Korean J. Chem. Eng., 39(2), 275-283 (2022)
DOI: 10.1007/s11814-021-0901-6

pISSN: 0256-1115
eISSN: 1975-7220

Dynamic plant-wide process monitoring based on distributed slow feature analysis with inter-unit dissimilarity

Ruoyu Huang*,, Zetao Li*,†, and Bin Cao*****

*The Electrical Engineering College, Guizhou University, Guiyang 550025, China
**Guiyang Aluminum Magnesium Design and Research Institute Co., Ltd., Guiyang 550081, China
***Chinalco Intelligent Technology Development Co., Ltd., Hangzhou 311199, China
(*Received 3 April 2021 • Revised 17 June 2021 • Accepted 12 July 2021*)

Abstract−In order to overcome the dynamic and large-scale characteristics of the plant-wide processes, this paper proposed a distributed slow feature analysis (SFA) with inter-unit dissimilarity method for process monitoring task. Firstly, to highlight the local dynamic features, the whole process is decomposed into several units according to the prior knowledge. Based on this, SFA monitoring model is built parallelly to handle the dynamic features. Considering the possible information loss caused by the process decomposition, the inter-unit dissimilarity index is carried out to monitor the variations between adjacent units. Finally, the fusion center is conducted by Bayesian inference to combine the results of SFA monitoring models and inter-unit dissimilarity statistics. The effectiveness of the proposed method is tested on the Tennessee Eastman process and an aluminum electrolysis process.

Keywords: Distributed Monitoring, Slow Feature Analysis, Dynamic Process, Fault Detection, Inter-unit Dissimilarity

INTRODUCTION

The general public is paying more attention to the quality, productivity, and safety in large-scale industrial processes. It is necessary to monitor the real-time operating status and to discover the abnormalities. Owing to developments in computer science and sensor technologies, large process data have been recorded and collected. As a result, multivariate statistical process monitoring (MSPM) methods have been widely discussed [1-6].

Despite the successful application of the traditional MSPM methods, process dynamics are not considered by the static monitoring methods. In practice, industrial processes are always under a closed-loop control, resulting in dynamic variations in the process variables. To explore the dynamic characteristics, dynamic principal component analysis (DPCA) [7] methods have been proposed, in which time correlations are introduced to extend data matrices. However, the extracted latent variables in DPCA fail to describe the essential dynamics. To improve this situation, a concurrent monitoring method based on slow feature analysis (SFA) was proposed by Shang et al. to extract the temporal invariant features [8]. Moreover, other modified SFA algorithms, such as recursive SFA [9] and integrated SFA [10], have been proposed. However, centralized methods may be unsuitable for large-scale process monitoring. The issue is that industrial processes consist of several operation units, resulting in more or less correlations. The centralized methods lack representing of local correlations. To this end, distributed monitoring strategies have been developed by first dividing the large-scale process into several units, and then monitoring each unit parallelly.

Generally, there are two categories for the distributed methods: data-based process decomposition methods [11-13] and knowledge-based process decomposition methods [14,15]. In the data-based process decomposition methods, the block division strategy refers to data characteristics according to some given rules [12,13]. Thus, the process data in the same block show similar characteristics. As an alternative, multiblock SFA is developed to gather the highly correlated variables [16,17]. However, the data-based distributed methods, especially the correlation-based decomposition methods, overemphasize the strong correlation between a few variables, neglecting the physical connections. Once the fault happens, the correlations may change with the fault. For knowledge-based process decomposition methods, the whole process is divided based on the domain knowledge [14,15]. Ge and Chen also pointed out that process knowledge would be helpful to process decomposition [18]. In real industrial processes, each operation unit can accomplish a definite subtask, which is the basis of the process decomposition. Distributed methods have solved the problem that industrial processes are submitted to large-scale characteristic. Through process decomposition, a monitoring model is built for each unit. However, the communications among units are cut off directly, that is, the inter-unit information, especially the information between the physically adjacent units, could be neglected by the distributed methods.

To solve the above problems, a distributed slow feature analysis with inter-unit dissimilarity (DSFA-IUD) algorithm is proposed for the large-scale dynamic process monitoring. According to the process knowledge, the large-scale process is divided into several operation units, in which the SFA monitoring models are parallelly built. Moreover, a dissimilarity index between the adjacent units is developed by calculating the variations of the inter-unit correlations. Finally, the fusion center is carried out by combining the results of

†To whom correspondence should be addressed.
E-mail: ztli@gzu.edu.cn
Copyright by The Korean Institute of Chemical Engineers.

SFA monitoring models and inter-unit dissimilarity statistics under a Bayesian framework. The main contributions of this paper are summarized as follows:

(1) A knowledge-based distributed strategy is proposed to divide the large-scale process into several meaningful operation units, so that each unit can be monitored parallelly. As a fault may occur in only one unit or a part of the process, the distributed strategy is more sensitive by highlighting the local information.

(2) Compared with traditional monitoring methods, the SFA model built in each unit isolates temporal dynamics from steady conditions, which makes it capable to handle process dynamics.

(3) Considering that process decomposition may result in information loss among units, a new inter-unit dissimilarity index is carried out to monitor the variations between adjacent units.

METHODOLOGY

1. Distributed Slow Feature Analysis

The process data matrix can be written as $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m] \in R^{n\times m}$, where m and n represent the numbers of variables and samples, respectively. In large-scale processes, each unit consists of many variables, which can accomplish a definite task. Because of the large-scale characteristic, the whole process should be decomposed into several units. The process decomposition may be based on operation tasks, physical constraints, physical correlations, process topologies, feedback constraints, and so on, which is called prior knowledge. Based on the prior knowledge, suppose that the process can be decomposed into p units. Then the dataset $\mathbf{X}$ can be expressed as $\mathbf{X}=[\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_p]$, in which $X_I(I=1, 2, \cdots, p) \in R^{n\times m_I}$ with m_I variables is the dataset for unit I.

