Preprints of the
9th International Symposium on Advanced Control of Chemical Processes
The International Federation of Automatic Control
June 7-10, 2015, Whistler, British Columbia, Canada

A Nonlinear Quality-relevant Process Monitoring Method with Kernel Input-output Canonical Variate Analysis

Huang Linzhe, Cao Yuping, Tian Xuemin, Deng Xiaogang

College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China (Correspondent: CAO Yu-ping, E-mail: caoyp@upc.edu.cn)

Abstract: Traditional process monitoring methods based on kernel canonical variate analysis do not extract variances. They cannot judge whether a process fault that is detected affects product quality. A nonlinear quality-relevant process monitoring method based on kernel input-output canonical variate analysis (KIOCVA) is proposed. Firstly, Process variables and quality variables are mapped into higher-dimensional linear feature spaces via unknown nonlinear mappings respectively. The higher-dimensional linear feature spaces are projected to three subspaces, an input-output correlated subspace that captures correlations between process data and quality data, an uncorrelated input subspace and an uncorrelated output subspace. To monitoring the variances of the uncorrelated input subspace and the uncorrelated output subspace, principal component analysis is performed. Correlations and variances in the higher-dimensional linear feature spaces are extracted by means of nonlinear kernel functions. The proposed KIOCVA method can judge the process fault that is detected affects product quality or not. The effectiveness of the proposed method is demonstrated by case studies of Tennessee Eastman process.

Keywords: Kernel input-output canonical variate analysis, process monitoring, quality monitoring, principal component analysis.

1. INTRODUTION

For complex modern industrial processes, process monitoring is of great importance (Zhou et al., 2013). Traditional data driven process monitoring methods (Ding, 2014; Ge et al., 2013; Li and Xiao, 2011; Qin, 2012; Yin et al., 2014) include principal component analysis (PCA) (Garcia-Alvarez et al., 2012), independent component analysis (ICA), canonical variate analysis (CVA) (Juricek et al., 2004; Russell et al., 2000; Stubbs et al., 2012) etc. On the basis of above traditional methods, many modified methods have been proposed to solve problems like nonlinearity (Deng and Tian, 2006a, 2006b; Lee et al., 2004; Tian et al., 2009), serial correlation (Ku et al., 1995; Stefatos and Hamza, 2010), outliers (Cai et al., 2014; Deng and Tian, 2008; Wang and Romagnoli, 2005), etc. Process variables are usually sampled on-line, and a large amount of data can be stored and used. While quality variables are usually sampled off-line with a low frequency and a time delay. Most data driven process monitoring methods analyze the data from a single layer process data set, and can only judge whether a process is abnormal or not. They cannot judge whether product quality is abnormal or not (Qin and Zheng, 2013). If the product quality is not abnormal, monitoring alarm is usually viewed as a false alarm, which reduces the reliability of a fault detection system to a certain extent.

Industrial faults can be divided into three types: 1) process measurements are abnormal, and quality measurements are abnormal casually; 2) only process measurements are abnormal, quality measurements are not affected due to the compensation of controllers; 3) quality measurements are abnormal, process measurements are normal. Hence, a process data space can be decomposed into two subspaces, one is relevant to quality data, and the other is not. a quality data space can also be decomposed similarly. How to monitor above subspaces needs to be studied. Projection to latent structures (PLS) based monitoring methods are capable of using quality data to guide decomposition of a process data space (Gunther et al., 2009). Zhou et al. (2010) decomposed the residual subspace of process data and proposed total projection to latent structures (T-PLS). Qin et al. (2013) proposed a bi-layer method concurrent projection to latent structures (CPLS) to provide complete monitoring of quality data and concise decomposition of a process data space.

Traditional CVA as well as kernel CVA (KCVA) based process monitoring methods extract state vectors by maximizing a correlation statistic between past data and future data. They are unable to judge whether a process fault that is detected affects product quality. (Deng et al., 2006b, 2008; Juricek et al., 2004; Russell et al., 2000; Stubbs et al., 2012). CVA as well as KCVA process monitoring measures include the T_s^2 statistic which measures the variations inside a state space, the squared prediction error (SPE) statistic which measures the variations inside a residual space, and the T_r^2 statistic which measures the variations outside the state space. The residual space can still include large variations, which may influence the effect of the SPE statistic. The T_r^2 statistic is over sensitive to the inversion of a covariance matrix when small values are included (Russell et al., 2000).

Motivated by realizing quality oriented process monitoring and avoiding statistic problems in CVA and KCVA, a nonlinear quality and process monitoring method based on kernel input-output canonical variate analysis (KIOCVA) is proposed. The proposed method uses kernel trick and CVA to perform nonlinear correlation analysis, and maps quality data and process data to three subspaces, an input-output correlated subspace (IOCS), an uncorrelated input subspace (UIS), and an uncorrelated output subspace (UOS). Each subspace will be monitored with proper measures.

The remainder of this paper is organized as follows. Traditional process monitoring based on KCVA is first reviewed in section 2. KIOCVA method and process monitoring measures are presented in section 3. In section 4, comparison of dynamic kernel PCA and the proposed method is implemented in Tennessee Eastman process. The concluding remarks are summarized in section 5.

