Korean J. Chem. Eng., **31**(11), 1935-1942 (2014)
DOI: 10.1007/s11814-014-0151-y

pISSN: 0256-1115
eISSN: 1975-7220

Improved fault detection in nonlinear chemical processes using WKPCA-SVDD

Qingchao Jiang and Xuefeng Yan†

Key Laboratory of Advanced Control and Optimization for Chemical Processes of Ministry of Education,
East China University of Science and Technology, Shanghai 200237, P. R. China
(*Received 11 July 2013 • accepted 2 June 2014*)

Abstract−Conventional kernel principal component analysis (KPCA) does not always perform well for nonlinear process monitoring because the beneficial information for fault detection may be submerged under the retained kernel principal components (KPCs). To overcome this deficiency, an adaptively weighted KPCA integrated with support vector data description (WKPCA-SVDD) monitoring method is proposed. In WKPCA-SVDD, the importance of each KPC is evaluated online by the change rate of T^2 statistic and then distinguished weighting values are set on the KPCs. The behaviors of all KPCs are comprehensively evaluated by the SVDD technique. Since the beneficial information is highlighted, the monitoring performance of the statistic in the dominant subspace can be improved. The proposed WKPCA-SVDD is applied to both a numerical process and the complicated Tennessee Eastman benchmark process. Monitoring results have indicated the efficiency of the WKPCA-SVDD method.

Keywords: Fault Detection, Nonlinear Process Monitoring, Weighted Kernel Principal Component Analysis, Support Vector Data Description

INTRODUCTION

Multivariate statistical process control (MSPC) methods have progressed significantly in monitoring of large-scale industrial processes, and among them principal component analysis (PCA) is the most widely used [1-10]. PCA as a dimensionality reduction technique can map the high-dimensional and correlated data into two subspaces: the dominant subspace and the residual subspace. The dominant subspace contains most variance of training data, while the residual subspace has the least [11-14]. Correspondingly, two statistics are constructed to interpret the mean and variance information of process in the two subspaces: T^2 statistic and Q statistic. The T^2 statistic measures the variations of the PCs in the dominant subspace; however, some of these PCs' scores may not change fast and significantly when a fault occurs, i.e., there is not much fault information reflected on these PCs. The fault information may be submerged by the PCs with useless information if the fault information is limited, leading to poor monitoring performance [15-17]. Therefore, it is very important to highlight the useful information exhibited from the process for fault detection and diagnosis, and the weighting strategies can be employed.

Concerning the weighting strategy and multivariate statistical process monitoring, numerous researches have been reported [12,18-22]. Jiang and Yan [12] suggested a weighted principal component analysis (WPCA) method which directly examines the direction of each principal component. It takes fault information into consideration online, and it determines the weighting values according to the importance of the PCs objectively. The WPCA method has significantly improved the performance of fault detection and diagnosis of linear process; however, for some complicated cases in industrial chemical and biological processes with particular nonlinear characteristics, linear PCA performs poor [23].

To address the nonlinear monitoring problem, several nonlinear extensions of PCA have been reported [24-26]. Schölkopf et al. [27] proposed the kernel principal component analysis (KPCA), which overcame limitations of the neural-network-based nonlinear PCA and has been the most widely used nonlinear process monitoring technology. The basic idea of KPCA is first to project the input space into a feature space via a nonlinear map, which makes data structure more linear, and then to extract principal components in the feature space. By introducing a kernel function, the nonlinear mapping and the inner product computation can be avoided. Similar to the conventional PCA, two monitoring statistics, T^2 statistic and Q statistic, have been constructed to monitor the dominant subspace and the residual subspace, respectively [4,28-30]. In KPCA monitoring, the first several kernel principal components (KPCs) are employed to construct the dominant subspace, and there still exists the useful information being submerged problem, which may affect the process monitoring performance. To address this problem in KPCA monitoring, Jiang and Yan [21] developed a probability density-based weighted KPCA (WKPCA) for monitoring performance improvements. The WKPCA considers the statistical distributions of the KPCs and significantly improves the monitoring performance of the T^2 statistic; however, the densities in each point need to be calculated.

This work focuses on extending the weighted WPCA into the nonlinear form and improving the KPCA-based nonlinear chemical process monitoring performance. Different from the previously developed methods [12,21], the current work combines the weighting strategy and the support vector data description (SVDD) technology for monitoring performance improvements. In PCA (KPCA) monitoring, T^2 statistic has the ability to measure the variation directly

†To whom correspondence should be addressed.
E-mail: xfyan@ecust.edu.cn

along each principal component and that is well used by WKPCA-SVDD for useful information indexing. First, WKPCA-SVDD uses normal operational data to build a conventional KPCA model. Second, the change rate of the T^2 statistic along each kernel principal component is constructed to capture the most useful information. The useful principal component trends to own larger change rates, and distinct weighting values are then set on the kernel components. Then the behaviors of all components are comprehensively evaluated by the SVDD technology. The remainder of the article is organized as follows. The KPCA algorithm is briefly reviewed in the next section, as well as the SVDD technology. Second, the WKPCA-SVDD is introduced and details are presented. The illustrative examples are given to demonstrate the effect of the proposed methods. Finally, conclusions are drawn in the last section.