After the process decomposition, several units are obtained. Given the input sample $\mathbf{x}(t) \in R^{m_I\times 1}$ in unit I, the optimization of SFA is to find a linear mapping to the slow feature (**SF**) $s_i(t)=w_i^T x$ $(i=1, 2, \cdots m_I)$, such that [19,20]

$$\min_{\boldsymbol{w}_i(\cdot)} \langle \dot{s}_i^2 \rangle_t \tag{1}$$

with the constraints

$$\langle s_i \rangle_t = 0 \quad \text{(zero mean)} \tag{2}$$

$$\langle s_i^2 \rangle_t = 1 \quad \text{(unit variance)} \tag{3}$$

$$\forall i \neq j: \quad \langle s_i s_j \rangle_t = 0 \quad \text{(decorrelation and order)} \tag{4}$$

where $\dot{s}$ indicates the first-order derivative of s_i and $\mathbf{w}_i$ is the weight vector.

As the data matrix is scaled to be zero mean, the extracted SF will be zero mean because of the linear mapping. Then, constraint (2) is satisfied. If the covariance matrices of $\dot{\mathbf{x}}(t)$ and $\mathbf{x}(t)$ are respectively expressed as $\mathbf{R}=\langle \dot{\mathbf{x}}(t)\dot{\mathbf{x}}(t)^T \rangle_t$ and $\mathbf{C}=\langle \mathbf{x}(t)\mathbf{x}(t)^T \rangle_t$, the optimization objective (1) and constraint (3) are simultaneously transformed as

$$\begin{aligned} &\min_{\boldsymbol{w}_i(\cdot)} \langle \dot{s}_i^2 \rangle_t = \min_{\boldsymbol{w}_i} \mathbf{w}_i^T \mathbf{R} \mathbf{w}_i \\ &\text{s.t.} \quad \langle s_i^2 \rangle_t = \mathbf{w}_i^T \mathbf{C} \mathbf{w}_i = 1 \end{aligned} \tag{5}$$

The above optimization is solved by the following generalized eigenvalue decomposition (GED):

$$\mathbf{R}\mathbf{w}_i = \lambda_i \mathbf{C}\mathbf{w}_i \tag{6}$$

where the eigenvalue λ_i satisfies $\lambda_i=\langle \dot{s}_i^2 \rangle_t$. Then constraint (4) can be satisfied automatically through GED. To reach the minimum of the first-order derivative of the SF, the eigenvalues should be in an ascending order, in which the smallest eigenvalue λ_1 corresponds to slowest features $\langle \dot{s}_1^2 \rangle_t$, and so on. Statistics, S^2 and squared prediction error (SPE), are defined to monitor the SF space and residual space:

$$S^2 = \mathbf{s}_k^T \mathbf{s}_k \tag{7}$$

$$\text{SPE} = \mathbf{e}^T \mathbf{e} = (\mathbf{x}-\hat{\mathbf{x}})^T (\mathbf{x}-\hat{\mathbf{x}}) \tag{8}$$

where $\mathbf{s}_k$ is the chosen important SFs and $\hat{\mathbf{x}}$ is the reconstructed $\mathbf{x}$. The SFs are chosen based on the rule that the large L_2 norm of $\mathbf{w}_i$ represents important process variation. The thresholds can be estimated by kernel density estimation (KDE). KDE [21] of the data $\mathbf{y}$ is given by

$$\hat{f}(y) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{\mathbf{y}-y_i}{h}\right) \tag{9}$$

where n is the number of observations and h is the bandwidth and y_i is *i*th sample in the data set. In this study, K represents the Gaussian kernel function

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}u^2\right)} \tag{10}$$

2. The Inter-unit Dissimilarity Statistic

Considering that inter-unit information could be neglected through the process decomposition, a dissimilarity statistic is developed to identify the inter-unit variations. Given two physically adjacent units, $\mathbf{X}_I \in R^{n\times m_I}$ and $\mathbf{X}_J \in R^{n\times m_J}$, the correlation coefficient between variable $x_i(i=1, 2, \cdots, m_I)$ in $\mathbf{X}_I$ and variable $x_j(j=1, 2, \cdots, m_J)$ in $\mathbf{X}_J$ is calculated by

$$r_{ij}^{IJ} = \frac{\frac{1}{n}\sum_{l=1}^{n}[x_i-\mu_i][x_j-\mu_j]}{\sqrt{v_i v_j}} \tag{11}$$

where μ_i and μ_j are the mean value of $\mathbf{x}_i$ and $\mathbf{x}_j$, v_i and v_j are the variance of $\mathbf{x}_i$ and $\mathbf{x}_j$. Then the correlation matrix can be written as

$$D^{IJ} = \begin{bmatrix} r_{11}^{IJ}, r_{12}^{IJ}, \cdots, r_{1m_J}^{IJ} \\ r_{12}^{IJ}, r_{22}^{IJ}, \cdots, r_{2m_J}^{IJ} \\ \vdots \\ r_{m_I1}^{IJ}, r_{m_I2}^{IJ}, \cdots, r_{m_Im_J}^{IJ} \end{bmatrix} \tag{12}$$

Note that the correlation coefficient is calculated based on two vectors. To explore the online variation of the correlation coefficient, a range of samples rather than a monitored sample are taken into consideration. Therefore, inter-unit dissimilarity statistic is to measure the correlation variation over a period, which shows more robustness compared with the statistics monitoring the current sample. For the online datasets $\mathbf{X}_I' \in R^{l\times m_I}$ and $\mathbf{X}_J' \in R^{l\times m_J}$ with l samples, the correlation coefficient between monitored vectors $\mathbf{x}_i'$ and

$\mathbf{x}_j'$ is recorded as $r_{ij}'^{IJ}$. The online variation of the correlation coefficient is calculated as

$$R_{ij}^{IJ}=\left|r_{ij}'^{IJ}-r_{ij}^{IJ}\right| \tag{13}$$

where $|\cdot|$ indicates the absolute value. The inter-unit dissimilarity statistic IUD_{IJ} between units I and J can be calculated by the average variation the correlation coefficient, which is expressed as

$$IUD_{IJ}=\frac{1}{m_I\times m_J}\sum_{i=1}^{m_I}\sum_{j=1}^{m_J}R_{ij}^{IJ} \tag{14}$$

The IUD_{IJ} statistic shows the overall deviation level for the correlation between units I and J. The IUD values under normal condition can be obtained based on some validation datasets. Then, the threshold can also be estimated by kernel density estimation (KDE).