2. KCVA FOR PROCESS MONITORING

Consider a normalized process input vector $\tilde{\boldsymbol{u}} \in \mathbb{R}^m$ and an output vector $\tilde{\boldsymbol{v}} \in \mathbb{R}^n$, at a time instant i, the past vector $\boldsymbol{p}_i = [\tilde{\boldsymbol{v}}_{i-1}^{\mathrm{T}}, \tilde{\boldsymbol{v}}_{i-2}^{\mathrm{T}}, \cdots, \tilde{\boldsymbol{v}}_{i-h}^{\mathrm{T}}, \tilde{\boldsymbol{u}}_{i-1}^{\mathrm{T}}, \tilde{\boldsymbol{u}}_{i-2}^{\mathrm{T}}, \cdots \tilde{\boldsymbol{u}}_{i-h}^{\mathrm{T}}]^{\mathrm{T}}$ and the future vector $\boldsymbol{f}_i = [\tilde{\boldsymbol{v}}_{i}^{\mathrm{T}}, \tilde{\boldsymbol{v}}_{i+1}^{\mathrm{T}}, \cdots \tilde{\boldsymbol{v}}_{i+l-1}^{\mathrm{T}}]^{\mathrm{T}}$ can be formed, where h and l are the number of lags. Collect normal data, and form the past matrix $\boldsymbol{P} \in \mathbb{R}^{N \times h(m+n)}$ and the future matrix $\boldsymbol{F} \in \mathbb{R}^{N \times n(l+1)}$ consisting of N samples. Under nonlinear cases, nonlinear transformations $\phi_p(\boldsymbol{p})$ and $\phi_f(\boldsymbol{f})$ can be used to mapping $\boldsymbol{p}$ and $\boldsymbol{f}$ to two higher-dimensional linear feature spaces. KCVA computes two linear projections $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\boldsymbol{\beta}}$ to maximize the following correlation:

$$\begin{cases} \max\limits_{\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\beta}}} \ \rho = \dfrac{\tilde{\boldsymbol{\alpha}}^{\mathrm{T}} \boldsymbol{C}_{\phi_p(\boldsymbol{p})\phi_f(\boldsymbol{f})} \tilde{\boldsymbol{\beta}}}{\sqrt{\tilde{\boldsymbol{\alpha}}^{\mathrm{T}} \boldsymbol{C}_{\phi_p(\boldsymbol{p})\phi_p(\boldsymbol{p})} \tilde{\boldsymbol{\alpha}}} \sqrt{\tilde{\boldsymbol{\beta}}^{\mathrm{T}} \boldsymbol{C}_{\phi_f(\boldsymbol{f})\phi_f(\boldsymbol{f})} \tilde{\boldsymbol{\beta}}}} \\ s.t. \ \operatorname{var} \boldsymbol{c} = \tilde{\boldsymbol{\alpha}}^{\mathrm{T}} \phi_p(\boldsymbol{p}) = 1, \quad \operatorname{var} \boldsymbol{d} = \tilde{\boldsymbol{\beta}}^{\mathrm{T}} \phi_f(\boldsymbol{f}) = 1 \end{cases} \tag{1}$$

where $\boldsymbol{c}$ and $\boldsymbol{d}$ are canonical variables, $\boldsymbol{C}_{\phi_p(\boldsymbol{p})\phi_f(\boldsymbol{f})}$ is the covariance matrix of $\phi_p(\boldsymbol{p})$ and $\phi_f(\boldsymbol{f})$, $\boldsymbol{C}_{\phi_p(\boldsymbol{p})\phi_p(\boldsymbol{p})}$ is the covariance matrix of $\phi_p(\boldsymbol{p})$, and $\boldsymbol{C}_{\phi_f(\boldsymbol{f})\phi_f(\boldsymbol{f})}$ is the covariance matrix of $\phi_f(\boldsymbol{f})$. There exist mapping vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ such that $\tilde{\boldsymbol{\alpha}} = \phi_p(\boldsymbol{P})^{\mathrm{T}} \boldsymbol{\alpha}$, $\tilde{\boldsymbol{\beta}} = \phi_f(\boldsymbol{F})^{\mathrm{T}} \boldsymbol{\beta}$.

By using kernel functions technique, equation (1) can be rewritten as

$$\begin{cases} \max\limits_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \ \rho = \dfrac{\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{K}_p \boldsymbol{K}_f \boldsymbol{\beta}}{\sqrt{\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{K}_p^{\,2} \boldsymbol{\alpha}} \sqrt{\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{K}_f^{\,2} \boldsymbol{\beta}}} \\ s.t. \ \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{K}_p^{\,2} \boldsymbol{\alpha} = 1, \quad \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{K}_f^{\,2} \boldsymbol{\beta} = 1 \end{cases} \tag{2}$$

where Gram matrices $\left[\boldsymbol{K}_p\right]_{i,j} = \mathrm{k}_p(\boldsymbol{p}_i, \boldsymbol{p}_j) = \left\langle \phi_p(\boldsymbol{p}_i) \cdot \phi_p(\boldsymbol{p}_j) \right\rangle$ and $\left[\boldsymbol{K}_f\right]_{i,j} = \mathrm{k}_f(\boldsymbol{f}_i, \boldsymbol{f}_j) = \left\langle \phi_f(\boldsymbol{f}_i) \cdot \phi_f(\boldsymbol{f}_j) \right\rangle$, $\mathrm{k}_p(\cdot,\cdot)$ and $\mathrm{k}_f(\cdot,\cdot)$ are kernel functions, $\langle \cdot \rangle$ denotes dot product. The optimization problem (2) can be solved via singular value decomposition (SVD) or eigenvalue decomposition (Deng et al., 2006b; Lai and Fyfe, 2000). The canonical vector for $\boldsymbol{p}$ can be computed as

$$\boldsymbol{c} = \tilde{\boldsymbol{A}}^{\mathrm{T}} \phi_p(\boldsymbol{p}) = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{K}_p(\boldsymbol{P}, \boldsymbol{p}) \tag{3}$$

where $\tilde{\boldsymbol{A}} = \left[\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \cdots, \tilde{\boldsymbol{\alpha}}_N\right]$, $\boldsymbol{A} = \left[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_N\right]$, and $\boldsymbol{K}_p(\boldsymbol{P}, \boldsymbol{p}) = \left[\mathrm{k}_p(\boldsymbol{p}_1, \boldsymbol{p}), \mathrm{k}_p(\boldsymbol{p}_1, \boldsymbol{p}), \cdots, \mathrm{k}_p(\boldsymbol{p}_N, \boldsymbol{p})\right]^{\mathrm{T}}$. The T_s^2, T_r^2 and SPE statistics are monitored with the following calculations