PRELIMINARIES

1. KPCA

In KPCA, observations are nonlinearly mapped into a high-dimensional feature space F and then linear PCA is employed to extract the nonlinear correlation between the variables [4,27,30]. Let the normalized training set be $x_1, \ldots, x_N \in R^m$ with N observations consisting of m measured process variables. The covariance matrix in the feature space F is calculated as

$$\mathbf{C}^F = \frac{1}{N}\sum_{j=1}^{N} \Phi(\mathbf{x}_j)\Phi(\mathbf{x}_j)^T \quad (1)$$

where $\Phi(\bullet)$ is the nonlinear mapping function and $\sum_{j=1}^{N}\Phi(\mathbf{x}_j)=0$ is assumed. The kernel principal component can be obtained by the eigenvalue problem [4,27,30]

$$\lambda\mathbf{v} = \mathbf{C}^F\mathbf{v} = \frac{1}{N}\sum_{j=1}^{N}\langle \Phi(\mathbf{x}_j)^T, \mathrm{v}\rangle \Phi(\mathbf{x}_j)\lambda\mathbf{v} \quad (2)$$

where λ and $\mathbf{v}$ denote the eigenvalue and the eigenvector of the covariance matrix $\mathbf{C}^F$, respectively, and $\langle \mathbf{x}, \mathbf{y}\rangle$ denotes the dot product between $\mathbf{x}$ and $\mathbf{y}$. Then, $\lambda\mathbf{v}=\mathbf{C}^F\mathbf{v}$ is equivalent to $\lambda\langle \Phi(\mathbf{x}_k), \mathbf{v}\rangle = \langle \Phi(\mathbf{x}_k), \mathbf{C}^F\mathbf{v}\rangle$(k=1, …, N). Defining an N×N matrix $\mathbf{K}$ by $[\mathbf{K}]_{ij} = K_{ij} = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j)\rangle$. We can obtain that [4,27,30]

$$\frac{1}{N}\sum_{i=1}^{N}\alpha_i\langle \Phi(\mathbf{x}_k), \sum_{j=1}^{N}\Phi(\mathbf{x}_j)\rangle\langle \Phi(\mathbf{x}_j), \Phi(\mathbf{x}_i)\rangle = \frac{1}{N}\sum_{i=1}^{N}\alpha_i\sum_{j=1}^{N}K_{kj}K_{ji} \quad (k=1, \ldots, N) \quad (3)$$

Regarding kernel functions, representative forms include polynomial kernel, sigmoid kernel and radial basis kernel, and it has been suggested that the above functions give similar results if appropriate parameters are chosen [31]. In the current work, the radial basis kernel is used and the parameter selection can refer to the work by Lee et al. [4]. The principal component $\mathbf{t}$ of a test vector $\mathbf{x}$ is then extracted by projecting $\Phi(\mathbf{x})$ onto the eigenvectors $\mathbf{v}_k$ in F, as follows:

$$t_k = \langle \mathbf{v}_k, \Phi(\mathbf{x})\rangle = \sum_{i=1}^{N}\alpha_i^k\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x})\rangle \qquad (k=1, \ldots, p) \quad (4)$$

where p is the number of KPCs retained in the dominant subspace [27,30].

In KPCA monitoring, T^2 statistic is constructed as [4,27,30]

$$T^2 = [t_1, \ldots, t_p]\Lambda^{-1}[t_1, \ldots, t_p]^T \quad (5)$$

where t_k is obtained from Eq. (4) and Λ^{-1} is the diagonal matrix of the inverse of the eigenvalues with the retained PCs [4,27,30]. The monitoring of T^2 is based on the assumption that the training data are multivariate normal in the feature space.

2. SVDD

SVDD is a recently developed classifier method which tries to find a sphere with minimum volume containing all (or most of) the data objects. Given its efficiency, SVDD has been introduced into process monitoring and control fields and several studies have been reported [10,32,33]. With a nonlinear transformation function Φ, the SVDD is to solve the following optimization problem [34,35]:

$$\begin{aligned} &\min_{R,\boldsymbol{a},\xi} R^2 + C\sum_{i=1}^{n}\xi_i \\ &\text{s.t.} \|\Phi(\mathbf{y}_i) - \mathbf{a}\|^2 \le R^2 + \xi_i \end{aligned} \quad (6)$$

where $\mathbf{a}$ is the center of the hypersphere; C gives the trade-off between the volume of the hypersphere and the number of errors; R^2 denotes the squared distance from the center $\mathbf{a}$ to the boundary; and ξ_i represents the slack variable which allows the probability that some of the training samples can be wrongly classified. The dual form of the optimization problem can be obtained as [34,35]

$$\begin{aligned} &\min_{\alpha_i}\sum_{i=1}^{n}\alpha_i K_2(\mathbf{y}_i, \mathbf{y}_j) - \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j K_2(\mathbf{y}_i, \mathbf{y}_j) \\ &\text{s.t.}\ 0 \le \alpha_i \le C, \sum_{i=1}^{n}\alpha_i = 1 \end{aligned} \quad (7)$$

where $K_2(\mathbf{y}_i, \mathbf{y}_j) = \langle \Phi(\mathbf{y}_i), \Phi(\mathbf{y}_j)\rangle$ is a kernel function which is introduced to compute the inner product in the feature space, α_i is a Lagrange multiplier. Solving Eq. (7) provides a set α_i and the samples $\mathbf{y}_i$ with $\alpha_i>0$ are called support vectors (SVs) of the description. The squared distance R^2 from the center $\mathbf{a}$ to the boundary is calculated as [34,35]

$$R^2 = K_2(\mathbf{y}, \mathbf{y}) - 2\sum_{i=1}^{n}\alpha_i K_2(\mathbf{y}_i, \mathbf{y}) + \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j K_2(\mathbf{y}_i, \mathbf{y}_j) \quad (8)$$

for any $\mathbf{y} \in$ SV.