3. Fusion Decision for Distributed SFA and Inter-unit Dissimilarity

3-1. Fusion Decision for S^2 and SPE Statistics

In each unit, the monitoring results can be obtained independently. To integrate all the results to give a final decision, Bayesian inference [15] is used in a probabilistic framework. Given an online sample $\mathbf{x}$, it is divided into p units according to the prior knowledge, which can be written as $\mathbf{x}=[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p]$. For each unit $\mathbf{x}_I$(I=1, 2, ⋯, p), the fault probabilities of S^2 and SPE statistics are given by

$$P_{S^2}(F|\mathbf{x}_I)=\frac{P_{S^2}(\mathbf{x}_I|F)P_{S^2}(F)}{P_{S^2}(\mathbf{x}_I)} \tag{15}$$

$$P_{S^2}(\mathbf{x}_I)=P_{S^2}(\mathbf{x}_I|N)P_{S^2}(N)+P_{S^2}(\mathbf{x}_I|F)P_{S^2}(F) \tag{16}$$

$$P_{SPE}(F|\mathbf{x}_I)=\frac{P_{SPE}(\mathbf{x}_I|F)P_{SPE}(F)}{P_{SPE}(\mathbf{x}_I)} \tag{17}$$

$$P_{SPE}(\mathbf{x}_I)=P_{SPE}(\mathbf{x}_I|N)P_{SPE}(N)+P_{SPE}(\mathbf{x}_I|F)P_{SPE}(F) \tag{18}$$

in which N and F represent the normal and faulty conditions, $P_{S^2}(N)$ and $P_{SPE}(N)$ are the prior probabilities of S^2 and SPE statistics under normal condition, respectively, and and are the prior probabilities of and SPE statistics under faulty condition, respectively. Given a confidence level α, the prior probabilities of normal and faulty conditions are equal to α and $1-\alpha$, respectively. $P_{S^2}(\mathbf{x}_I|N)$, $P_{S^2}(\mathbf{x}_I|F)$, $P_{SPE}(\mathbf{x}_I|N)$ and $P_{SPE}(\mathbf{x}_I|F)$ are calculated as

$$P_{S^2}(\mathbf{x}_I|N)=\exp(-S_I^2/S_{I,\,lim}^2) \tag{19}$$

$$P_{S^2}(\mathbf{x}_I|F)=\exp(-S_{I,\,lim}^2/S_I^2) \tag{20}$$

$$P_{SPE}(\mathbf{x}_I|N)=\exp(-SPE_I/SPE_{I,\,lim}) \tag{21}$$

$$P_{SPE}(\mathbf{x}_I|F)=\exp(-SPE_{I,\,lim}/SPE_I) \tag{22}$$

where S_I^2 and $S_{I,\,lim}^2$ are the S^2 statistic and its control limit for unit I. Similarly, SPE_I and $SPE_{I,\,lim}$ are the SPE statistic and its control limit for unit I.

When the fault probabilities for all the units are calculated, a fusion decision is carried out to integrate the parallel monitoring results into a final decision. The BI-S^2 and BI-SPE statistics are derived by, respectively, combining the S^2 and SPE statistics for all units in a weighted form as follows:

$$\text{BI-}S^2=\sum_{I=1}^{p}\frac{P_{S^2}(F|\mathbf{x}_I)^2}{\sum_{I=1}^{p}P_{S^2}(F|\mathbf{x}_I)} \tag{23}$$

$$\text{BI-SPE}=\sum_{I=1}^{p}\frac{P_{SPE}(F|\mathbf{x}_I)^2}{\sum_{I=1}^{p}P_{SPE}(F|\mathbf{x}_I)} \tag{24}$$

In general, the thresholds of BI-S^2 and BI-SPE are equal to $1-\alpha$.

3-2. Fusion Decision for Inter-unit Dissimilarity

For each pair of adjacent units, the monitoring results of the dissimilarity can be obtained. Bayesian inference is also used to integrate all the dissimilarity results to give a final decision. Suppose that there are a total of q pairs of adjacent units. IUD_I(I=1, 2, ⋯, q) and $IUD_{I,\,lim}$ indicate the Ith IUD statistic and its threshold, respectively. Similarly, the fault probability is given by

$$P(F|IUD_I)=\frac{P(IUD_I|F)P(F)}{P(IUD_I)} \tag{25}$$

$$P(IUD_I)=P(IUD_I|N)P(N)+P(IUD_I|F)P(F) \tag{26}$$

$$P(IUD_I|N)=\exp(-IUD_I/IUD_{I,\,lim}) \tag{27}$$

$$P(IUD_I|F)=\exp(-IUD_{I,\,lim}/IUD_I) \tag{28}$$

in which N and F represent the normal and faulty conditions, P(N) and P(F) are equal to α and $1-\alpha$, respectively. To give an overall decision about the inter-unit dissimilarity, BI-IUD statistic is calculated by

$$\text{BI-IUD}=\sum_{I=1}^{q}\frac{P(F|IUD_I)^2}{\sum_{I=1}^{q}P(F|IUD_I)} \tag{29}$$

3-3. Fusion Decision for the DSFA-IUD Monitoring Model

The statistics are built to monitor operation units and inter-unit dissimilarity, respectively. Then, the fusion decision for the DSFA-IUD monitoring model can be derived by the BIC statistic, which is given by

$$BIC=\frac{\sum_{I=1}^{p}P_{S^2}(F|\mathbf{x}_I)^2+\sum_{I=1}^{p}P_{SPE}(F|\mathbf{x}_I)^2+\sum_{I=1}^{q}P(F|IUD_I)^2}{\sum_{I=1}^{p}P_{S^2}(F|\mathbf{x}_I)^2+\sum_{I=1}^{p}P_{SPE}(F|\mathbf{x}_I)+\sum_{I=1}^{q}P(F|IUD_I)} \tag{30}$$

4. DSFA-IUD Fault Detection Scheme

The diagram of DSFA-IUD monitoring method is shown in Fig. 1. The detailed steps of the DSFA-IUD are as follows.

DSFA-IUD model training:

Step 1: The scaled benchmark data $\mathbf{X}$ is recorded;

Step 2: The benchmark dataset is decomposed into several units on the basis of the process knowledge;

Step 3: For each unit, SFA is performed to generated S^2 and SPE statistics;

Step 4: The thresholds S_{lim}^2 and SPE_{lim} are estimated by KDE;

Step 5: For each pair of adjacent units, the correlation matrix is calculated; and

Step 6: A validation matrix is introduced to estimate the threshold of IUD.

