one.

A. RESEARCH MOTIVATION AND MODEL FRAMEWORK

As mentioned in the introduction part, the basic SVDD monitoring method for batch processes has the intrinsic limitations because it does not carry out the deep local information mining and omits the process dynamic analysis. In order to develop a new SVDD algorithm to overcome the above limitations, it incurs three related problems as follows. (1) How do we perform deep local information mining? The present works have developed some multi-block and multi-phase methods. However, these methods are restricted to one dimension, which are not enough to describe the local behavior sufficiently. (2) How do we capture the dynamic information in the local data segment? When fault occurs, it may destroy the dynamic or/and static relationships. If the data relationship change can be clearly investigated, it is beneficial to ameliorate monitoring performance. (3) Once multiple local models are developed, how do we integrate them to form a global monitoring result? To inspect the local data behavior is necessary to the model developers. However, the production operators often require one global monitoring chart. Therefore, the final monitoring charts should integrate all the local models rather than demonstrate the numerous local results.

Aiming at the above three problems, a two-dimensional localized dynamic SVDD (TLDSVDD) method is designed for batch process fault detection. The whole method

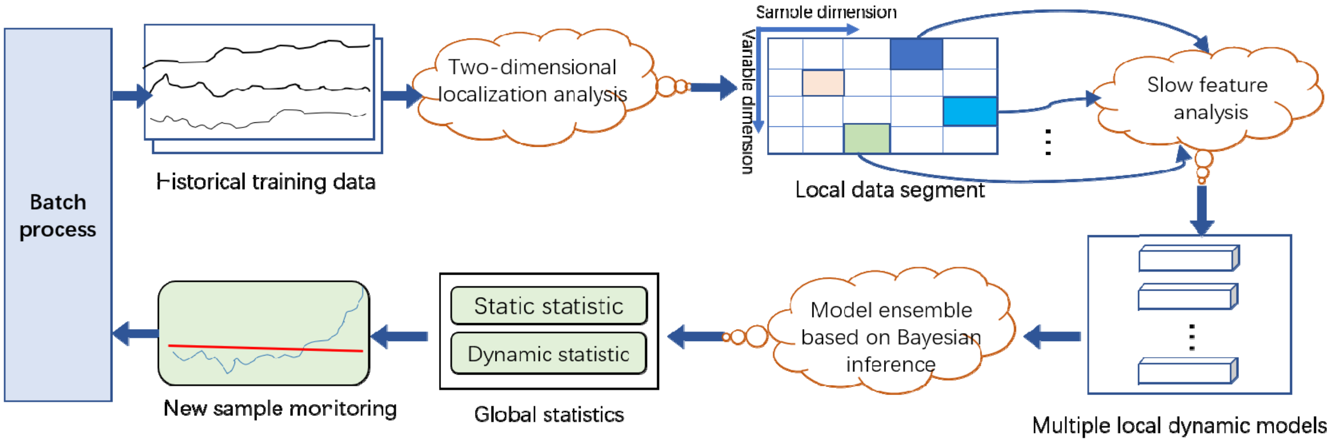

FIGURE 2. **Framework of the proposed TLD-SVDD method.**

framework is depicted as Fig. 2. Firstly, the two-dimensional localization procedure is performed on the training data. In this step, hierarchical clustering based on the distance correlation coefficient is applied in the variable dimension for the variable block division, and then the phases of the variable block are divided by spectral clustering based on mutual information as the local information mining of the sample dimension. Secondly, slow feature analysis method is introduced to capture the process dynamic changes so that the dynamic and static SVDD sub-models are developed respectively. Finally, Bayesian inference is adopted to fuse multiple local SVDD models for the global statistics. The global statistics include the dynamic and static statistics, which indicate the process changes from different perspectives.

B. TWO-DIMENSIONAL LOCALIZATION SCHEME WITH MULTI-BLOCK AND MULTI-PHASE DIVISION

This section describes in detail the two-dimensional localization analysis, including the two steps of local variable analysis for multi-block division and local sample analysis for multi-phase partition.

The first-dimension localization is performed on the variables. This step is to investigate the similarity between variables and divide all the variables into several sub-blocks for monitoring. After this step, the variables with high correlations are divided into a group according to the similarity, and variable coupling in different sub-blocks is reduced. The key point of this step is to determine some criterion to judge the relationships between local variables. Pearson correlation coefficient is a common method to measure the variable relationship, but it only describes the linear correlation between two variables. Even if the Pearson correlation coefficient is equal to zero, it only means that the two variables are not linearly correlated. Different from Pearson correlation coefficient, the distance correlation coefficient can measure the nonlinear correlation between two random vectors [37]. Therefore, this article selects the distance correlation coefficient as the criterion. However, the distance correlation coefficient only provides the measure on the nonlinear correlation degree, but can not indicate the correlation direction. Considering that the cosine metric is one effective tool to determine the variable correlation direction, a new similarity measurement method is proposed for the sub-block division by combining the distance correlation coefficient and cosine similarity. The similarity between two variables z_i and z_j is given by:

$$\boldsymbol{S}_{ij} = \boldsymbol{Dr}_{ij} \times \boldsymbol{Cos}_{ij}, \tag{8}$$

where $\boldsymbol{Dr}_{ij}$ is the distance correlation coefficient between variables; $\boldsymbol{Cos}_{ij}$ is the cosine similarity between variables. Distance correlation between the vector z_i and z_j is computed by [37]

$$\boldsymbol{Dr}_{ij} = \begin{cases} \dfrac{\mathrm{cov}_d(z_i, z_j)}{\sqrt{\mathrm{cov}_d(z_i)\mathrm{cov}_d(z_j)}}, & \mathrm{cov}_d(z_i)\mathrm{cov}_d(z_j) > 0 \\ 0 & \mathrm{cov}_d(z_i)\mathrm{cov}_d(z_j) = 0, \end{cases} \tag{9}$$

where $\mathrm{cov}_d(z_i, z_j)$ represents the distance covariance between the vector z_i and z_j. The distance covariance is calculated by

$$\mathrm{cov}_d^2(z_i, z_j) = \hat{S}_1 + \hat{S}_2 - 2\hat{S}_3, \tag{10}$$

$$\hat{S}_1 = \frac{1}{n^2}\sum_{l=1}^{n}\sum_{k=1}^{n}\left\|z_i^l - z_i^k\right\|\left\|z_j^l - z_j^k\right\|, \tag{11}$$

$$\hat{S}_2 = \frac{1}{n^2}\sum_{l=1}^{n}\sum_{k=1}^{n}\left\|z_i^l - z_i^k\right\| \cdot \frac{1}{n^2}\sum_{l=1}^{n}\sum_{k=1}^{n}\left\|z_j^l - z_j^k\right\|, \tag{12}$$

$$\hat{S}_3 = \frac{1}{n^3}\sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{m=1}^{n}\left\|z_i^l - z_i^m\right\|\left\|z_j^k - z_j^m\right\|. \tag{13}$$

$\boldsymbol{Dr}_{ij} \in [0, 1]$, and $Dr_{ij} = 0$ only if z_i and z_j are independent [37]. Cosine similarity of variables z_i and z_j is defined as

$$\boldsymbol{Cos}_{ij} = \frac{z_i \cdot z_j}{\|z_i\|\,\|z_j\|}. \tag{14}$$

After the similarity matrix $\boldsymbol{S}$ for all variables is obtained. The division of variable blocks is realized as follows:

1) Compute the similarity for variable pairs by (8).
2) Take each variable as one individual cluster and merge closest variables into the same cluster based on similarity analysis.