WKPCA-SVDD METHOD

Given a weighting matrix $\mathbf{W}$, the weighted kernel principal component $\mathbf{t}_W$ becomes

$$\mathbf{t}_W = \mathbf{W}\mathbf{t} \quad (9)$$

where the weighting matrix $\mathbf{W}=\mathrm{diag}(w_1, w_2, \ldots, w_p)$ is a diagonal matrix. The following task is to determine the weighting values w_i in the weighting matrix $\mathbf{W}$. The KPCs containing more fault information should be high-weighted and the corresponding weighting values should be set large. To describe the variation of the k^{th} principal component and quantitatively evaluate the importance of the k^{th} principal component, the change rate $RT_k^2(i)$ of T_k^2 on the i^{th} sample point is defined as follows:

$$RT_k^2(i) = \frac{\sum_{r=i-h}^{i-1} T_k^2(r)}{\sum_{j=1}^{N} T_k^2(j)} \quad (10)$$

where $T_k^2(i)$ is the T_k^2 statistic on the i[th] sample point and $T_k^2(i)=t_k(i)\lambda_k^{-1}t_k^T(i)$; h is the window width considered around the current sample and N is the number of points in the normal training samples. In this work, the weighting values are determined as

$$w_k(i)=RT_k^2(i) \quad (11)$$

After weighting, the distribution of the component scores become complicated and the threshold cannot be determined directly from a normal distribution. The SVDD does not need the Gaussian distribution assumption and is employed to examine the behavior of the scores. Based on normal training data, the SVDD mode can be constructed, and the input variables to the SVDD model is

$$\mathbf{Y}=[\mathbf{t}_{W1}, \mathbf{t}_{W2}, \ldots, \mathbf{t}_{Wp}]^T \quad (12)$$

Assuming that the online obtained weighted principal components score is denoted as

$$\mathbf{z}=[\mathbf{t}_{W1}, \mathbf{t}_{W2}, \ldots, \mathbf{t}_{Wp}]^T \quad (13)$$

The squared distance from the center of the sphere is calculated as

$$D^2=\|\mathbf{z}-\mathbf{a}\|^2=K_2(\mathbf{z},\mathbf{z})-2\sum_{i=1}^{n}\alpha_i K_2(\mathbf{z},\mathbf{y}_i)+\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j K_2(\mathbf{y}_i,\mathbf{y}_j) \quad (14)$$

For fault detection purpose, the index is defined as follows:

$$DR=\frac{\|\mathbf{z}-\mathbf{a}\|^2}{R^2} \quad (15)$$

and the control limit $DR_l=1$. That is, the sample **z** is accepted when the distance is smaller or equal than R^2, and otherwise, it is rejected and regarded as a fault point.

IMPLEMENTATION

Two case studies, on the numerical nonlinear process and the TE process, were performed to illustrate the efficiency of the proposed WKPCA algorithm.

1. Case Study on the Numerical Nonlinear Process

Consider the process which is with three variables but only one

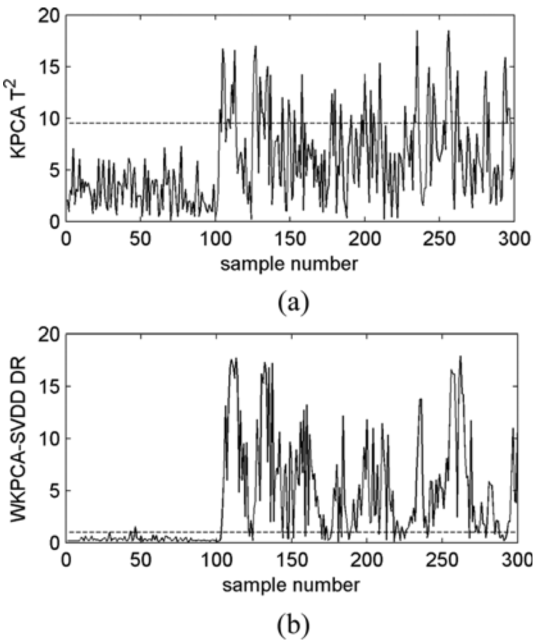

Fig. 1. Monitoring results of fault 1 (a) KPCA (b) WKPCA-SVDD.