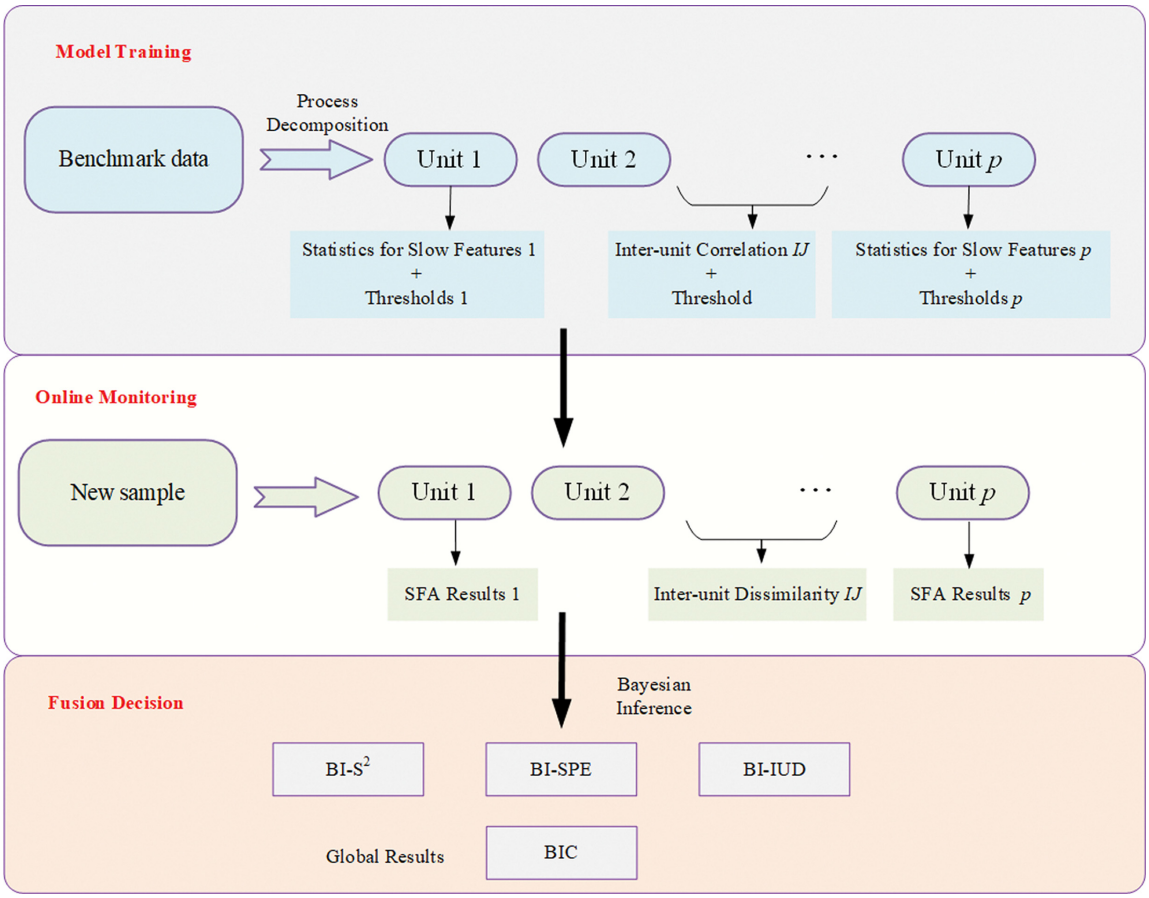

Fig. 1. The diagram of DSFA-IUD process monitoring method.

DSFA-IUD model testing:

Step 1: The online samples are scaled;

Step 2: The samples are decomposed the same with training stage;

Step 3: The S^2 and SPE statistics for each unit are calculated;

Step 4: The variation of the correlation coefficient is calculated and IUD statistic is generated;

Step 5: Bayesian inference is used to integrate the monitoring results of DSFA and IUD strategy;

Step 6: BI-S^2 and BI-SPE statistics are generated to monitor the units and BI-IUD is generated to monitor the inter-unit variation; and

Step 7: The integrated statistic BIC is generated for the proposed DSFA-IUD method.

CASE STUDY

1. Tennessee Eastman (TE) Benchmark Process

The TE process is a widely used benchmark simulation, which was built by Downs and Vogel [22] and restructured by Lyman and Georgakis [23]. It contains five major unit operations: a reactor, a product condenser, a vapor-liquid separator, a recycle compressor and a product stripper. There are 33 measurements in the process. A set of 21 different faults are simulated for monitoring, in which each simulated dataset consists 960 samples and the fault is introduced from 161st sample.

First, the TE process is divided into several units, in which each one implements a specific function or an operation. Since the TE process contains five major units, it can be simply divided into five units. However, there are only two variables in the condenser unit and the compressor unit, respectively; they can be assigned to other related blocks. Therefore, TE process is divided into three different blocks. The detailed variables in each unit of are shown in Table 1 [18]. Then, SFA monitoring models are built for three units. As units 1 and 2, units 2 and 3 are adjacent units, the inter-unit dissimilarity statistics for these two pairs of units are calculated, respectively. Finally, to combine the monitoring results, global monitoring statistics are generated. The fault detection rates (FDRs) of PCA, SFA, Bayesian weigh PCA [18], and DSFA-IUD are listed in Table 2. The cumulative percentage variance criteria for PCA and Bayesian weigh PCA are selected as 85%. The confidence levels for all methods in this paper are set as 0.97. For each fault and the aver-

Table 1. The unit division for TE process

Unit	1	2	3
Variables	1, 2, 3, 5, 6, 7, 8, 9, 21, 23, 24, 25, 32	10, 11, 12, 13, 14, 20, 22, 27, 28, 29, 33	4, 15, 16, 17, 18, 19, 26, 30, 31