3) Calculate the similarity between the new cluster and the current cluster and perform step 2) again until all clusters are grouped together.
4) Determine the variable sub-blocks B according to the cluster diagram and give the final clustering results.

Generally, the block number B can be obtained by observing the cluster diagram. For the complicated case, the optimal block number B can be also determined by some quantitative indices such as the Calinski Harabasz index [38].

The second-dimension localization is performed on the samples, which is to divide all the monitored samples into several phases. The real batch processes often involve multiple operating stages. In this work, the improved spectral clustering is utilized to implement the multi-phase division.

The basis of spectral clustering is graph theory, where the data clustering is transformed into the optimal partitioning of undirected graphs [39]. The crux of undirected graphs partitioning is the similarity matrix. The traditional spectral clustering algorithm constructs an adjacent matrix with the fully connected graph, and applies the Gaussian kernel function to define the similarity measure between sample $\boldsymbol{x}_i$ and sample $\boldsymbol{x}_j$.

$$\boldsymbol{W}_{ij} = \exp(-\frac{\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2}{2\sigma^2}), \tag{15}$$

where σ is a parameter of Gaussian kernel width. Intrinsically, the above similarity depends on the Euclidean distance between samples. However, a simple Euclidean distance cannot accurately describe the samples similarity for phase division because it does not consider the variable correlation degree. To build a comprehensive similarity evaluation for batch process phase division, a mixed similarity based on the mutual information and the Euclidean distance is defined as

$$\boldsymbol{W}_{ij} == \beta \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2 + (1-\beta)/\boldsymbol{MI}_{ij}, \tag{16}$$

where the weight coefficient $\beta \in [0,\ 1]$ is to balance the distance and correlation, $\boldsymbol{MI}_{ij}$ denotes the mutual information between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, computed by [40]

$$\boldsymbol{MI}_{ij} = -\frac{1}{2}\log(1-\rho_{ij}^2), \tag{17}$$

where ρ_{ij} is cross correlation coefficient of sample $\boldsymbol{x}_i$ and sample $\boldsymbol{x}_j$.

A degree matrix $\boldsymbol{\Omega}$ is defined as

$$\boldsymbol{\Omega}_{ii} = \sum_j \boldsymbol{W}_{ij}. \tag{18}$$

Further, the Laplace matrix $\boldsymbol{L}$ is given by

$$\boldsymbol{L} = \boldsymbol{\Omega} - \boldsymbol{W}. \tag{19}$$

The eigenvalue decomposition on the Laplace matrix $\boldsymbol{L}$ leads to the eigenvectors corresponding to the k maximum eigenvalues, and the K-means clustering algorithm is applied to these eigenvectors. Based on this, the whole data set is partitioned into several phases. It should be noted that the focus of this article is to design a two-dimensional localization strategy, but not just to emphasize the improvement of one localization analysis method.

C. DYNAMIC CHARACTERISTIC ANLYSIS USING SLOW FEATURE ANALYSIS

Slow feature analysis (SFA) is a recently rising dynamic data analysis technology, which extracts slowly changing features from the temporal signals for further statistical analysis. The present studies have demonstrated its effectiveness in dynamic data analysis [11].

Given the input signal $\boldsymbol{x}(t) = [x_1(t), \cdots, x_J(t)]^{\mathrm{T}}$ from J sensors with $t \in [t_0, t_1]$, SFA is to find a transformation function $\boldsymbol{g}(\boldsymbol{x}(t)) = [g_1(\boldsymbol{x}(t)), g_2(\boldsymbol{x}(t)) \cdots g_J(\boldsymbol{x}(t))]^T$ to make the output signal $s_j(t) = g_j(\boldsymbol{x}(t)), j \in \{1, 2, \cdots J\}$ vary as slowly as possible. An optimal objective function of SFA is to minimize the temporal variation of the output signal, as follows [41]:

$$\min \Delta(s_j) = \min\left\langle \dot{s}_j^2 \right\rangle, \tag{20}$$

under the constraints

$$\left\langle s_j \right\rangle = 0, \tag{21}$$

$$\left\langle s_j^2 \right\rangle = 1, \tag{22}$$

$$\forall i \neq j \quad \left\langle s_i \cdot s_j \right\rangle = 0, \tag{23}$$

where $\Delta(\cdot)$ is the temporal variation of the output signal, $\dot{s}_j$ is the first-order derivative of the output signal s_j. Here, $\dot{s}_j(t) = s_j(t) - s_j(t-1)$. The angle brackets express temporal averaging, that is

$$\langle f \rangle = \frac{1}{t_1 - t_0}\int_{t_0}^{t_1} f(t)dt. \tag{24}$$

The constraints would avoid a constant solution and make solution less arbitrary. If $g_j(\cdot)$ is a linear transform, $s_j(t) = g_j(\boldsymbol{x}(t)) = \boldsymbol{v}_i^T\boldsymbol{x}(t)$ is expressed by a transform vector $\boldsymbol{v}_i$. The objective function is then re-written as

$$\min\left\langle \dot{s}_j^2 \right\rangle = \min \boldsymbol{v}_j^T \left\langle \dot{\boldsymbol{x}}(t)\dot{\boldsymbol{x}}(t)^T \right\rangle \boldsymbol{v}_j = \min \boldsymbol{v}_j^T \boldsymbol{A}\boldsymbol{v}_j, \tag{25}$$

$$s.t \ \left\langle s_j^2 \right\rangle = \boldsymbol{v}_j^T \left\langle \boldsymbol{x}(t)\boldsymbol{x}(t)^T \right\rangle \boldsymbol{v}_j = \boldsymbol{v}_j^T \boldsymbol{B}\boldsymbol{v}_j = 1, \tag{26}$$

where $\boldsymbol{A}$ is a covariance matrix of $\dot{\boldsymbol{x}}(t)$, $\boldsymbol{B}$ is a covariance matrix of $\boldsymbol{x}(t)$. The matrix $\boldsymbol{V} = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_J]$ is obtained by solving the generalized eigenvalue problem

$$\boldsymbol{AV} = \boldsymbol{BV\Lambda}, \tag{27}$$

where $\boldsymbol{\Lambda}$ is the eigenvalue matrix. Some dynamic information is extracted by SFA in the phase matrix, that is

$$\boldsymbol{S}_p = \boldsymbol{X}_p \boldsymbol{V}. \tag{28}$$

Here, the matrix $\boldsymbol{S}_p$ expresses static information of each phase data for batch processes. The temporal slow feature $\dot{\boldsymbol{S}}_p$ represents the speed of signal variance which describes the dynamic information of every phase data.

To apply the SVDD modeling on the static part $\boldsymbol{S}_p$ and the dynamic part $\dot{\boldsymbol{S}}_p$ can bring the static monitoring statistic SD_{ist} and the dynamic monitoring statistic DD_{ist}, respectively.