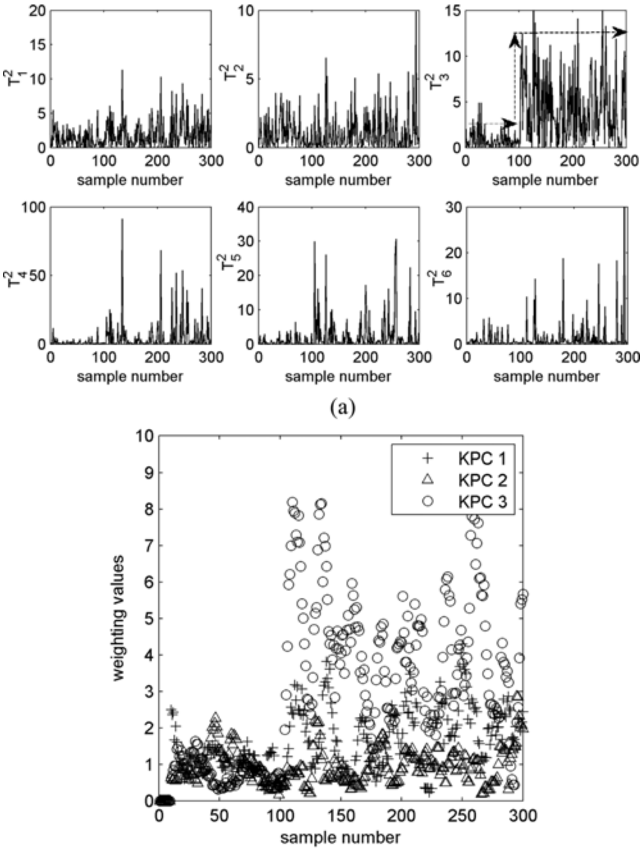

Fig. 2. (a) T^2 of each principal component; (b) weighting values when monitoring fault 1.

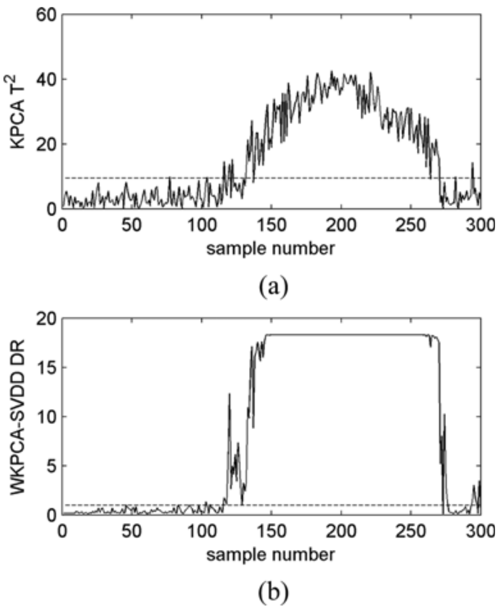

Fig. 3. KPCA and WKPCA-SVDD monitoring charts for (a) fault 1 (b) fault 2.

factor [25]:

$$x_1=t+e1 \quad (16)$$

$$x_2=t^2-3t+e2 \quad (17)$$

$$x_3=-t^3+3t^2+e3 \quad (18)$$

where e1, e2 and e3 are independent noise variables N(1, 0.01), and $t\in[0.01, 2]$. Normal data comprising 100 samples are generated according to these equations. Two sets of test data comprising 300 samples each are also generated. The following two faults are applied separately during generation of the test data sets.

Fault 1: a step change of x_2 by -0.3 is introduced starting from sample 101.

Fault 2: x_1 is linearly increased from sample 101 to 270 by adding $0.01\times(i-100)$ to the x_1 value of each sample in this range, where i is the sample number.

When building the KPCA model, three kernel principal components are retained in the dominant subspace according to the cumulative variance contribution rule. The monitoring performances of the conventional KPCA and the WKPCA-SVDD are presented in Fig. 1. Fig. 1(a) shows the T^2 monitoring results of the KPCA and Fig. 1(b) shows the monitoring results of the WKPCA-SVDD. As is shown in Fig. 1(a), T^2 can detect the fault 1; however, the statistic did not stay above the confidence limit, which resulted in high non-detection rate. The monitoring result of the T^2 statistic is poor. From Fig. 1(b), we can see that the fault 1 can be successfully detected by the WKPCA-SVDD method and the non-detection rate is much lower. To explore the cause of the poor monitoring performance of T^2 in KPCA, the monitoring performances of the first six KPCs are examined and illustrated in Fig. 2(a), and the weighting values when monitoring fault 1 are presented in Fig. 2(b).