Table 2. The monitoring results for TE process

Fault No.	PCA		SFA		Bayesian weigh PCA		DSFA-IUD			
	T^2	SPE	S^2	SPE	T^2	SPE	BI-S^2	BI-SPE	BI-IUD	BIC
1	0.992	0.997	0.996	0.996	0.993	0.995	**0.999**	0.998	0.818	**0.999**
2	0.980	0.986	0.955	0.985	**0.988**	0.985	0.985	0.986	0.533	**0.988**
3	0.002	0.009	0.011	0.173	0.021	0.178	0.081	0.201	**0.280**	**0.280**
4	0.544	0.962	0.998	0.968	0.541	0.999	**1**	0.289	0.524	**1**
5	0.225	0.254	0.999	**1**	0.278	0.339	0.359	0.484	0.624	0.585
6	0.992	**1**	**1**	**1**	0.993	**1**	**1**	0.998	0.999	**1**
7	**1**	**1**	**1**	0.751	**1**	**1**	**1**	0.989	0.664	**1**
8	0.975	0.975	0.668	**0.991**	0.975	0.973	0.984	0.976	0.973	0.980
9	0.006	0.019	0.006	0.140	0.039	0.114	0.069	0.019	**0.198**	0.186
10	0.334	0.341	0.554	0.749	0.443	0.683	0.864	0.891	0.790	**0.938**
11	0.206	0.644	0.524	0.754	0.545	0.651	0.869	0.273	0.645	**0.908**
12	0.971	0.975	0.929	0.996	0.988	0.991	0.994	0.986	0.990	**0.998**
13	0.940	0.955	0.938	0.950	0.943	0.958	0.958	**0.999**	0.946	**0.999**
14	**1**	0.989	0.953	0.988	**1**	0.964	0.976	0.998	0.224	**1**
15	0.012	0.027	0.003	0.181	0.093	0.164	0.153	0.415	0.280	**0.441**
16	0.157	0.245	0.810	0.710	0.281	0.658	0.915	0.323	0.390	**0.951**
17	0.741	0.892	0.961	0.856	0.814	0.965	0.973	0.608	0.909	**0.976**
18	0.887	0.899	0.901	0.905	0.899	0.916	0.911	**0.938**	0.349	**0.938**
19	0.139	0.28	**0.915**	0.088	0.083	0.364	0.749	0.190	0.603	0.859
20	0.315	0.602	0.765	0.789	0.469	0.626	0.694	0.818	0.854	**0.895**
21	0.264	0.430	0.291	0.498	0.334	0.625	0.420	0.615	0.533	**0.764**
Average	0.556	0.642	0.723	0.737	0.606	0.721	0.760	0.666	0.625	**0.842**

age, the best FDR is marked in bold type. In general, the average FDR of DSFA-IUD reaches to 0.842, while those of other state-of-art methods are lower than 0.75. Moreover, DSFA-IUD shows the better performance when handling faults 10, 11, 15, 16, 20 and 21. Taking faults 10 and 21 as typical examples, the monitoring results of the proposed DSFA-IUD are presented in detail.

The comparative monitoring charts for faults 10 and 21 are given in Figs. 2 and 3, respectively. For the random generated fault 10, some of the faulty samples are too small to be detected. It is obvious that PCA and SFA can only detect a part of the faulty samples and the FDRs are lower than 80%. As shown in Figs. 2(c) and (d), the FDR of the proposed method is 0.938, which is a remarkable improvement. Fault 21 is a constant position on the steam 4 valve, which means that the fault will show continuous impact on the process and the fault altitude will get bigger with time going by. Apparently, the monitoring statistics of PCA and SFA rise after 600^{th} sample. Thus, the FDRs of these methods are lower than 50%. For the proposed method, the BI-S^2 and BI-SPE statistics detect the fault before 600^{th} sample. As a supplement, BI-IUD statistic has slight increases between 200^{th} and 300^{th} samples and between 500^{th} and 600^{th} samples. By combining the three statistics, the FDR of the integrate statistic BIC reaches to 0.764. The proposed DSFA-IUD method highlights the local dynamic features by the distributed strategy and avoids information loss simultaneously by generating an IUD index, making it appropriate to monitor process faults.

2. Aluminum Electrolysis Process

Aluminum electrolysis is carried out in aluminum reduction cell. The raw material used for electrolysis is alumina, the electrolyte is molten cryolite, and carbon anode is used. Electrolytic operation is carried out at 950-980 °C. The process data used in this paper are collected from an aluminum corporation in Guizhou, China. The diagram of the aluminum reduction cell can be seen in Fig. 4(a). Alumina raw materials are dissolved into the electrolyte near the feeder and then spread to the whole aluminum reduction cell. Due to the influence of feeding strategy, physical and chemical properties of alumina raw materials and magnetic field, the dissolution and diffusion speeds of alumina in each area are different, resulting in the uneven distribution of alumina concentration in the electrolyte and the different current flowing through each anode. Anode effect (AE) is an abnormal phenomenon that often occurs in the process of direct current electrolytic aluminum. The occurrence of AE has a great impact on normal production, resulting in the decrease of current efficiency, aluminum output and a large amount of fluoride gas, which is harmful for the production and environment [24,25]. AE is generally caused by the lack of local alumina concentration, which is mainly determined by the feeding strategy. In the modern aluminum electrolysis industry, the global AE is mostly judged by the phenomenon of great increase of anode voltage. The global AE is generally caused by the spread of local AE. Because the local AE would not cause a great change of anode voltage, it is difficult to detect the local AE. If the local AE is not intervened, it will gradually turn into global AE. Therefore, to reduce the global AE, it is necessary to detect and eliminate the local AE.