D. MULTIPLE LOCAL MODELS ENSEMBLE BY BAYESIAN INFERENCE

By combining the two-dimensional localization processing and SFA modeling, a series of sub-block monitoring models are developed for each block of each phase. Therefore, at each operation phase, we can obtain the corresponding sub-block monitoring statistics $\{SD_{ist}^{(b)},\ DD_{ist}^{(b)},\ 1 \leq b \leq B\}$. However, a holistic monitoring indication is needed for judging the system status of the whole process. In this article, we adopt the Bayesian strategy to fuse the multiple sub-blocks.

For a given testing sample $\boldsymbol{x}_t$ at the t-th sample instant, the static and dynamic fault probability induced by the b-th sub-model at the corresponding phase is defined as [42]

$$P_S^{(b)}(F|\boldsymbol{x}_t) = \frac{P(F)P_S^{(b)}(\boldsymbol{x}_t|F)}{P(N)P_S^{(b)}(\boldsymbol{x}_t|N) + P(F)P_S^{(b)}(\boldsymbol{x}_t|F)}, \quad (29)$$

$$P_D^{(b)}(F|\boldsymbol{x}_t) = \frac{P(F)P_D^{(b)}(\boldsymbol{x}_t|F)}{P(N)P_D^{(b)}(\boldsymbol{x}_t|N) + P(F)P_D^{(b)}(\boldsymbol{x}_t|F)}, \quad (30)$$

where $P(F)$ and $P(N)$ are the prior probabilities of fault and normal state, respectively. With the confidence level $\alpha = 0.01$, $P(N) = 1 - \alpha = 0.99$ and $P(F) = \alpha = 0.01$. The occurring probabilities of the testing sample under the fault and normal state are defined as [42], [43]

$$P_S^{(b)}(\boldsymbol{x}_b|F) = \exp(-\frac{SD_{limit}^{(b)}}{SD_{ist}^{(b)}}), \quad (31)$$

$$P_S^{(b)}(\boldsymbol{x}_b|N) = \exp(-\frac{SD_{ist}^{(b)}}{SD_{limit}^{(b)}}), \quad (32)$$

$$P_D^{(b)}(\boldsymbol{x}_b|F) = \exp(-\frac{DD_{limit}^{(b)}}{DD_{ist}^{(b)}}), \quad (33)$$

$$P_D^{(b)}(\boldsymbol{x}_b|N) = \exp(-\frac{DD_{ist}^{(b)}}{DD_{limit}^{(b)}}), \quad (34)$$

where $SD_{limit}^{(b)}, DD_{limit}^{(b)}$ are the control limits corresponding to $SD_{ist}^{(b)},\ DD_{ist}^{(b)}$, respectively.

To combine the monitoring statistic of all sub-blocks, the final global statistics are calculated by

$$SBIC = \sum_{b=1}^{B} w_b P_S^{(b)}(F|\boldsymbol{x}_t)\frac{P_S^{(b)}(\boldsymbol{x}_t|F)}{\sum\limits_{i=1}^{B} P_S^{(i)}(\boldsymbol{x}_t|F)}, \quad (35)$$

$$DBIC = \sum_{b=1}^{B} w_b P_D^{(b)}(F|\boldsymbol{x}_t)\frac{P_D^{(b)}(\boldsymbol{x}_t|F)}{\sum\limits_{i=1}^{B} P_D^{(i)}(\boldsymbol{x}_t|F)}, \quad (36)$$

where w_b is a weight coefficient used to highlight the fault blocks, which is defined as

$$w_b = \frac{Con_b}{\sum\limits_{b=1}^{B} Con_b} B, \quad (37)$$

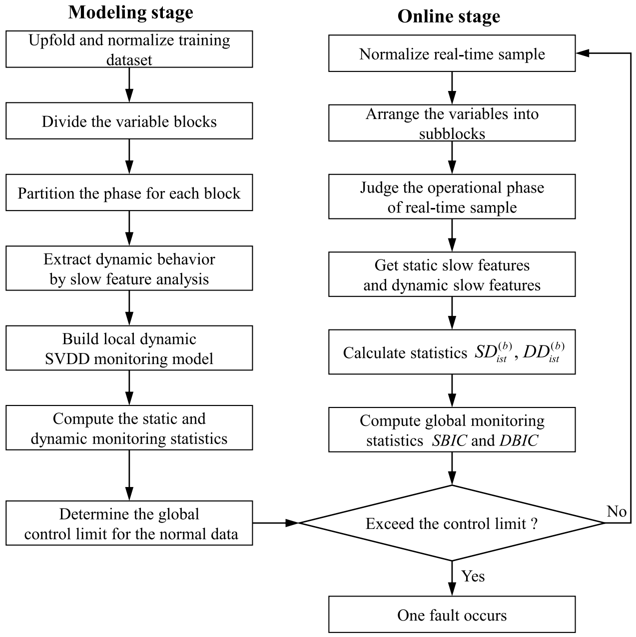

FIGURE 3. Implementation of the proposed TLD-SVDD method.

where Con_b expresses the contribution degree, defined by

$$Con_b = \begin{cases} 1 & D_{ist}^{(b)} > D_{limit}^{(b)} \\ 0 & D_{limit}^{(b)} \geq D_{ist}^{(b)}. \end{cases} \quad (38)$$

Monitoring threshold $SBIC_{limit}$ and $DBIC_{limit}$ are determined by the kernel density estimation.

E. PROCESS MONITORING PROCEDURE

The whole process monitoring includes two stages: modeling stage and online detection stage. The flow chart of the proposed TLD-SVDD method is shown in Fig. 3

1) MODELING STAGE

1) Gather the normal data from multiple batches as the training dataset $\boldsymbol{X}$ and upfold it.
2) Carry out the two-dimensional localization processing, which leads to the phase division and the variable block partition results.
3) For each block at each operation phase, perform the slow feature analysis to build the local dynamic monitoring model.
4) Compute the static and dynamic monitoring statistics $\{SD_{ist}^{(b)},\ DD_{ist}^{(b)},\ 1 \leq b \leq B\}$ for each block.
5) Determine global control limit according to the normal data.

2) ONLINE STAGE

1) Obtain the real-time sample and preprocess it.
2) Divide the variable group according the variable-dimension localization results.
3) Judge the operational phase of real-time sample for each variable sub-block according to the sample-dimension localization results.
4) Conduct SFA to extract static slow features and temporal slow features.

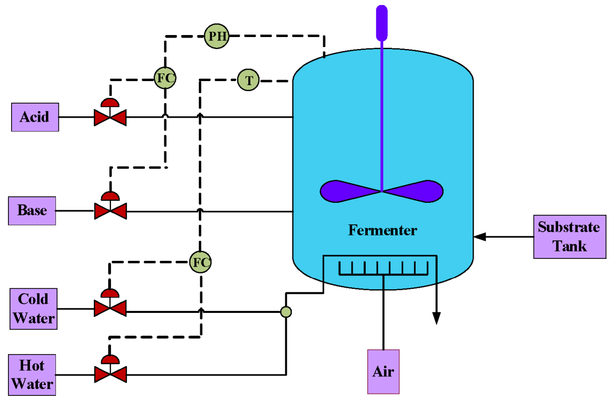

FIGURE 4. Schematic diagram of penicillin fermentation process.

5) Calculate the static and dynamic monitoring statistic $SD_{ist}^{(b)}$, $DD_{ist}^{(b)}$, and combine them through Bayesian inference to get global monitoring statistics $SBIC$ and $DBIC$.