Fig. 2(a) reveals that when fault 1 occurs, the six principal com-

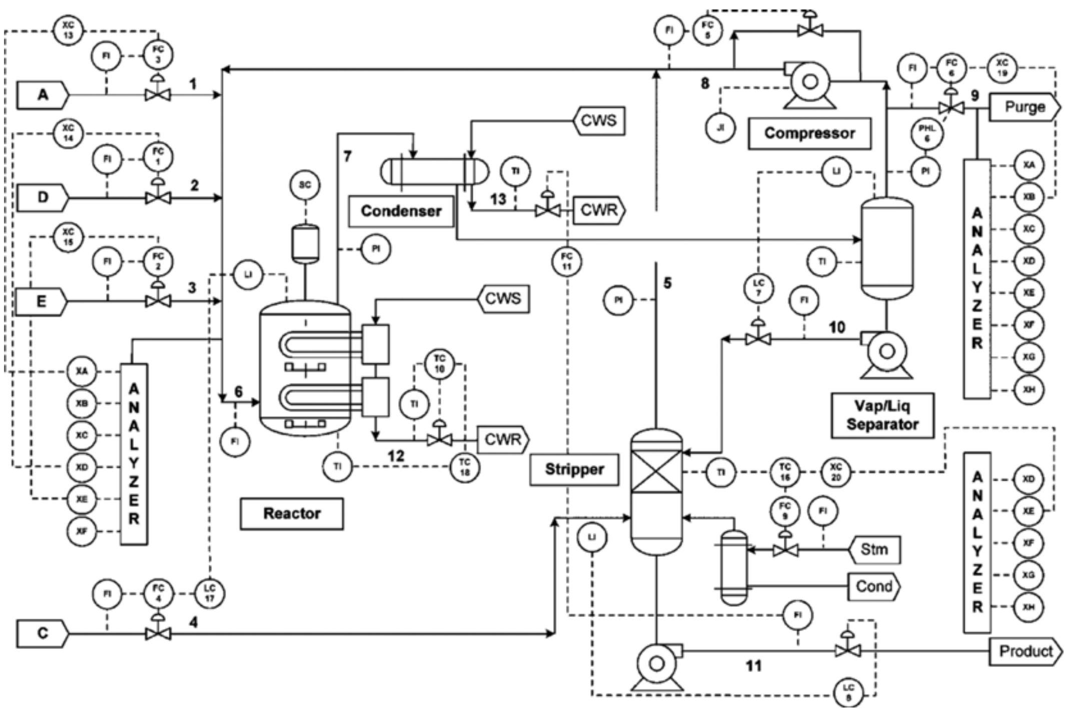

Fig. 4. Control scheme for the Tennessee Eastman process [11,37].

Table 1. Measured process variables in TE process [11,37]

Variables No.	Process measurements	Variables No.	Process measurements
1	A feed (stream 1)	12	Product separator level
2	D feed (stream 2)	13	Product separator pressure
3	E feed (stream 3)	14	Product separator underflow
4	A and C feed (stream 4)	15	Stripper level
5	Recycle flow (stream 8)	16	Stripper pressure
6	Reactor pressure	17	Stripper underflow (stream 11)
7	Reactor level	18	Stripper temperature
8	Reactor temperature	19	Stripper steam flow
9	Purge rate (stream 9)	20	Reactor cooling water outlet temp
10	Product separator temperature	21	Separator cooling water outlet temp
11	A feed (stream 1)	22	Reactor cooling water valve

Table 2. Process faults for the tennessee eastman process [11,37]

No.	Disturbance State	Type
IDV (1)	A/C feed ratio, B composition constant (stream 4)	Step
IDV (2)	B composition, A/C ratio constant (stream 4)	Step
IDV (3)	D feed temperature (stream 2)	Step
IDV (4)	Reactor cooling water inlet temperature	Step
IDV (5)	Condenser cooling water inlet temperature	Step
IDV (6)	A feed loss (stream 1)	Step
IDV (7)	C header pressure loss–reduced availability (stream 4)	Step
IDV (8)	A,B,C feed composition (stream 4)	Random variation
IDV (9)	D feed temperature (stream 2)	Random variation
IDV (10)	C feed temperature (stream 4)	Random variation
IDV (11)	Reactor cooling water inlet temperature	Random variation
IDV (12)	Condenser cooling water inlet temperature	Random variation
IDV (13)	Reaction kinetics	Slow drift
IDV (14)	Reactor cooling water valve	Sticking
IDV (15)	Condenser cooling water valve	Sticking
IDV (16)	Unknown	Unknown
IDV (17)	Unknown	Unknown
IDV (18)	Unknown	Unknown
IDV (19)	Unknown	Unknown
IDV (20)	Unknown	Unknown

ponents are not equally informative to the fault and the T^2 statistics perform differently with each other. The T^2 statistics of the third kernel principal components change most hugely when the fault occurs, indicating that the most fault information is reflected on the KPC. The fault information is quite limited and the useful information is more likely to be inhibited by the useless ones, leading to the situation of useful information being submerged. This is a major cause of the poor performance of T^2 statistics when monitoring fault 1. Consequently, it is very important to highlight the limited fault information reflected on the third kernel principal component to improve monitoring performance. From Fig. 2(b), we can see that along the 3rd KPC the weighting values become large when the fault occurs and the information reflected on the third KPC is highlighted significantly.

The monitoring results of the KPCA and WKPCA-SVDD are presented in Figs. 3(a) and (b), from which both KPCA and WKPCA-SVDD provided enjoyable results; however, the boundary between normal and abnormal statuses is clearer in WKPCA-SVDD, providing a more reliable indication of the process status.

2. Case Study on TE Process

The Tennessee Eastman process is a typical nonlinear benchmark in process engineering, which has been widely used for process monitoring scheme testing [36]. The base control scheme for the TE process is shown in Fig. 4, and the simulation code for the open loop can be downloaded from http://brahms.scs.uiuc.edu. The second plant-wide control structure described by Lyman and Georgakis [37] is implemented to simulate the realistic conditions. The process has 22 continuous process measurements, 19 compositions and 12 manipulated variables, and in the present work, 22 variables are selected for demonstration (Table 1). The simulator can generate 20 types of different faults, listed in Table 2. All the process measurements include Gaussian noise. Once a fault enters the process, it affects almost all state variables in the process [11,37]. For offline modeling, 500 samples under normal condition are collected with a sampling interval of 3 minutes. Eighteen KPCs occupying about 80% cumulative variance contribution are retained in the dominant subspace. The window with h in the WKPCA-SVDD is empirically set as 6. The confidence interval in the T^2 statistics of KPCA is set as 97.5%. Five sets of testing data are generated which are described in the following.