$$\begin{cases} T_s^2 = \boldsymbol{c}_s^{\ \mathrm{T}} \boldsymbol{c}_s \\ T_r^2 = \boldsymbol{c}_r^{\ \mathrm{T}} \boldsymbol{c}_r \\ SPE = \mathrm{k}_p(p,p) - 2\left\| \boldsymbol{A}_s^{\ \mathrm{T}} \boldsymbol{K}_p(P,p) \right\|^2 \\ \qquad + \boldsymbol{K}_p(P,p)^{\mathrm{T}} \boldsymbol{A}_s \boldsymbol{A}_s^{\ \mathrm{T}} \boldsymbol{K}_p \boldsymbol{A}_s \boldsymbol{A}_s^{\ \mathrm{T}} \boldsymbol{K}_p(P,p) \end{cases} \tag{4}$$

where $\boldsymbol{c}_s$ contains the first s canonical variables, $\boldsymbol{c}_r$ contains the last $r = N - s$ canonical variables, and $\boldsymbol{A}_s$ contains the first s columns of $\boldsymbol{A}$. The T_r^2 statistic is over sensitive to the inversion of $\boldsymbol{C}_{\phi_p(\boldsymbol{p})\phi_p(\boldsymbol{p})}$ when small values exist in $\boldsymbol{C}_{\phi_p(\boldsymbol{p})\phi_p(\boldsymbol{p})}$. Because CVA and KCVA extract correlations but variances, large variations may still exist in the space monitored by SPE. Consequently, it is inappropriate to use T_r^2 and SPE statistics to monitor corresponding spaces.

3. KIOCVA FOR QUALITY AND PROCESS MONITORING

To monitoring nonlinear quality data and process data simultaneously, a kernel input-output canonical variate analysis is proposed. At a time instant i, the input vector $\boldsymbol{x}_i = [\boldsymbol{u}_i^{\mathrm{T}}, \boldsymbol{u}_{i-1}^{\mathrm{T}}, \cdots \boldsymbol{u}_{i-h}^{\mathrm{T}}]^{\mathrm{T}}$ and the output vector $\boldsymbol{y}_i = \boldsymbol{v}_i$ can be constructed by using a process vector $\boldsymbol{u} \in \mathbb{R}^m$ with h lags and a quality vector $\boldsymbol{v} \in \mathbb{R}^n$. Then, with N samples of normal data, an input data matrix $\boldsymbol{X} \in \mathbb{R}^{N \times (h+1)m}$ and an output data matrix $\boldsymbol{Y} \in \mathbb{R}^{N \times n}$ could be obtained. For the proposed KIOCVA method, vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are first mapped into high-dimensional linear spaces $\phi_x(\boldsymbol{x})$ and $\phi_y(\boldsymbol{y})$ via unknown nonlinear mappings $\phi_x(\cdot)$ and $\phi_y(\cdot)$ respectively. Then, CVA is performed to extract canonical variables by maximizing correlations between $\phi_x(\boldsymbol{x})$ and $\phi_y(\boldsymbol{y})$. The canonical variables of $\phi_x(\boldsymbol{x})$, which are relevant to output data, can be used as features of the input-output correlated subspace (IOCS). The predictive residual of $\phi_x(\boldsymbol{x})$ forms the uncorrelated input subspace (UIS), and the predictive residual of $\phi_y(\boldsymbol{y})$ forms the uncorrelated output subspace (UOS). To monitoring abnormal variations in UIS and UOS, PCA is performed.

3.1 Quality-Relevant Process Monitoring in IOCS

The nonlinear mappings $\phi_x(\cdot)$ and $\phi_y(\cdot)$ are unknown, so the canonical variables of $\phi_x(\boldsymbol{x})$ and $\phi_y(\boldsymbol{y})$ can be extracted by KCVA algorithm. The objective function is:

$$\begin{cases} \max\limits_{\tilde{\alpha},\tilde{\beta}} \ \rho = \dfrac{\tilde{\boldsymbol{\alpha}}^{\mathrm{T}} \boldsymbol{C}_{\phi_x(x)\phi_y(y)} \tilde{\boldsymbol{\beta}}}{\sqrt{\tilde{\boldsymbol{\alpha}}^{\mathrm{T}} \boldsymbol{C}_{\phi_x(x)\phi_x(x)} \tilde{\boldsymbol{\alpha}}} \sqrt{\tilde{\boldsymbol{\beta}}^{\mathrm{T}} \boldsymbol{C}_{\phi_y(y)\phi_y(y)} \tilde{\boldsymbol{\beta}}}} \\ s.t. \ \operatorname{var} \boldsymbol{c} = \tilde{\boldsymbol{\alpha}}^{\mathrm{T}} \phi_x(\boldsymbol{x}) = 1, \quad \operatorname{var} \boldsymbol{d} = \tilde{\boldsymbol{\beta}}^{\mathrm{T}} \phi_y(\boldsymbol{y}) = 1 \end{cases} \tag{5}$$

where $\boldsymbol{c}$ and $\boldsymbol{d}$ are canonical variables, $\boldsymbol{C}_{\phi_x(x)\phi_y(y)}$ is the covariance matrix of $\phi_x(\boldsymbol{x})$ and $\phi_y(\boldsymbol{y})$, $\boldsymbol{C}_{\phi_x(x)\phi_x(x)}$ is the covariance matrix of $\phi_x(\boldsymbol{x})$, and $\boldsymbol{C}_{\phi_y(y)\phi_y(y)}$ is the covariance matrix of $\phi_y(\boldsymbol{y})$. There exist mapping vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that $\tilde{\boldsymbol{\alpha}} = \phi_x(\boldsymbol{X})^{\mathrm{T}} \boldsymbol{\alpha}$ and $\tilde{\boldsymbol{\beta}} = \phi_y(\boldsymbol{Y})^{\mathrm{T}} \boldsymbol{\beta}$.