By investigating the process monitoring procedure of the proposed TLDSVDD method, the main advantages of the proposed method lie in two aspects. On the one hand, a two-dimensional localization strategy is designed to mine the local behaviors of process data. On the other hand, for each local data segment, the slow feature analysis is applied to build the local dynamic sub-models, which can monitor the static and dynamic process changes simultaneously. With these advantages, TLDSVDD provides the greater potential to monitor the batch process faults compared with the basic MSVDD, multi-block SVDD, and multi-phase SVDD methods.

IV. CASE STUDIES

The proposed strategy based on TLDSVDD is tested with the penicillin fermentation process. It is compared with MPCA, MSVDD, MPSVDD, MBSVDD and MBMPSVDD. Here, MSVDD is the basic multiway SVDD method. MPSVDD is the multi-phase SVDD method by combining multiple phases partition. MBSVDD is the multi-block SVDD method, which integrates the variable block division and MSVDD. This method only involves the localization analysis at the variable dimension. MBMPSVDD is the improved version of MBSVDD, which further considers the sample dimension localization. However, it does not involve the dynamic data analysis. Among these methods, our proposed TLDSVDD method utilizes the local and dynamic data information sufficiently.

A. THE PENICILLIN PRODUCTION SYSTEM

Penicillin is a well-known antibiotic in modern medicine. Penicillin fermentation is a fed-batch fermentation process with nonlinear, dynamic, and multi-phase characteristics. By controlling the PH value and the temperature in the fermentation reactor, the reaction can be carried out under optimal conditions. A schematic diagram of penicillin fermentation process is shown in Fig. 4. The penicillin simulation platform Pensim 2.0 [44] generates 20 batches of normal production

TABLE 1. The monitored process variables.

No.	Variable description	No.	Variable description
1	Aeration rate	6	Co2 concentration
2	Agitator power	7	PH
3	Substrate feed temperature	8	Reactor temperature
4	Culture volume	9	Generated heat
5	Dissolved oxygen concentration	10	Coiling water flow rate

TABLE 2. The tested Fault patterns.

Fault	Fault Description
F1	A step change with the amplitude of +0.2% in the aeration rate between 100 and 150 h
F2	A ramp change with the slope of -0.2L/h^2 in aeration rate between 100 and 200 h
F3	A step change with the amplitude of +5% in the agitator power between 10 and 40 h
F4	A ramp change with the slope of -0.02W/h in the agitator power between 100 and 400 h
F5	A step change with the amplitude of +10% in the substrate feed rate between 100 and 400 h
F6	A ramp change with the slope of -0.02L/h^2 in the substrate feed rate between 150 and 250 h

data as modeling data and 10 batches of data as test data. The reaction time of each batch is 400 h, and the samples are recorded every 0.5 hour. 10 process variables, as shown in Table 1, are selected as the monitored variables.

The fault patterns of penicillin fermentation are shown in Table 2. Fault F1, F3, F5 are step change of process variables, while fault F2, F4, F6 are ramp change of process variables. In fact, all these faults belong to the process condition changes. As the Pensim 2.0 only provides these faults, no other kinds of faults are discussed in this article. In the following monitoring charts, control limits are indicated by dotted lines and the monitoring statistics are plotted by solid lines. If the statistics of five consecutive sampling moments exceed the control limit, a fault is judged and the first sampling moment is thought to the fault detection time. Two indices, including false alarming rate (FAR) and fault detection rate (FDR), are used to evaluate different methods, which are defined as follows.

$$FAR = \frac{N_{n\to f}}{N_n} \times 100\%, \tag{39}$$

$$FDR = \frac{N_{f\to f}}{N_f} \times 100\%, \tag{40}$$

where N_n, N_f are the numbers of normal and faulty samples, respectively, $N_{n\to f}$ is the number of normal samples exceeding the confidence limit, and $N_{f\to f}$ is the number of faulty samples exceeding the confidence limit from the first detected sample.

B. MONITORING MODEL DEVELOPMENT

First, the normal training dataset is preprocessed. The three-dimensional matrix of training set $\boldsymbol{X}(20 \times 10 \times 800)$ is unfolded along the batch direction into a two-dimensional matrix $\overline{\boldsymbol{X}}(20 \times 8000)$. After it is normalized, $\overline{\boldsymbol{X}}$ is unfolded along the variable direction into $\overline{\overline{\boldsymbol{X}}}(16000 \times 10)$.

Next, the local information of variable dimension is mined through multi-block division method. The improved

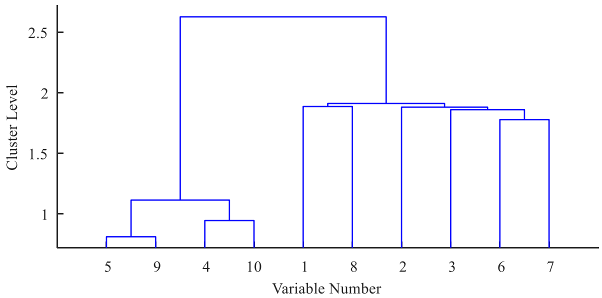

FIGURE 5. Hierarchy dendrogram of process variables.

TABLE 3. Sample phase division results.

Phase Number	1st block	2nd block	3th block	All variable
Phase 1	1-160	1-185	1-151	1-151
Phase 2	161-287	186-313	152-257	152-231
Phase 3	288-486	314-458	258-361	232-509
Phase 4	487-800	459-800	362-800	510-800

hierarchical clustering method is used to cluster all variables into three sub-blocks. The result of multi-block division is given in Fig. 5. The first sub-block contains the variables 2, 3, 6, 7. The second sub-block has the variables 1, 8. The third sub-block is made up of variables 4, 5, 9, 10.

Then the local information at the data sample dimension is mined by the phase division. The spectral clustering method based on mutual information is applied to the phase division of sub-data sets. The parameter β in the similarity matrix is set as 0.5 by cross validation. The division results of operational phases are shown in Table 3, which are consistent with the four physical stages of penicillin fermentation process. A validity index S_Dbw [45], called scattering and density between the clusters, is used to evaluate the results of clustering algorithms. The optimal phase number for the selected historical data is 4 by calculating the value of S_Dbw. We also apply this index to compare different clustering methods including k-means, spectral clustering, and MI-spectral clustering respectively. The smaller value indicates the well-separated and compact cluster. The clustering evaluation indices of phase division with k-means algorithm, spectral clustering algorithm, and MI-spectral clustering algorithm are 1.8340, 0.1708, 0.1310, respectively. By these results, the improved spectral clustering method achieves better separation between the clusters and has lower average scattering within a cluster.

C. FAULT DETECTION RESULTS

In this section, the proposed method and other five methods are applied in fault detection. For all the SVDD models, Gaussian kernel function is selected with the kernel width $\sigma = 6$, and the penalty constant C is set as 0.08. In MPCA, the number of principal components is selected according to the principle that the contribution rate of principal component variance is above 95%, and the confidence limit of 99% is selected to determine the threshold value of monitoring statistics.