1. Case Study on Fault 0

Fault 0, which is the normal operating condition in the TE process, is used for testing the false alarm performance [11,37]. The

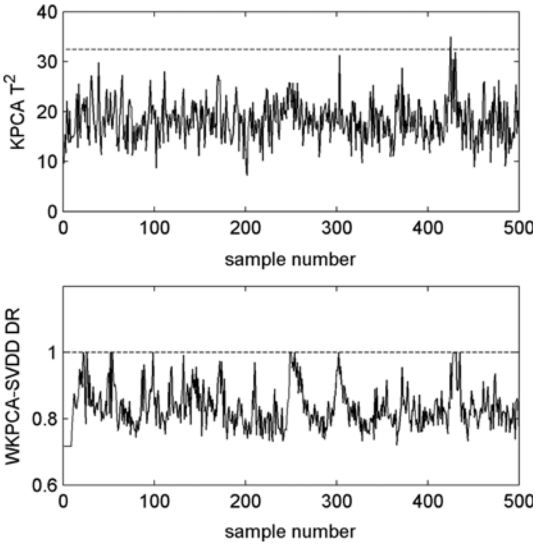

Fig. 5. Monitoring performance of fault 0 using KPCA and WKPCA-SVDD.

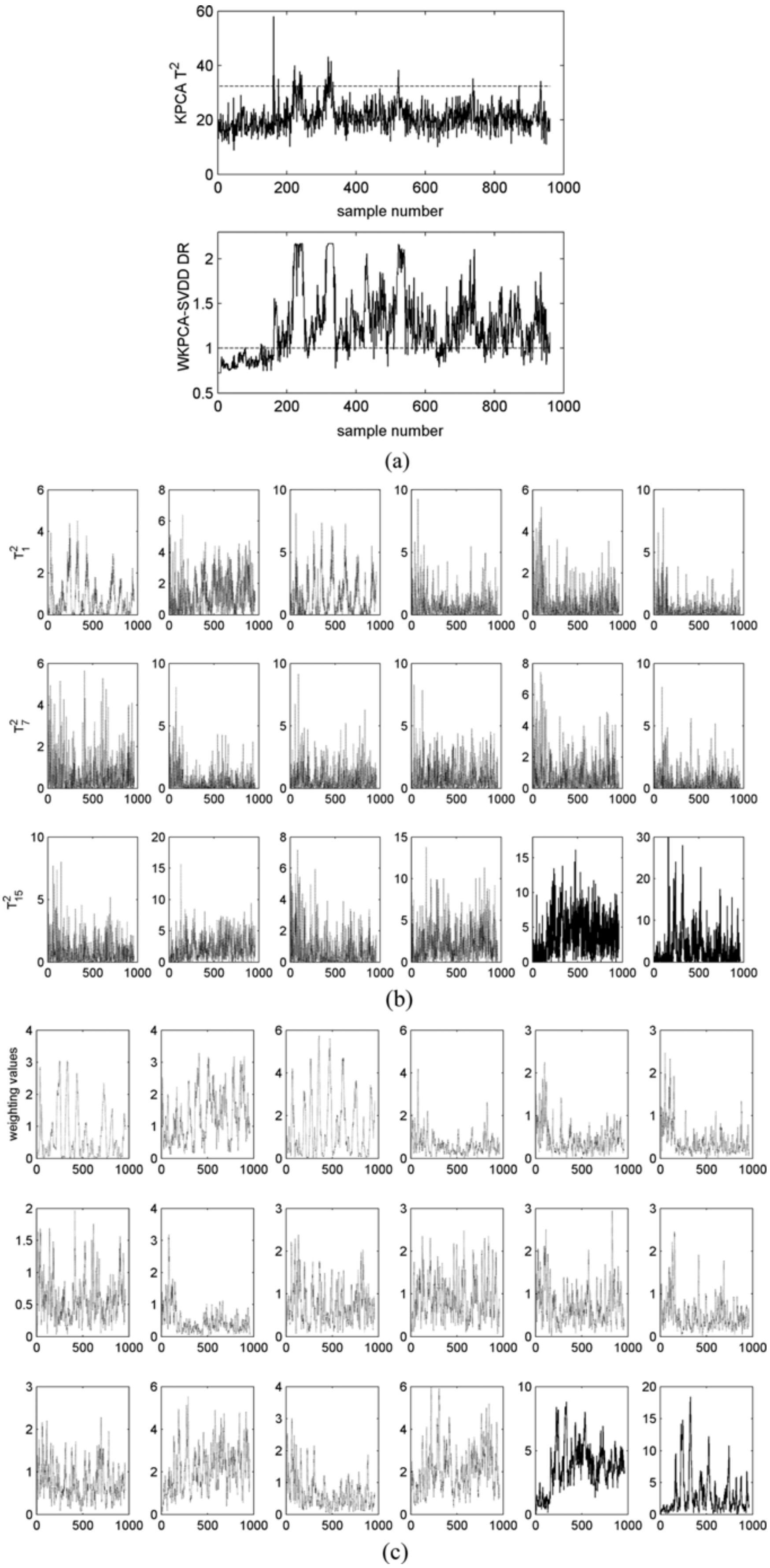

Fig. 6. Monitoring results of fault 4 (a) KPCA and WKPCA-SVDD (b) T^2 of each kernel principal component (c) weighting values.