By using kernel function technique, optimization problem (5) can be rewritten as

$$\begin{cases} \max\limits_{\alpha,\beta} \ \rho = \dfrac{\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{K}_x \boldsymbol{K}_y \boldsymbol{\beta}}{\sqrt{\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{K}_x^{\ 2} \boldsymbol{\alpha}} \sqrt{\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{K}_y^{\ 2} \boldsymbol{\beta}}} \\ s.t. \ \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{K}_x^{\ 2} \boldsymbol{\alpha} = 1, \quad \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{K}_y^{\ 2} \boldsymbol{\beta} = 1 \end{cases} \tag{6}$$

where Gram matrices $[\boldsymbol{K}_x]_{i,j} = \mathrm{k}_x(\boldsymbol{x}_i, \boldsymbol{x}_j) = \langle \phi_x(\boldsymbol{x}_i) \cdot \phi_x(\boldsymbol{x}_j) \rangle$ and $[\boldsymbol{K}_y]_{i,j} = \mathrm{k}_y(\boldsymbol{y}_i, \boldsymbol{y}_j) = \langle \phi_y(\boldsymbol{y}_i) \cdot \phi_y(\boldsymbol{y}_j) \rangle$, $\mathrm{k}_x(\cdot,\cdot)$ and $\mathrm{k}_y(\cdot,\cdot)$ are kernel functions. The optimization problem (6) can be transferred to the generalized eigenvalue problem:

$$\begin{bmatrix} 0 & \boldsymbol{K}_x \boldsymbol{K}_y \\ \boldsymbol{K}_y \boldsymbol{K}_x & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix} = \lambda \begin{bmatrix} \boldsymbol{K}_x \boldsymbol{K}_x & 0 \\ 0 & \boldsymbol{K}_y \boldsymbol{K}_y \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix} \tag{7}$$

The problem (7) is often ill-posed. Therefore, $\boldsymbol{K}_x \boldsymbol{K}_x + \eta \boldsymbol{I}_1$ and $\boldsymbol{K}_y \boldsymbol{K}_y + \eta \boldsymbol{I}_2$ are used to replace $\boldsymbol{K}_x \boldsymbol{K}_x$ and $\boldsymbol{K}_y \boldsymbol{K}_y$ respectively, where η is a regularization constant, $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$ are identity matrices with suitable dimension. Then for a new sample of $\boldsymbol{x}$ and a new sample of $\boldsymbol{y}$, canonical vectors can be computed as

$$\begin{cases} \boldsymbol{c} = \tilde{\boldsymbol{A}}^{\mathrm{T}} \phi_x(\boldsymbol{x}) = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{K}_x(\boldsymbol{X}, \boldsymbol{x}) \\ \boldsymbol{d} = \tilde{\boldsymbol{B}}^{\mathrm{T}} \phi_y(\boldsymbol{y}) = \boldsymbol{B}^{\mathrm{T}} \boldsymbol{K}_y(\boldsymbol{Y}, \boldsymbol{y}) \end{cases} \tag{8}$$

where $\tilde{\boldsymbol{A}} = [\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \cdots, \tilde{\boldsymbol{\alpha}}_N]$, $\boldsymbol{A} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_N]$, $\boldsymbol{K}_x(\boldsymbol{X}, \boldsymbol{x}) = [\mathrm{k}_x(\boldsymbol{x}_1, \boldsymbol{x}), \mathrm{k}_x(\boldsymbol{x}_2, \boldsymbol{x}), \cdots, \mathrm{k}_x(\boldsymbol{x}_N, \boldsymbol{x})]^{\mathrm{T}}$, $\tilde{\boldsymbol{B}} = [\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \cdots, \tilde{\boldsymbol{\beta}}_N]$, $\boldsymbol{B} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_N]$, and $\boldsymbol{K}_y(\boldsymbol{Y}, \boldsymbol{y}) = [\mathrm{k}_y(\boldsymbol{y}_1, \boldsymbol{y}), \mathrm{k}_y(\boldsymbol{y}_2, \boldsymbol{y}), \cdots, \mathrm{k}_y(\boldsymbol{y}_N, \boldsymbol{y})]^{\mathrm{T}}$.

The canonical variables of $\phi_x(\boldsymbol{x})$ and $\phi_y(\boldsymbol{y})$ can be monitored with the following T_s^2 statistic

$$T_s^2 = \boldsymbol{c}_s^{\ \mathrm{T}} \boldsymbol{c}_s \tag{9}$$

where $\boldsymbol{c}_s$ contains the first s canonical variables. If T_s^2 statistic violates control limit, it is shown that a fault may occur in process, and product quality may be affected.

3.2 Quality-Uncorrelated Process Monitoring in UIS

The higher-dimensional linear space $\phi_x(\boldsymbol{x})$ can be reconstructed by using $\boldsymbol{c}_s$. Predictive residual $\phi_{ex}(\boldsymbol{x}) = \phi_x(\boldsymbol{x}) - \hat{\phi}_x(\boldsymbol{x})$ is uncorrelated to quality variables, where $\hat{\phi}_x(\boldsymbol{x})$ is the reconstruction of $\phi_x(\boldsymbol{x})$. Considering that large variations may exist in the uncorrelated higher-dimensional linear space UIS, PCA algorithm can be used to analyze variances. However, it is difficult to know nonlinear projection function $\phi_x(\cdot)$. Hence predictive residual $\phi_{ex}(\boldsymbol{x})$ is unknown. PCA algorithm could not be directly performed in UIS. Predictive residuals $\phi_{ex}(\boldsymbol{x})$ can be seen as a nonlinear transformation of $\boldsymbol{x}$. PCA processing on $\phi_{ex}(\boldsymbol{x})$ is equivalent to resolving the eigenvalue problem $\lambda^{ex} \boldsymbol{v}^{ex} = \boldsymbol{C}^F \boldsymbol{v}^{ex}$, where $\boldsymbol{C}^F$ stands for the covariance matrix of $\phi_{ex}(\boldsymbol{x})$, λ^{ex} is an eigenvalue and $\boldsymbol{v}^{ex} = \phi_{ex}(\boldsymbol{X})^{\mathrm{T}} \boldsymbol{\alpha}^{ex}$ is an eigenvector. It can be obtained that

$$N \lambda^{ex} \boldsymbol{A}^{ex} = \boldsymbol{K}_x^{ex} \boldsymbol{A}^{ex} \tag{10}$$