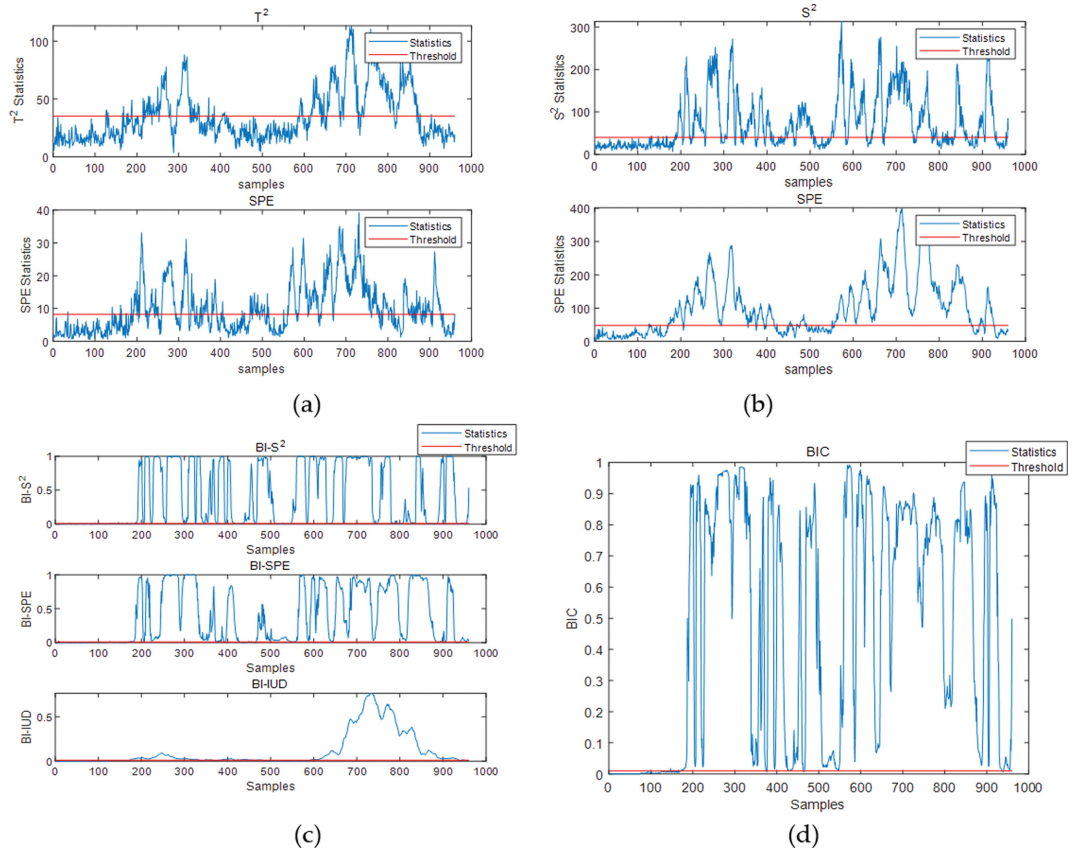

Fig. 2. The monitoring results of (a) PCA, (b) SFA, (c) BI-S2, BI-SPE, and BI-IUD statistics, and (d) BIC statistic of DSFA-IUD method for Fault 10.

Table 3. The unit division results for aluminum electrolysis process

Unit	1	2	3	4	5	6
Anode	A1-A4 B1-B4	A5-A8 B5-B8	A9-A12 B9-B12	A13-A16 B13-B16	A17-A20 B17-B20	A21-A24 B21-B24

There are 6 alumina feeders and 48 anodes in each aluminum reduction cell. The 48 anodes are arranged in two rows in parallel, which are commonly called A1-A24 and B1-B24. The six feeders are distributed between the two rows of anodes, which are marked as F1-F6. The alumina feeders and anodes can be seen in Fig. 4(b). The anode currents are closely related to the concentration of alumina in the electrolyte. AE can be reflected by the changes of the anode currents. For the aluminum electrolysis process, the currents near each feeder may show similar changes when the local or global AE occurs. Therefore, according to the physical locations of the feeders and anodes, the anodes are divided into six units. The unit division results for aluminum electrolysis process are given in Table 3. In this case, a total of 48 anode currents are collected for AE detection and the sample time is 10 s. The training set consists of 1000 samples under normal condition. Meanwhile, 200 samples are collected for testing. According to the daily report, a global AE (detected by the increase of anode voltage) occurs after 160th sample. The monitoring results of PCA, SFA, and DSFA-IUD are shown in Fig. 5. It can be seen that the PCA T^2 statistic can only detect a few faulty samples (between 120th sample and 160th sample). The PCA SPE, as well as SFA S^2 and SPE statistics has a significant increase after the 120th sample, which means that the local AE is detected 400 s before the detected global AE. The best monitoring results are achieved by the proposed DSFA-IUD method. As shown in Figs. 5(c) and (d), the local AE is detected after the 110th sample. In terms of detection time, DSFA-IUD is 100 s earlier than the traditional SFA method, and 500 s earlier than the anode voltage method in the real factory. If early interventions are taken in the 500 s, the global AE can be avoided efficiently. In conclusion, the proposed DSFA-IUD method show better monitoring performance for the large-scale aluminum electrolysis task.

CONCLUSION

A large-scale dynamic process monitoring method based on distributed slow feature analysis and inter-unit dissimilarity is proposed to improve monitoring performance. First, a knowledge-based distributed strategy is used to highlight the local dynamic features.