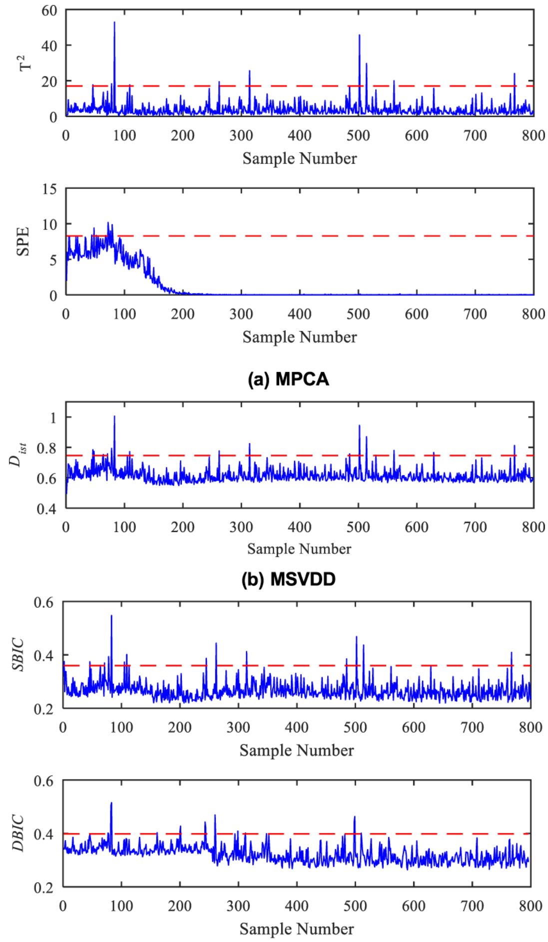

FIGURE 6. Process monitoring results on normal data.

Firstly, the normal test set is monitored. The simulation results of MPCA, MSVDD and TLDSVDD are compared in Fig. 6. By this figure, most of monitoring statistics are below the confidence limit, which means the batch process runs under normal operating conditions. In Fig. 6(a), MPCA has the false alarm rates (FARs) of 1.37% and 1.13% for T^2 and SPE, respectively. The FAR of MSVDD method is 2.67%, while the FARs of TLDSVDD method are 1.88% and 2.14% for the static and dynamic statistics, respectively. Three other methods of MPSVDD, MBSVDD and MBMPSVDD are also tested. Their FARs are 1.5%, 2%, and 1.38%, respectively. All the methods can give the ideal monitoring of the normal operations.

Fault F1 is an aeration rate fault with small step change. The fault detection results for F1 are presented in Fig. 7. MPCA method is a classic method for batch process fault detection. Its monitoring results are shown in Fig. 7(a), where the T^2 statistic detects the fault F1 at 122 h with the fault detection rate (FDR) of 30%. However, MPCA SPE statistic doesn't detect the fault F1 effectively, and there are some false alarms during the non-fault period. In Fig. 7(b), the FDR of MSVDD method is 38%, which is a little higher than MPCA method. Considering the multi-phase nature of the

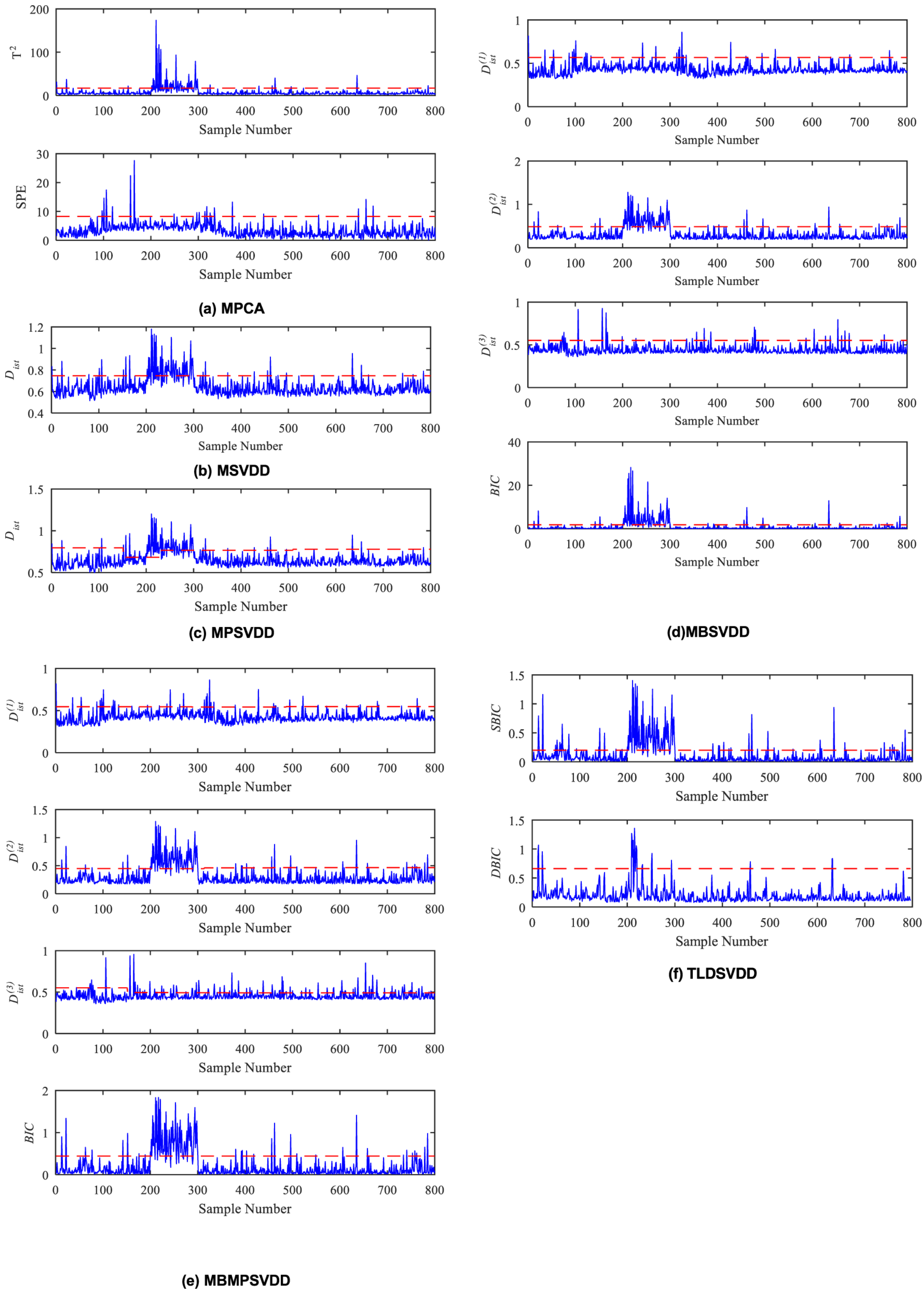

FIGURE 7. Process monitoring results on fault F1.

batch process, the penicillin fermentation process is divided into four phases for monitoring, and the monitoring results of MPSVDD are shown Fig. 7(c). After phase division, each phase has different change trends, and the control limit of the local model is tighter than that of the global model. These factors lead to the higher FDR of 75% in the MPSVDD monitoring chart. When MBSVDD method is applied, the variable groups is divided by the improved hierarchical clustering, and then the SVDD sub-models are constructed for each variable block. The fault detection results obtained by