where $\boldsymbol{A}^{ex} = [\boldsymbol{\alpha}_1^{ex}, \boldsymbol{\alpha}_2^{ex}, \cdots, \boldsymbol{\alpha}_N^{ex}]$, kernel matrix $[\boldsymbol{K}_x^{ex}]_{i,j} = \mathrm{k}_x^{ex}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \langle \phi_{ex}(\boldsymbol{x}_i) \cdot \phi_{ex}(\boldsymbol{x}_j) \rangle$. Combing the kernel matrix $\boldsymbol{K}_x$ in (6), it can be yield that

$$\begin{aligned} &\mathrm{k}_x^{ex}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \phi_x(\boldsymbol{x}_i)^{\mathrm{T}} \phi_x(\boldsymbol{x}_j) - \phi_x(\boldsymbol{x}_i)^{\mathrm{T}} \hat{\phi}_x(\boldsymbol{x}_j) \\ &\quad - \hat{\phi}_x(\boldsymbol{x}_i)^{\mathrm{T}} \phi_x(\boldsymbol{x}_j) + \hat{\phi}_x(\boldsymbol{x}_i)^{\mathrm{T}} \hat{\phi}_x(\boldsymbol{x}_j) \\ &= \boldsymbol{K}_x(\boldsymbol{x}_i, \boldsymbol{x}_j) - 2\boldsymbol{K}_x(\boldsymbol{X}, \boldsymbol{x}_i)^{\mathrm{T}} \boldsymbol{A}_s \boldsymbol{A}_s^{\ \mathrm{T}} \boldsymbol{K}_x(\boldsymbol{X}, \boldsymbol{x}_j) \\ &\quad + \boldsymbol{K}_x(\boldsymbol{X}, \boldsymbol{x}_i)^{\mathrm{T}} \boldsymbol{A}_s \boldsymbol{A}_s^{\ \mathrm{T}} \boldsymbol{K}_x \boldsymbol{A}_s \boldsymbol{A}_s^{\ \mathrm{T}} \boldsymbol{K}_x(\boldsymbol{X}, \boldsymbol{x}_j) \end{aligned} \tag{11}$$

where matrix $\boldsymbol{A}_s$ is the first s columns of matrix $\boldsymbol{A}$.

To ensure $\|\boldsymbol{v}^{ex}\|^2 = 1$, $\boldsymbol{\alpha}^{ex}$ should be scaled as $\|\boldsymbol{\alpha}^{ex}\|^2 = \dfrac{\lambda^{ex}}{N}$.

For a new sample of $\boldsymbol{x}$, the first s^{ex} principal component scores can be calculated as $t_i^{ex} = (\boldsymbol{\alpha}_i^{ex})^{\mathrm{T}} \boldsymbol{K}_x^{ex}(\boldsymbol{X}, \boldsymbol{x})$, $i = 1, 2, \cdots, s^{ex}$.

Variations in UIS can be monitoring by using the statistics

$$T_{ex}^2 = [t_1^{ex}, t_2^{ex}, \cdots, t_{s^{ex}}^{ex}] \boldsymbol{\Lambda}_{ex}^{-1} [t_1^{ex}, t_2^{ex}, \cdots, t_{s^{ex}}^{ex}]^{\mathrm{T}} \tag{12}$$

$$\begin{aligned} SPE_{ex} &= \boldsymbol{K}_x^{ex}(\boldsymbol{X}, \boldsymbol{x})^{\mathrm{T}} \boldsymbol{K}_x^{ex}(\boldsymbol{X}, \boldsymbol{x}) \\ &\quad - [t_1^{ex}, t_2^{ex}, \cdots, t_{s^{ex}}^{ex}][t_1^{ex}, t_2^{ex}, \cdots, t_{s^{ex}}^{ex}]^{\mathrm{T}} \end{aligned} \tag{13}$$

where $\boldsymbol{\Lambda}_{ex} = \mathrm{diag}(\lambda_1^{ex}, \lambda_2^{ex}, \cdots, \lambda_{s^{ex}}^{ex})$. If the T_{ex}^2 statistic or the SPE_{ex} statistic violates control limit, it is shown that a fault occurs in process, but product quality is normal. The fault may be compensated by controllers.

3.3 Process-Uncorrelated Quality Monitoring in UOS

Similarly, variations in UOS can also be monitored by using PCA algorithm. Predictive residual $\phi_{ey}(\boldsymbol{y}) = \phi_y(\boldsymbol{y}) - \hat{\phi}_y(\boldsymbol{y})$ is uncorrelated to process variables, where $\hat{\phi}_y(\boldsymbol{y})$ is the reconstruction of $\phi_y(\boldsymbol{y})$. PCA processing on $\phi_{ey}(\boldsymbol{y})$ is equivalent to resolving the eigenvalue problem $N\lambda^{ey}\boldsymbol{A}^{ey} = \boldsymbol{K}_y^{ey}\boldsymbol{A}^{ey}$, where $\boldsymbol{A}^{ey} = \left[\boldsymbol{\alpha}_1^{ey}, \boldsymbol{\alpha}_2^{ey}, \cdots, \boldsymbol{\alpha}_N^{ey}\right]$, kernel matrix $\left[\boldsymbol{K}_y^{ey}\right]_{i,j} = \mathrm{k}_y^{ey}(\boldsymbol{y}_i, \boldsymbol{y}_j) = \left\langle \phi_{ey}(\boldsymbol{y}_i) \cdot \phi_{ey}(\boldsymbol{y}_j) \right\rangle$. Combing the kernel matrix $\boldsymbol{K}_y$ in(6), it can be yield that

$$\begin{aligned}\mathrm{k}_y^{ey}(\boldsymbol{y}_i, \boldsymbol{y}_j) &= \boldsymbol{K}_y(\boldsymbol{y}_i, \boldsymbol{y}_j) - 2\boldsymbol{K}_y(\boldsymbol{Y}, \boldsymbol{y}_i)^{\mathrm{T}}\boldsymbol{B}_s\boldsymbol{B}_s^{\mathrm{T}}\boldsymbol{K}_y(\boldsymbol{Y}, \boldsymbol{y}_j) \\ &+ \boldsymbol{K}_y(\boldsymbol{Y}, \boldsymbol{y}_i)^{\mathrm{T}}\boldsymbol{B}_s\boldsymbol{B}_s^{\mathrm{T}}\boldsymbol{K}_y\boldsymbol{B}_s\boldsymbol{B}_s^{\mathrm{T}}\boldsymbol{K}_y(\boldsymbol{Y}, \boldsymbol{y}_j)\end{aligned} \tag{14}$$