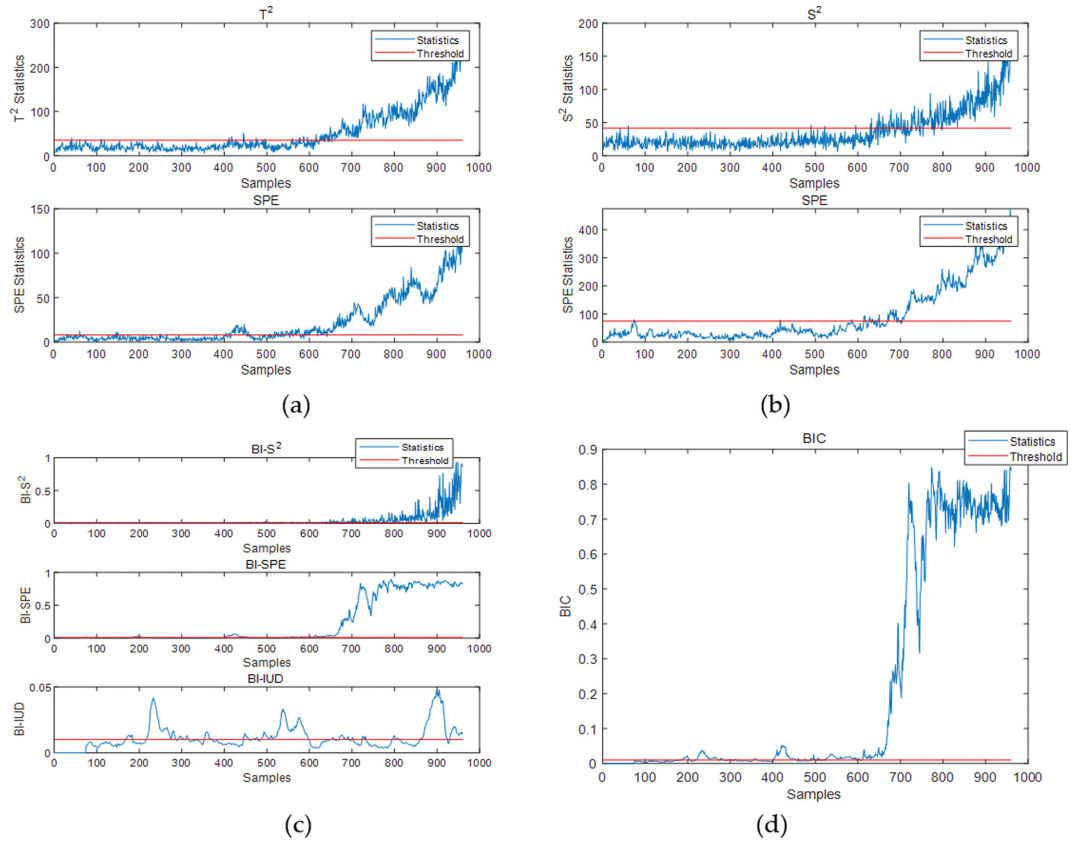

Fig. 3. The monitoring results of (a) PCA, (b) SFA, (c) BI-S2, BI-SPE, and BI-IUD statistics, and (d) BIC statistic of DSFA-IUD method for Fault 21.

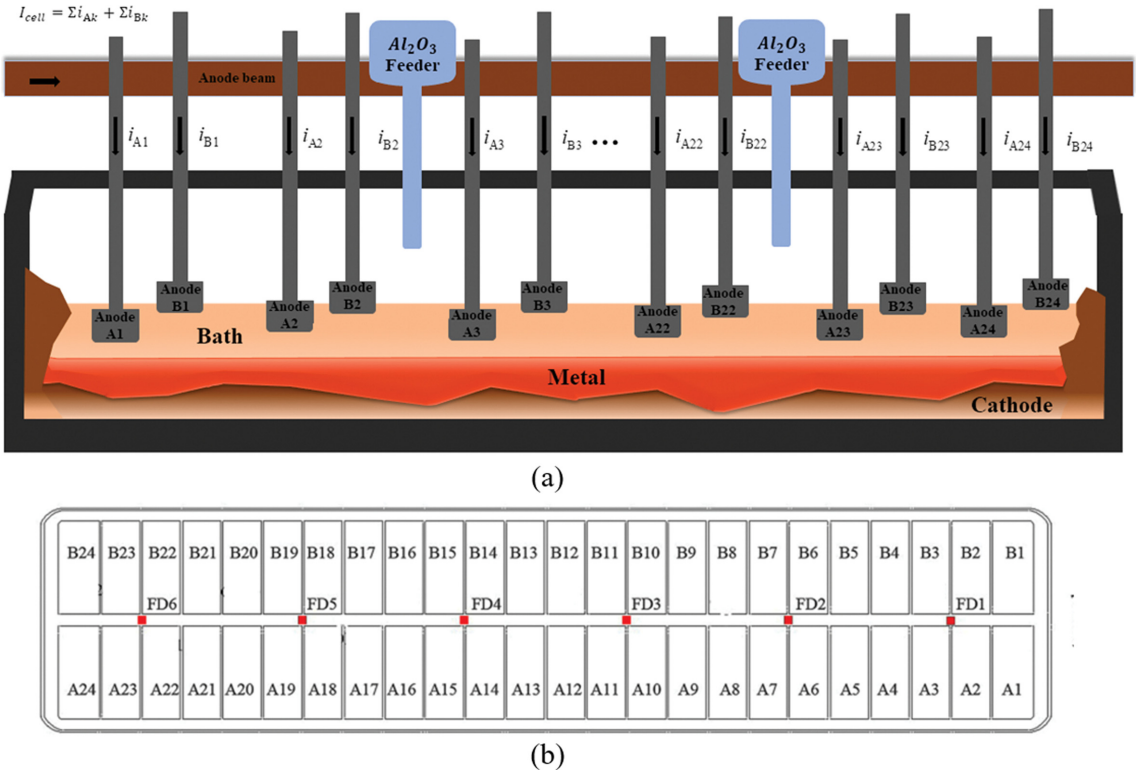

Fig. 4. The diagrams of (a): the aluminum reduction cell, and (b) the locations of the feeders and anodes.