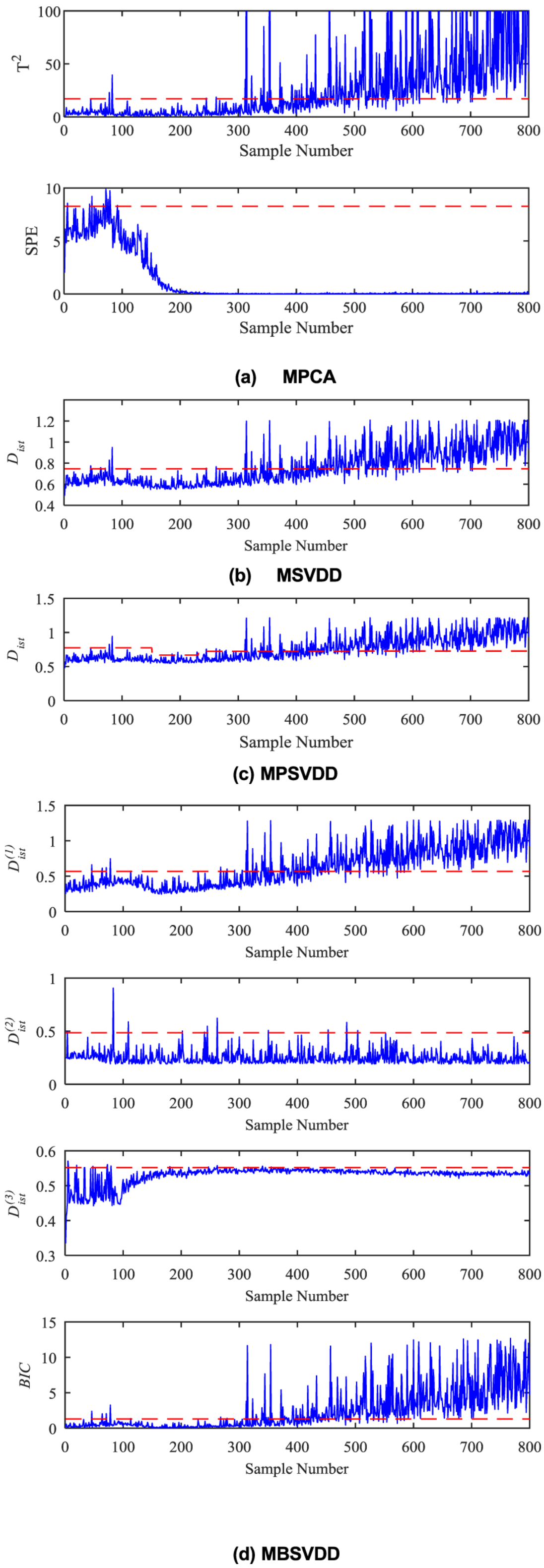

FIGURE 8. Process monitoring results on fault F4.

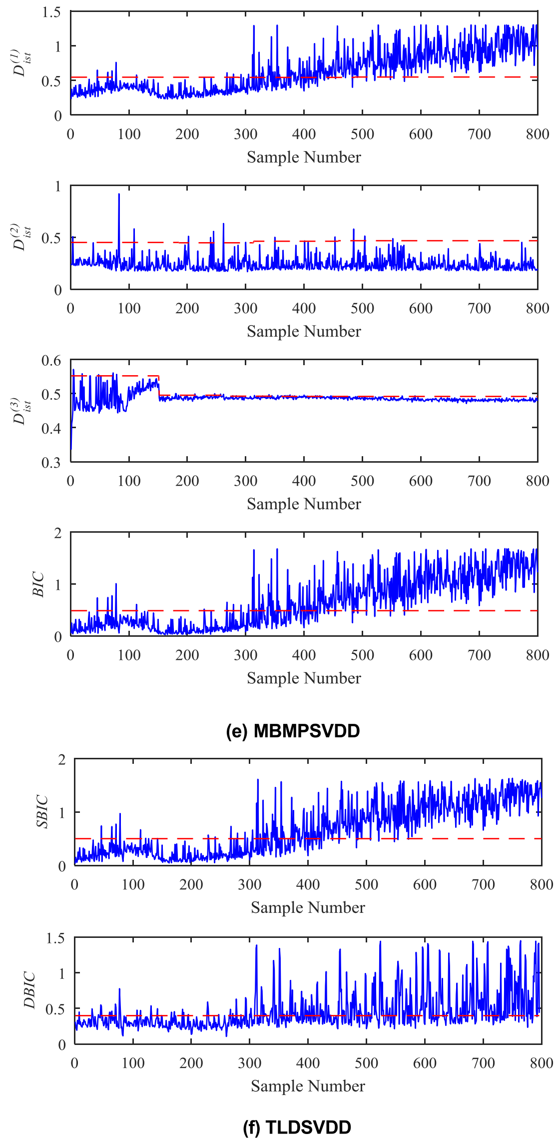

FIGURE 8. ***(Continued)*** **Process monitoring results on fault F4.**

MBSVDD are demonstrated in Fig. 7(d), where FDR is increased to 78%. The statistics in the first variable sub-block and the third variable sub-block do not exceed the threshold, showing that the variables contained in the two groups have nothing to do with the fault F1. The fault is detected only in the second sub-block, which contains the aeration rate and reactor temperature. The multi-block method can not only improve the efficiency of fault detection, but also help to identify the causes of faults. On the basis of MBSVDD method, MBMPSVDD divides the sub-block into multiple phases, and then the individual monitoring is performed through the local SVDD model. A global monitoring statistic *BIC* of MBMPSVDD method is shown in Fig. 7(e). The FDR of MBMPSVDD method is further prompted to 85%. In the MBMPSVDD monitoring charts, monitoring performance has been improved significantly due to the two-dimensional localization analysis. Considering the dynamic behavior with the transient states, the static fault detection rate of TLDSVDD method reaches 86% in Fig. 7(f). The SFA algorithm extracts the slowly changing components to mine more dynamic information. The alarm signal is firstly obtained at 101 h. This means that the proposed method could capture the change of dynamic behavior earlier. It can be observed that most of dynamic statistic *DBIC* values are

TABLE 4. Comparison of fault detection rates of the six methods.

Fault	*MPCA*		*MSVDD*	*MPSVDD*	*MBSVDD*	*MBMPSVDD*	*TLDSVDD*	
	T^2	*SPE*					*SBIC*	*DBIC*
F1	30%	0	38%	75%	78%	85%	86%	0
F2	94%	0	94%	94%	94%	94.5%	97%	82%
F3	100%	100%	100%	100%	100%	100%	100%	100%
F4	45.5%	0	49.67%	53.17%	57.17%	63.83%	65.33%	35.8%
F5	100%	0	100%	100%	100%	100%	100%	100%
F6	59%	68%	69%	70.5%	79.5%	86%	97%	69.5%
Average	71.42%	28%	75.11%	82.11%	84.78%	88.22%	90.89%	64.55%

under the control limit, and only a few samples exceed the threshold at the beginning of the failure. This explains that the fault F1 only destroys the static data relationship and does not influence the system dynamic behavior seriously. To compare all these six methods, TLDSVDD has the earlier fault detection with the highest detection rate for the fault F1. Therefore, the TLDSVDD method is superior to the other five methods because of the considerations of the local information and dynamic information.