where the matrix $\boldsymbol{B}_s$ is the first s columns of the matrix $\boldsymbol{B}$. For a new sample of $\boldsymbol{y}$, the first s^{ey} principal component scores can be calculated as $t_j^{ey} = \left(\boldsymbol{\alpha}_j^{ey}\right)^{\mathrm{T}}\boldsymbol{K}_y^{ey}(\boldsymbol{Y}, \boldsymbol{y})$, $j = 1, 2, \cdots, s^{ey}$. Variations in UOS can be monitoring by using the statistics

$$T_{ey}^2 = [t_1^{ey}, t_2^{ey}, \cdots, t_{s^{ey}}^{ey}]\boldsymbol{\Lambda}_{ey}^{-1}[t_1^{ey}, t_2^{ey}, \cdots, t_{s^{ey}}^{ey}]^{\mathrm{T}} \tag{15}$$

$$\begin{aligned}SPE_{ey} &= \boldsymbol{K}_y^{ey}(\boldsymbol{Y}, \boldsymbol{y})^{\mathrm{T}}\boldsymbol{K}_y^{ey}(\boldsymbol{Y}, \boldsymbol{y}) \\ &\quad - [t_1^{ey}, t_2^{ey}, \cdots, t_{s^{ey}}^{ey}][t_1^{ey}, t_2^{ey}, \cdots, t_{s^{ey}}^{ey}]^{\mathrm{T}}\end{aligned} \tag{16}$$

where $\boldsymbol{\Lambda}_{ey} = \mathrm{diag}(\lambda_1^{ey}, \lambda_2^{ey}, \cdots, \lambda_{s^{ey}}^{ey})$. If the T_{ey}^2 statistic or the SPE_{ey} statistic violates control limit, it shows that quality samples is abnormal, but process measurements are normal. Some key process variables may not be measured.

4. CASE STUDIES

The Tennessee Eastman process (TEP) (Lyman and Georgakis, 1995) is used to evaluate the effectiveness of the proposed KIOCVA method. The process consists of five major units: a reactor, condenser, compressor, separator and stripper. And it contains eight components: A~H. A simulation system was downloaded from *http://web.mit.edu/braatzgroup/links.html*. A total of 52 measurements are collected for each data set of length $N = 960$. The TEP simulation contains 21 preprogrammed faults (Fault 1-21) and one runs under normal operation (Fault 0). Simulations started with no faults. All faults are introduced after 160 samples.

Both KIOCVA based monitoring and dynamic kernel PCA (DKPCA) based monitoring are performed. For KIOCVA, process variables XMEAS(1-36) and manipulated variables XMV(1-11) are treated as input variables, output variables are quality measurements XMEAS(37-41). Variables in DKPCA are XMEAS(1-36) and XMV(1-11).

The number of lags h in input can be determined by referring to methods in DPCA (Ku et al., 1995). The number of canonical variables (CV) s in IOCS is determined by λ in (7), because λ equals to correlation coefficient ρ in (5).

And the principal component (PC) number s^{ex} and s^{ey} can be determined by the cumulative variance contribution rate (CVCR) strategy. In addition, the data in higher-dimensional linear space should be centered and scaled (Lee et al., 2004). Considering that data may not obey Gaussian distribution, kernel density estimation (Odiowei and Cao, 2010) is used to calculate all control limits of the five monitoring statistics. In the case studies, the kernel function $K(\boldsymbol{x}_1, \boldsymbol{x}_2) = \exp(-\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2 / c)$ is used. CVs with $\rho > 0.5$ are kept in IOCS. Control limits with 99.73% confidence interval, CVCR>0.99 and $h = 2$ are adopted by both KIOCVA and DKPCA. From KIOCVA modeling, $s = 4$, $s^{ex} = 94$ and $s^{ey} = 12$ are obtained. The PC number of DKPCA is 94. Three faults are discussed as follows.