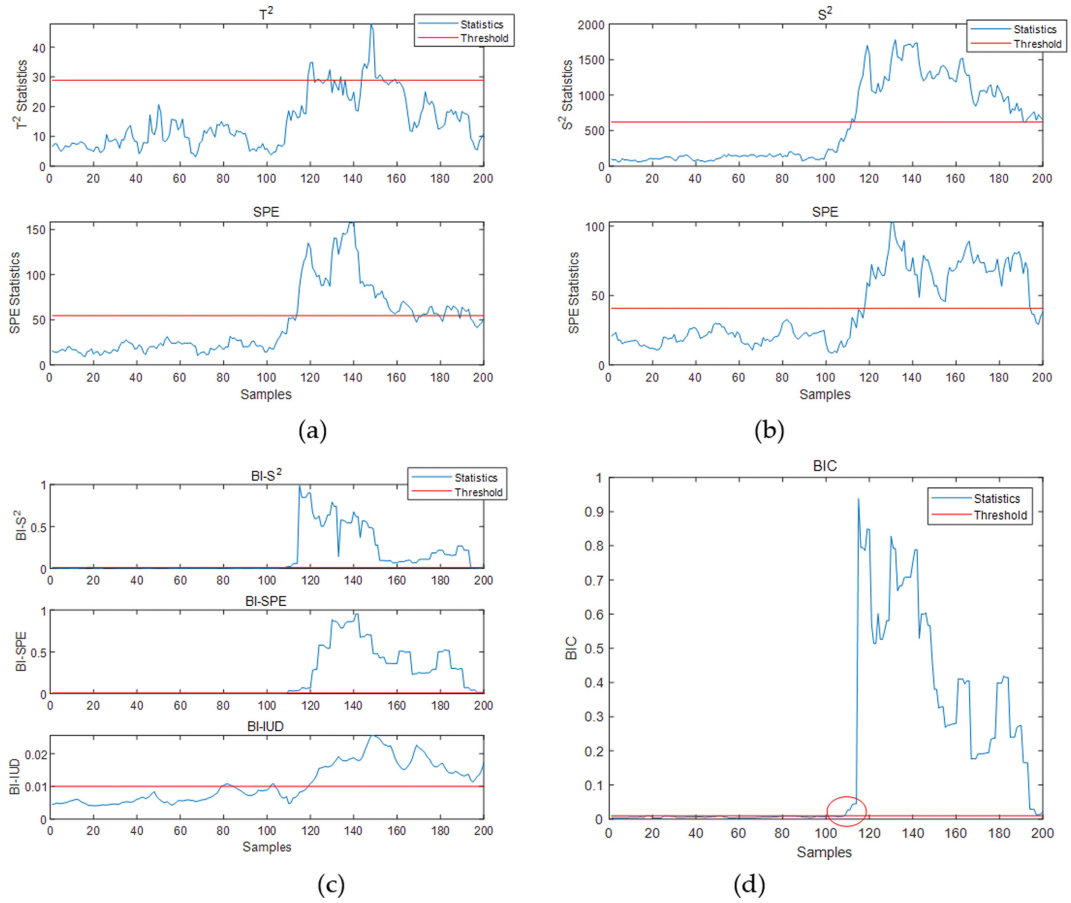

Fig. 5. The monitoring results for aluminum reduction cell: (a) PCA, (b) SFA, (c) BI-S^2, BI-SPE, and BI-IUD statistics, and (d) BIC statistic of DSFA-IUD method.

Then, SFA is parallelly built in each unit to handle process dynamics. Moreover, in case of the information loss by the distributed strategy, a new inter-unit dissimilarity index is carried out to monitor the variations between adjacent units. Finally, the proposed method is applied on the TE process and an aluminum electrolysis process. The monitoring results show the effectiveness of the proposed method. However, the process decomposition in this paper relies on the process knowledge. If biased knowledge is used, the monitoring performance of the distributed method may be affected. As an alternate, knowledge-based process decomposition should refer to process data to enhance the robustness.

ACKNOWLEDGEMENTS

This project is supported by Science and Technology Achievements Application and Industrialization Plan Project of Guizhou Province (Grant No. General 085 [2021]), special fund project of provincial governor for outstanding science and technology education talents in Guizhou Province (Grant number 4 [2010]).

REFERENCES

1. Z. W. Chen, Y. Cao, S. X. Ding, K. Zhang, T. Koenings, T. Peng, C. H. Yang and W. H. Gui, *IEEE Trans. Ind Inform.*, **15**, 2710 (2019).
2. J. Huang, X. Yang and K. X. Peng, *IEEE Trans. Ind. Inform.*, **17**, 6419 (2021).
3. S. Gajjar, M. Kulahci and A. Palazoglu, *Ind. Eng. Chem. Res.*, **59**, 15656 (2020).
4. J. Huang, X. Yang, Y. A. W. Shardt and X. F. Yan, *J. Taiwan Inst. Chem. E.*, **122**, 14 (2021).
5. S. Yin, J. J. Rodriguez-Andina and Y. C. Jiang, *IEEE Trans Ind. Electron.*, **13**, 38 (2019).
6. I. Jaffel, O. Taouali, M. F. Harkat and H. Messaoud, *ISA Trans.*, **64**, 184 (2016).
7. W. Ku, R. H. Store and C. Georgakis, *Chemometr. Intell. Lab.*, **30**, 179 (1995).
8. C. Shang, F. Yang, X. Q. Gao, X. L. Huang, J. A. K. Suykens and D. X. Huang, *AIChE J.*, **61**, 3666 (2015).
9. C. Shang, F. Yang, B. Huang and D. X. Huang, *IEEE Trans. Ind. Electron.*, **65**, 8895 (2018).
10. J. Huang, X. Yang and X. F. Yan, *Control. Eng. Pract.*, **102**, 104558 (2020).
11. Q. C. Jiang and X. F. Yan, *ISA Trans.*, **53**, 1516 (2014).
12. Q. C. Jiang, X. F. Yan and B. A. Huang, *IEEE Trans. Ind. Electron.*, **63**, 377 (2016).
13. C. D. Tong, Y. Song and X. F. Yan, *Ind. Eng. Chem. Res.*, **52**, 9897 (2013).
14. K. K. Huang, Y. M. Wu, H. F. Wen, Y. S. Liu, C. H. Yang and W. H.

Gui, *Control. Eng. Pract.*, **98**, 104386 (2020).
15. J. L. Zhu, Z. Q. Ge and Z. H. Song, *IEEE Trans. Ind. Inform.*, **13**, 1877 (2017).
16. J. Huang, O. K. Ersoy and X. F. Yan, *ISA Trans.*, **85**, 119 (2019).
17. H. Zheng, Q. Jiang and X. Yan, *J. Chemometrics*, **33**, 3110 (2019).
18. Z. Q. Ge and J. H. Chen, *IEEE Trans. Ind. Inform.*, **12**, 310 (2016).
19. C. Shang, B. Huang, F. Yan and D. X. Huang, *J. Process Contr.*, **39**, 21 (2016).
20. S. M. Zhang and C. H. Zhao, *IEEE Trans. Ind. Electron.*, **66**, 3773 (2019).
21. J. M. Lee, C. K. Yoo and I. B. Lee, *J. Process Contr.*, **14**, 467 (2004).
22. J. J. Downs and E. F. Vogel, *Comput. Chem. Eng.*, **17**, 245 (1993).
23. P. R. Lyman and C. Georgakis, *Comput. Chem. Eng.*, **19**, 321 (1995).
24. L. Dion, J. Marks, L. I. KISS, S. Poncsak and C. L. Lagace, *J. Clean. Prod.*, **164**, 357 (2017).
25. H. Pan, L. Kong, X. R. Chen, K. B. Zhou, J. Liu and Q. Xu, *Meas. Sci. Technol.*, **30**, 115105 (2019).