Fault F4 is the ramp fault with slow drift of agitator power. Fig. 8(a) shows that the MPCA T^2 statistic detects the fault at 249 h, and the FDR is 45.5%. The MPCA *SPE* statistic doesn't detect the fault F4 successfully. By contrast, the MSVDD monitoring statistic D_{ist} in Fig. 8(b) detects the fault at 236.5 h with the FDR of 49.67%. The MSVDD method detects faults earlier than the MPCA method. As is shown in Fig. 8(c), the fault is detected at 227.5h by the MPSVDD method and FDR is further increased to 53.17%. MBSVDD method does a little better than MPSVDD. The monitoring statistic of MBSVDD alarms the fault at 215 h with a higher FDR of 57.17%. Through the partition of variable blocks, the influence of local variables is highlighted and the MBSVDD method can detect the fault faster. Fig. 8 (e) is the result of the MBMPSVDD method, which gives the consecutive alarms from the 189 h. The FDR of MBMPSVDD is 63.83%, exceeding the MBSVDD and MPSVDD methods. In Fig. 8(f), the detection times of two TLDSVDD statistics are 189 h and 226 h, respectively. Especially, the *SBIC* gives an earlier fault detection. Furthermore, the FDRs of two statistics are 65.33% and 35.8%, respectively. As can be seen from the above comparison, the proposed TLDSVDD method can detect the fault F4 earlier, and its FDR is higher than other methods.

To clearly compare the monitoring performance of the six methods, more fault detection cases are listed in Table 4. As the basic methods, MPCA and MSVDD give the poor monitoring. Their average FDRs are below 80%. With the local data characteristic analysis, MBSVDD and MPSVDD achieve the clear improvement of average FDR. The average FDRs of these two methods exceed 80%. As the combination of MBSVDD and MPSVDD, MBMPSVDD obtains the average FDR of 88.22%. The proposed TLDSVDD not only makes use of all local variable and sample information, but also highlights the dynamic data changes. The average FDR by TLDSVDD *SBIC* statistic is 90.89%, which is the highest among the six methods. Generally, TLDSVDD can describe process information more sufficiently and achieves best fault detection performance.

In the monitoring results of TLDSVDD, two monitoring statistics *SBIC* and *DBIC* are applied. For the small step fault F1, it only breaks the static relationship. Therefore, *SBIC* gives the clear alarms but *DBIC* indicates no changes. As to the ramp faults F2, F4, and F6, they affect both relationships but the static relationship is destroyed more seriously. Therefore, the *SBIC* does better than *DBIC* in terms of fault detection. But for the faults F3 and F5 with the significant step changes, they destroy the static and the dynamic relationships simultaneously so that both *SBIC* and *DBIC* give the clear fault detection.

V. CONCLUSION

In this article, a two-dimensional localized dynamic SVDD monitoring strategy is proposed for batch processes. This method can deal with both the local information mining and dynamic data analysis effectively. A two-dimensional localization strategy is designed to improve the SVDD modeling by combining the variable sub-block division and the phase partition. SFA is applied to extract the static and dynamic features as the input of the SVDD monitoring model. Compared with the basic SVDD with original variables as the input, the SFA features can provide more plentiful information about process changes. Application to the penicillin fermentation process shows that the proposed method can detect faults sensitively and provide the meaning monitoring results for further fault source diagnosis. However, some related problems are also noteworthy. On the one hand, the local model is developed by the linear SFA method, which omits the possible nonlinearity in the local models. On the other, this paper assumes that all the batches are with the same time length, which may be not satisfied in the real applications. Therefore, the nonlinear local modeling and uneven batch process monitoring are two valuable problem deserving the deep studies in the future.

REFERENCES

[1] Z. Ge, Z. Song, S. X. Ding, and B. Huang, "Data mining and analytics in the process industry: The role of machine learning," *IEEE Access*, vol. 5, pp. 20590–20616, Sep. 2017.

[2] S. Zhang, C. Zhao, and F. Gao, "Incipient fault detection for multiphase batch processes with limited batches," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 1, pp. 103–117, Jan. 2019.

[3] A. Friedel, E. von Harbou, K. Munnemann, and H. Hasse, "Online process monitoring of a batch distillation by medium field NMR spectroscopy," *Chem. Eng. Sci.*, vol. 219, pp. 1–8, Feb. 2020.

[4] Q. Jiang, S. Yan, H. Cheng, and X. Yan, "Local-global modeling and distributed computing framework for nonlinear plant-wide process monitoring with industrial big data," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Apr. 21, 2020, doi: 10.1109/TNNLS.2020.2985223.

[5] P. Chang, C. Ding, and Q. Zhao, "Fault diagnosis of microbial pharmaceutical fermentation process with non-Gaussian and nonlinear coexistence," *Chemometrics Intell. Lab. Syst.*, vol. 199, pp. 1–9, Feb. 2020.

[6] J. Zhu, Y. Wang, D. Zhou, and F. Gao, "Batch process modeling and monitoring with local outlier factor," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 4, pp. 1552–1565, Jul. 2019.

[7] Y. Liu, L. Yuan, S. Huang, D. Huang, and B. Liu, "Integrated design of monitoring, analysis and maintenance for filamentous sludge bulking in wastewater treatment," *Measurement*, vol. 155, pp. 1–11, Apr. 2020.

[8] Z. Ge, Z. Song, and F. Gao, "Review of recent research on data-based process monitoring," *Ind. Eng. Chem. Res.*, vol. 52, no. 10, pp. 3543–3562, Mar. 2013.

[9] X. Deng, P. Cai, Y. Cao, and P. Wang, "Two-step localized kernel principal component analysis based incipient fault diagnosis for nonlinear industrial processes," *Ind. Eng. Chem. Res.*, vol. 59, no. 13, pp. 5956–5968, Mar. 2020.

[10] X. Deng, X. Tian, S. Chen, and C. J. Harris, "Nonlinear process fault diagnosis based on serial principal component analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 560–572, Mar. 2018.

[11] H. Zhang, X. Tian, X. Deng, and Y. Cao, "Batch process fault detection and identification based on discriminant global preserving kernel slow feature analysis," *ISA Trans.*, vol. 79, pp. 108–126, Aug. 2018.

[12] Q. Jiang, S. Yan, X. Yan, H. Yi, and F. Gao, "Data-driven two-dimensional deep correlated representation learning for nonlinear batch process monitoring," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2839–2848, Apr. 2020.

[13] L. Zhaomin, J. Qingchao, and Y. Xuefeng, "Batch process monitoring based on multisubspace multiway principal component analysis and time-series Bayesian inference," *Ind. Eng. Chem. Res.*, vol. 53, no. 15, pp. 6457–6466, Apr. 2014.

[14] Q. Jiang, X. Yan, H. Yi, and F. Gao, "Data-driven batch-end quality modeling and monitoring based on optimized sparse partial least squares," *IEEE Trans. Ind. Electron.*, vol. 67, no. 5, pp. 4098–4107, May 2020.

[15] Y. Cao, Y. Hu, X. Deng, and X. Tian, "Quality-relevant batch process fault detection using a multiway multi-subspace CVA method," *IEEE Access*, vol. 5, pp. 23256–23265, Nov. 2017.

[16] Z. Chen, C. Yang, T. Peng, H. Dan, C. Li, and W. Gui, "A cumulative canonical correlation analysis-based sensor precision degradation detection method," *IEEE Trans. Ind. Electron.*, vol. 66, no. 8, pp. 6321–6330, Aug. 2019.

[17] D. M. J. Tax and R. P. W. Duin, "Support vector domain description," *Pattern Recognit. Lett.*, vol. 20, nos. 11–13, pp. 1191–1199, Nov. 1999.