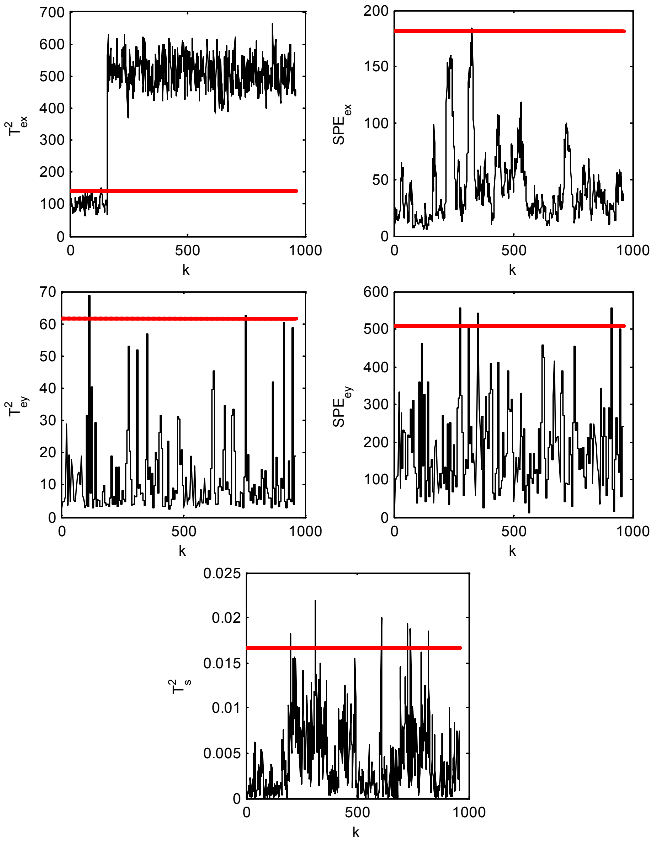

Fig. 1. KIOCVA monitoring results of Fault 4

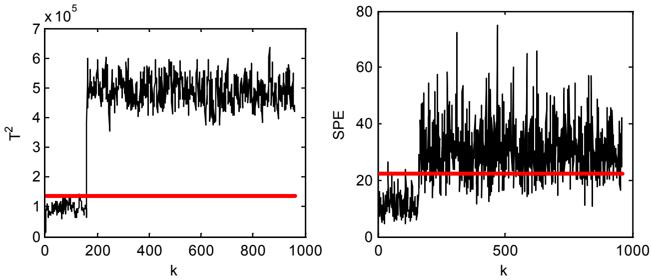

Fig. 2. DKPCA monitoring results of Fault 4

Fault 4 involves a step change in the reactor cooling water inlet temperature. When the fault occurs, there is a sudden increase in the reactor temperature. The fault is compensated by control loops, so other measurements and manipulated variables keep steady. Monitoring results of KIOCVA and DKCVA are showed in Fig. 1 and Fig. 2. For KIOCVA, the fault is alarmed only in UIS, which means that the process fault does not affect output product. However, both of monitored spaces of DKPCA alarm the fault. In fact, product quality is unaffected. So, DKPCA monitoring results may make users doubt the reliability of monitoring systems.

Fault 7 is the result of a step header pressure loss of component C in stream 4. Monitoring results of KIOCVA and DKPCA are shown in Fig. 3 and Fig. 4. KIOCVA detects the fault in all five subspaces. But, statistics T_s^2, T_{ey}^2 and SPE_{ey} go back to normal due to control actions. It is shown that DKPCA alarms all the time after the fault occurs.