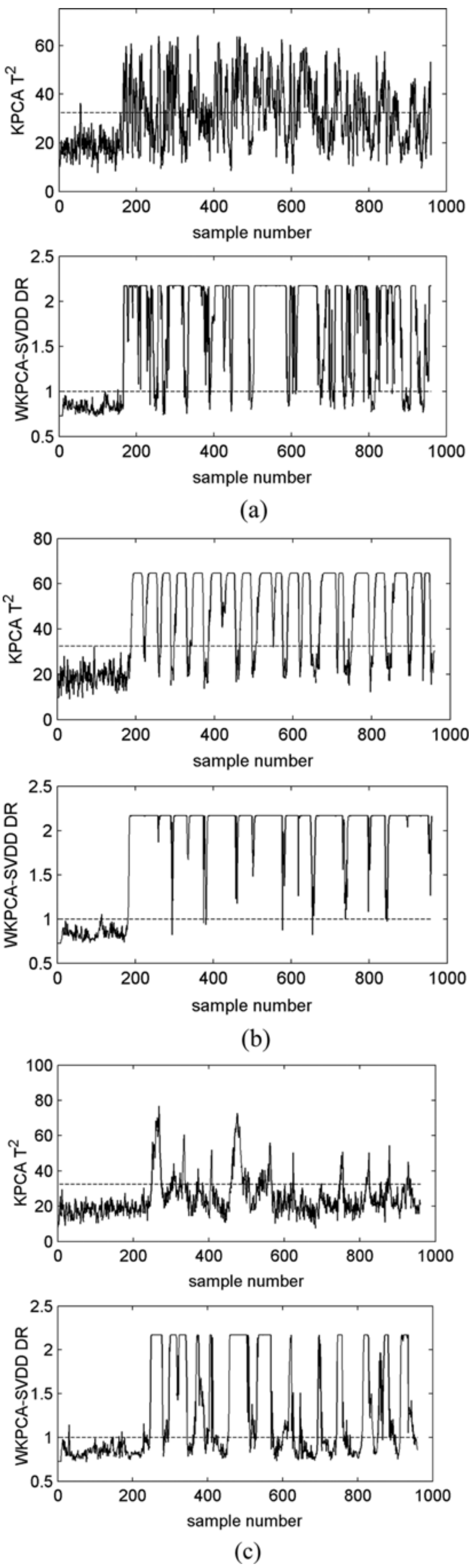

Fig. 7. Monitoring results using KPCA and WKPCA-SVDD (a) fault 11 (b) fault 17 and (c) fault 20.

monitoring performance of KPCA and WKPCA-SVDD for fault 0 is shown in Fig. 5. False alarms in both the KPCA and WKPCA-SVDD methods are low and the normal monitoring performance of the process has not been degraded by introducing the weighting strategy. Actually, the WKPCA-SVDD examines the directions of the first several KPCs with larger variances and these KPCs are not sensitive to process noise. Consequently, the monitoring performance is not easy to be influenced by the noise.

2. Case Studies on Faults 4, 11, 17 and 20

Faults 4 and 11 are the same kinds of faults, which influence the reactor cooling water inlet temperature. Fault 4 involves a step change in the reactor cooling water inlet temperature, and as a significant effect, a step change is introduced into the reactor cooling water flow rate [11,37]. When the fault occurs, there is a sudden temperature increase in the reactor, which is compensated by the control loops. Monitoring performance of fault 4 based on KPCA and WKPCA-SVDD is shown in Fig. 6(a). From Fig. 6(a), KPCA fails to detect the fault, whereas the fault is successfully detected by the WKPCA-SVDD method. To illustrate the cause, the T^2 of each KPC is presented in Fig. 6(b) and the weighting values when monitoring the fault are presented in 6(c). From Fig. 6(b) only few of the KPCs changed significantly while most of the KPCs did not figure out the fault. From Fig. 6(c), the weighting values along the 17^{th} KPC become significantly large when the fault occurs. The fault information reflected on the 17^{th} KPC is highlighted, thereby improving the monitoring performance in the dominant subspace.

Fault 11 is a random variation that induces large oscillations in the reactor cooling water flow rate, which results in a fluctuation of reactor temperature. The other variables are able to remain around the set-points and behave similarly as in the normal operating conditions. Monitoring performance of fault 11 using KPCA and WKPCA-SVDD is shown in Fig. 7(a). The monitoring performance of fault 17 and fault 20 is presented in Figs. 7(b) and (c). From Fig. 7, the monitoring performance has been improved by the WKPCA-SVDD in the monitoring of all the three faults. The non-detection rates have been significantly reduced.