[18] X. Deng and Z. Zhang, "Nonlinear chemical process fault diagnosis using ensemble deep support vector data description," *Sensors*, vol. 20, no. 16, pp. 1–19, Aug. 2020.

[19] Z. Ge, F. Gao, and Z. Song, "Batch process monitoring based on support vector data description method," *J. Process Control*, vol. 21, no. 6, pp. 949–959, Jul. 2011.

[20] Z. Ge and Z. Song, "Bagging support vector data description model for batch process monitoring," *J. Process Control*, vol. 23, no. 8, pp. 1090–1096, Sep. 2013.

[21] J. Wang, W. Liu, K. Qiu, T. Yu, and L. Zhao, "Dynamic hypersphere based support vector data description for batch process monitoring," *Chemometric Intell. Lab. Syst.*, vol. 172, pp. 17–32, Jan. 2018.

[22] C. Zhang, K. Peng, and J. Dong, "An incipient fault detection and self-learning identification method based on robust SVDD and RBM-PNN," *J. Process Control*, vol. 85, pp. 173–183, Jan. 2020.

[23] M. Yao, H. Wang, and W. Xu, "Batch process monitoring based on functional data analysis and support vector data description," *J. Process Control*, vol. 24, no. 7, pp. 1085–1097, Jul. 2014.

[24] Z. Ge and Z. Song, "Online monitoring and quality prediction of multiphase batch processes with uneven length problem," *Ind. Eng. Chem. Res.*, vol. 53, no. 2, pp. 800–811, Jan. 2014.

[25] L. Luo, S. Bao, J. Mao, and D. Tang, "Phase partition and phase-based process monitoring methods for multiphase batch processes with uneven durations," *Ind. Eng. Chem. Res.*, vol. 55, no. 7, pp. 2035–2048, Feb. 2016.

[26] J. Wang, K. Qiu, W. Liu, T. Yu, and L. Zhao, "Unsupervised-multiscale-sequential-partitioning and multiple-SVDD-model-based process-monitoring method for multiphase batch processes," *Ind. Eng. Chem. Res.*, vol. 57, no. 51, pp. 17437–17451, Dec. 2018.

[27] K. Peng, Q. Li, K. Zhang, and J. Dong, "Quality-related process monitoring for dynamic non-Gaussian batch process with multi-phase using a new data-driven method," *Neurocomputing*, vol. 214, pp. 317–328, Nov. 2016.

[28] X. Tang and Y. Li, "Between-phase nonlinear correlation analysis-based modeling and online monitoring for multiphase batch process with transitions," *IEEE Access*, vol. 7, pp. 158951–158968, Oct. 2019.

[29] C. Peng, R. Lu, O. Kang, and W. Kai, "Batch process fault detection for multi-stage broad learning system," *Neural Netw.*, vol. 129, pp. 298–312, Sep. 2020.

[30] R. Guo, K. Guo, and J. Dong, "Phase partition and online monitoring for batch process based on multiway BEAM," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 4, pp. 1582–1589, Oct. 2017.

[31] S. Zhang, C. Zhao, S. Wang, and F. Wang, "Pseudo time-slice construction using a variable moving window k nearest neighbor rule for sequential uneven phase division and batch process monitoring," *Ind. Eng. Chem. Res.*, vol. 56, no. 3, pp. 728–740, Jan. 2017.

[32] Z. Lv and X. Yan, "Hierarchical support vector data description for batch process monitoring," *Ind. Eng. Chem. Res.*, vol. 55, no. 34, pp. 9205–9214, Aug. 2016.

[33] Z. Lv, X. Yan, Q. Jiang, and N. Li, "Just-in-time learning–multiple subspace support vector data description used for non-Gaussian dynamic batch process monitoring," *J. Chemometrics*, vol. 33, no. 6, pp. 1–13, Jun. 2019.

[34] Y. Hui and X. Zhao, "Batch process monitoring based on WGNPE–GSVDD related and independent variables," *Chin. J. Chem. Eng.*, vol. 26, no. 12, pp. 2549–2561, Dec. 2018.

[35] B. Wang, Q. Jiang, and X. Yan, "Fault detection and identification using a Kullback-Leibler divergence based multi-block principal component analysis and Bayesian inference," *Korean J. Chem. Eng.*, vol. 31, no. 6, pp. 930–943, Jun. 2014.

[36] J. Huang and X. Yan, "Angle-based multiblock independent component analysis method with a new block dissimilarity statistic for non-Gaussian process monitoring," *Ind. Eng. Chem. Res.*, vol. 55, no. 17, pp. 4997–5005, Apr. 2016.

[37] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances," *Ann. Statist.*, vol. 35, no. 6, pp. 2769–2794, Dec. 2007.

[38] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1650–1654, Dec. 2002.

[39] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proc NIPS*, vol. 14, no. 4, pp. 849–856, Jan. 2001

[40] G. A. Darbellay and I. Vajda, "Estimation of the information by an adaptive partitioning of the observation space," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1315–1321, May 1999.

[41] L. Wiskott and T. J. Sejnowski, "Slow feature analysis: Unsupervised learning of invariances," *Neural Comput.*, vol. 14, no. 4, pp. 715–770, Apr. 2002.

[42] J. P. Huelsenbeck and F. Ronquist, "MRBAYES: Bayesian inference of phylogenetic trees," *Bioinformatics*, vol. 17, no. 8, pp. 754–755, Aug. 2001.

[43] P. Cai and X. Deng, "Incipient fault detection for nonlinear processes based on dynamic multi-block probability related kernel principal component analysis," *ISA Trans.*, to be published, doi: 10.1016/j.isatra.2020.05.029.

[44] G. Birol, C. Ündey, and A. Çinar, "A modular simulation package for fed-batch fermentation: Penicillin production," *Comput. Chem. Eng.*, vol. 26, no. 11, pp. 1553–1565, Nov. 2002.

[45] M. Halkidi and M. Vazirgiannis, "Clustering validity assessment: Finding the optimal partitioning of a data set," in *Proc. IEEE Int. Conf. Data Mining*, San Jose, CA, USA, Nov. 2001, pp. 187–194.

XIAOHUI WANG received the M.S. degree from North China Electric Power University, Baoding, China, in 2004. She is currently pursuing the Ph.D. degree with the China University of Petroleum (East China). She works as an Associate Professor at the College of Application Technology, Qingdao University. Her research interests include batch process modeling and monitoring and machine learning.

YANJIANG WANG received the Ph.D. degree from Beijing Jiaotong University, in 2001. He is currently a Professor with the College of Control Science and Engineering, China University of Petroleum (East China), China. His research interests include intelligent information processing, computer vision, and pattern recognition.

XIAOGANG DENG (Member, IEEE) received the Ph.D. degree in chemical engineering and technology from the China University of Petroleum, in 2008. He is currently an Associate Professor with the College of Control Science and Engineering, China University of Petroleum. He works at the National Engineering Laboratory of Offshore Geophysical and Exploration Equipment for the opening project research. His research interests include process modeling and simulation, industrial data mining, and complex system fault diagnosis.

YUPING CAO received the B.S. degree in engineering and the Ph.D. degree in control theory and control engineering from the China University of Petroleum, in 2004 and 2010, respectively. She is currently an Associate Professor with the College of Control Science and Engineering, China University of Petroleum. Her research interest includes process fault diagnosis and prediction.