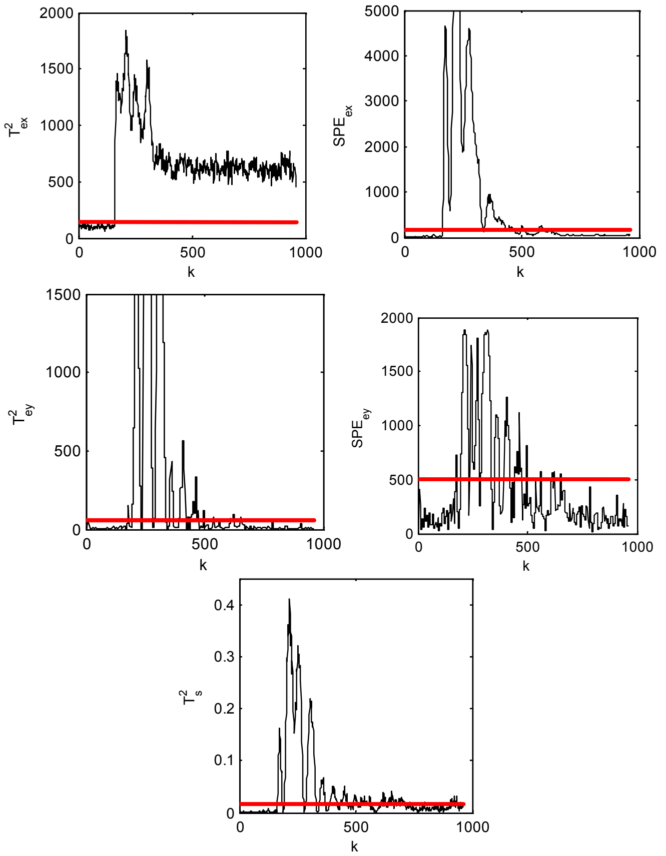

Fig. 3. KIOCVA monitoring results of Fault 7

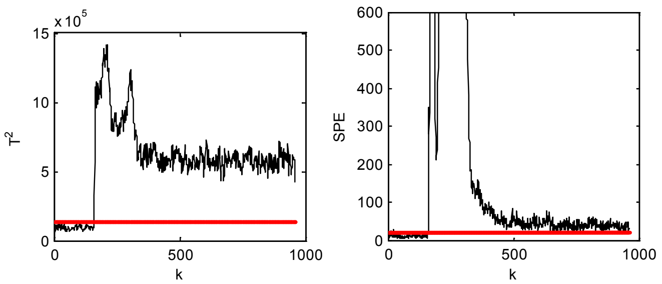

Fig. 4. DKPCA monitoring results of Fault 7

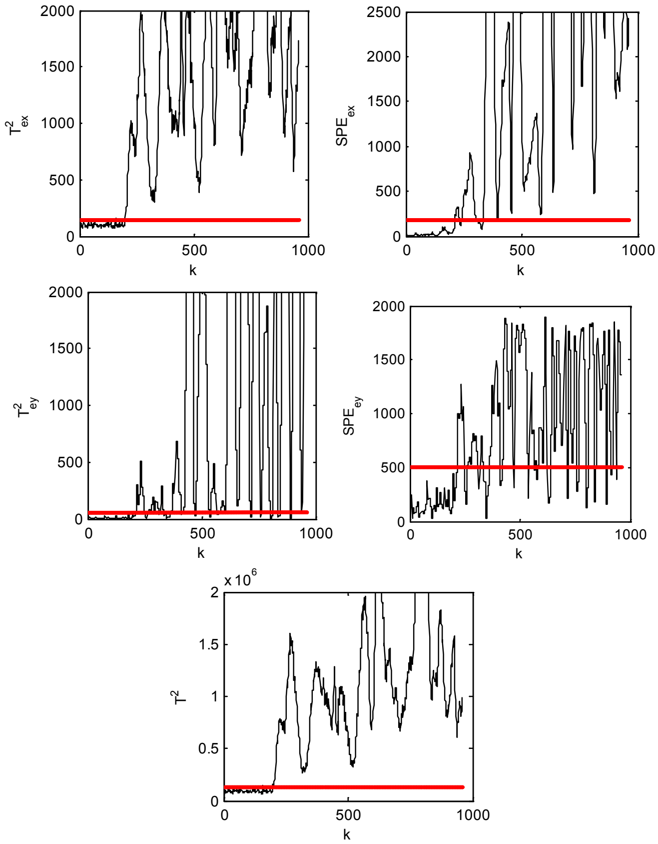

Fig. 5. KIOCVA monitoring results of Fault 13

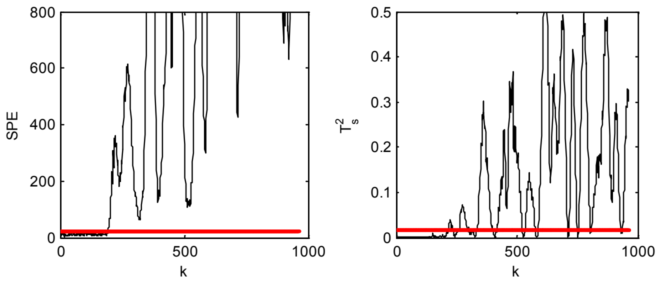

Fig. 6. DKPCA monitoring results of Fault 13

A slow drift in reaction kinetics leads to Fault 13. Fig. 5 and Fig. 6 show that the fault influences all subspaces obviously. Both KIOCVA and DKPCA can detect it effectively and timely.

5. CONCLUSIONS

A new process monitoring approach KIOCVA is proposed to analyze whether a process fault impacts product quality. KIOCVA provides a complete monitoring for nonlinear input data and output data, and captures nonlinear correlation characteristics and nonlinear variance characteristics simultaneously. Simulation results on TEP show that KIOCVA can determine whether a process fault impacts product quality effectively, which is valuable in practical industrial applications.

ACKNOWLEDGEMENTS

This work is supported by National Nature Science Foundation of China (61273160), Doctoral Fund of Shandong Province (BS2012ZZ011), and the Fundamental Research Funds for the Central Universities (14CX02174A).

REFERENCES

Cai, L., Tian, X., and Chen, S. (2014). A process monitoring method based on noisy independent component analysis. *Neurocomputing*, *127*, 231-246.

Deng, X., and Tian, X. (2006a). Multivariate statistical process monitoring using multi-scale kernel principal component analysis. In *Proc. 6th IFAC Symp. Fault Detection, Supervision and Safety of Technical Processes (Beijing, China), Aug* (pp. 108-112).

Deng, X., and Tian, X. (2006b). Nonlinear Process Fault Diagnosis Based on Kernel Canonical Variate Analysis. *Control and Decision*, *21*(10), 1109.

Deng, X., and Tian, X. (2008). Fault diagnosis method based on robust canonical variate analysis. *Control and Decision*, *23*(4), 415.

Ding, S. X. (2014). Data-driven design of monitoring and diagnosis systems for dynamic processes: a review of subspace technique based schemes and some recent results. *Journal of Process Control*, *24*(2), 431-449.

Garcia-Alvarez, D., Fuente, M. J., and Sainz, G. I. (2012). Fault detection and isolation in transient states using principal component analysis. *Journal of Process Control*, *22*(3), 551-563.

Ge, Z., Song, Z., and Gao, F. (2013). Review of recent research on data-based process monitoring. *Industrial & Engineering Chemistry Research*, *52*(10), 3543-3562.

Gunther, J. C., Conner, J. S., and Seborg, D. E. (2009). Process monitoring and quality variable prediction utilizing PLS in industrial fed-batch cell culture. *Journal of Process Control*, *19*(5), 914-921.

Juricek, B. C., Seborg, D. E., and Larimore, W. E. (2004). Fault detection using canonical variate analysis. *Industrial & engineering chemistry research*, *43*(2), 458-474.

Ku, W., Storer, R. H., and Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and intelligent laboratory systems*, *30*(1), 179-196.

Lai, P. L., and Fyfe, C. (2000). Kernel and nonlinear canonical correlation analysis. *International Journal of Neural Systems*, *10*(05), 365-377.

Lee, J. M., Yoo, C., Choi, S. W., Vanrolleghem, P. A., and Lee, I. B. (2004). Nonlinear process monitoring using kernel principal component analysis. *Chemical Engineering Science*, *59*(1), 223-234.

Lee, J. M., Yoo, C., and Lee, I. B. (2004). Fault detection of batch processes using multiway kernel principal component analysis. *Computers & chemical engineering*, *28*(9), 1837-1847.

Li, H., and Xiao, D. (2011). Survey on data driven fault diagnosis methods. *Control and Decision*, *26*(1), 1-9.

Lyman, P. R., and Georgakis, C. (1995). Plant-wide control of the Tennessee Eastman problem. *Computers & chemical engineering*, *19*(3), 321-331.

Odiowei, P. P., and Cao, Y. (2010). State-space independent component analysis for nonlinear dynamic process monitoring. *Chemometrics and intelligent laboratory systems*, *103*(1), 59-65.

Qin, S. J. (2012). Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, *36*(2), 220-234.

Qin, S. J., and Zheng, Y. (2013). Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures. *AIChE Journal*, *59*(2), 496-504.

Russell, E. L., Chiang, L. H., and Braatz, R. D. (2000). Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, *51*(1), 81-93.

Stefatos, G., and Hamza, A. B. (2010). Dynamic independent component analysis approach for fault detection and diagnosis. *Expert Systems with Applications*, *37*(12), 8606-8617.

Stubbs, S., Zhang, J., and Morris, J. (2012). Fault detection in dynamic processes using a simplified monitoring-specific CVA state space modelling approach. *Computers & Chemical Engineering*, *41*, 77-87.

Tian, X., Zhang, X., Deng, X., and Chen, S. (2009). Multiway kernel independent component analysis based on feature samples for batch process monitoring. *Neurocomputing*, *72*(7), 1584-1596.

Wang, D., and Romagnoli, J. A. (2005). Robust multi-scale principal components analysis with applications to process monitoring. *Journal of Process Control*, *15*(8), 869-882.

Yin, S., Ding, S., Xie, X., and Luo, H. (2014). A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, *61*(11), 6418-6428.

Zhou, D., Li, G., and Qin, S. J. (2010). Total projection to latent structures for process monitoring. *AIChE Journal*, *56*(1), 168-178.

Zhou, D., Liu, Y., and He, X. (2013). Review on fault diagnosis techniques for closed-loop systems. *Acta Automatica Sinica*, *39*(11), 1933-1943.