The monitoring results of conventional KPCA and WKPCA for all the faults in TE process are listed in Table 3. For each statistic, the non-detection rates for 17 (faults 3, 9, 15 are not included because there is no observable change in the mean or the variance that can be detected by visually comparing the faults with the normal condition) faults are calculated [11,37]. As shown in Table 3, the

Table 3. Non-detection rates of KPCA and WKPCA-SVDD

Fault No.	KPCA T^2	WKPCA-SVDD DR	Fault No.	KPCA T^2	WKPCA-SVDD DR
1	0.005	0.063	12	0.019	0.003
2	0.019	0.018	13	0.059	0.053
4	0.964	**0.151**	14	0	0.003
5	0.765	**0.619**	16	0.730	**0.408**
6	0	0.002	17	0.249	**0.034**
7	0.596	**0.458**	18	0.1	0.093
8	0.031	0.023	19	0.964	**0.633**
10	0.614	**0.406**	20	0.809	**0.460**
11	0.510	**0.153**			

proposed WKPCA-SVDD outperformed the conventional KPCA method in most cases [27].

CONCLUSIONS

This paper introduces a novel WKPCA-SVDD monitoring method for nonlinear process monitoring performance improvement. WKPCA-SVDD combines the superiority of KPCA for nonlinear monitoring and the advantage of weighting strategy in highlighting the fault information, which can significantly improve the process monitoring performance. In the WKPCA-SVDD method, the T^2 statistic of each kernel principal component is examined to index the most useful information for fault detection. Distinct weighting values are set according to the importance of the kernel principal component in fault detection, which can efficiently deal with the situation of useful information being submerged and reduce non-detection rates of the T^2 statistic. Additionally, the SVDD is employed to provide a comprehensive evaluation of all the kernel principal components, which is more suitable in practice. The proposed method is demonstrated to be effective through both the simulated nonlinear process and the TE benchmark process.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support from the following foundations: 973 project of China (2013CB733600), National Natural Science Foundation of China (21176073) and the Fundamental Research Funds for the Central Universities.

REFERENCES

1. D. Aguado, A. Ferrer, J. Ferrer and A. Seco, *Chemom. Intell. Lab. Syst.*, **85**, 82 (2007).
2. A. AlGhazzawi and B. Lennox, *Control Eng. Pract.*, **16**, 294 (2008).
3. L. H. Chiang, E. L. Russell and R. D. Braatz, *Chemom. Intell. Lab. Syst.*, **50**, 243 (2000).
4. J.-M. Lee, C. Yoo, S. W. Choi, P. A. Vanrolleghem and I.-B. Lee, *Chem. Eng. Sci.*, **59**, 223 (2004).
5. J.-M. Lee, C. Yoo and I.-B. Lee, *J. Process Control*, **14**, 467 (2004).
6. P. Odiowei and Y. Cao, *Chemom. Intell. Lab. Syst.*, **103**, 59 (2010).
7. E. Tomba, P. Facco, F. Bezzo, S. García-Muñoz and M. Barolo, *Chemom. Intell. Lab. Syst.*, **116**, 67 (2012).
8. K. Han, K. J. Park, H. Chae and E. S. Yoon, *Korean J. Chem. Eng.*, **25**, 13 (2008).
9. M. H. Kim and C. K. Yoo, *Korean J. Chem. Eng.*, **25**, 947 (2008).
10. Q. Jiang and X. Yan, *AIChE J.*, **60**, 949 (2014).
11. L. H. Chiang, E. Russell and R. D. Braatz, *Fault detection and diagnosis in industrial systems*, Springer Verlag, London (2001).
12. Q. Jiang and X. Yan, *Chemom. Intell. Lab. Syst.*, **119**, 11 (2012).
13. J. V. Kresta, J. F. Macgregor and T. E. Marlin, *Can. J. Chem. Eng.*, **69**, 35 (2009).
14. Q. Jiang, X. Yan and W. Zhao, *Ind. Eng. Chem. Res.*, **52**, 1635 (2013).
15. Q. Chen, U. Kruger, M. Meronk and A. Leung, *Control Eng. Pract.*, **12**, 745 (2004).
16. I. T. Jolliffe, *Applied Statistics*, 300 (1982).
17. H. Wang, Z. Song and P. Li, *Ind. Eng. Chem. Res.*, **41**, 2455 (2002).
18. S. Wold, *Chemom. Intell. Lab. Syst.*, **23**, 149 (1994).
19. X. B. He, Y. P. Yang and Y. H. Yang, *Chemom. Intell. Lab. Syst.*, **93**, 27 (2008).
20. D. L. Ferreira, S. Kittiwachana, L. A. Fido, D. R. Thompson, R. E. Escott and R. G. Brereton, *Analyst*, **134**, 1571 (2009).
21. Q. Jiang and X. Yan, *Chemom. Intell. Lab. Syst.*, **127**, 121 (2013).
22. C. Alzate and J. A. Suykens, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **32**, 335 (2010).
23. Q. Jiang, X. Yan, Z. Lv and M. Guo, *Korean J. Chem. Eng.*, **30**, 1181 (2013).
24. M. A. Kramer, *AIChE J.*, **37**, 233 (1991).
25. D. Dong and T. J. McAvoy, *Comput. Chem. Eng.*, **20**, 65 (1996).
26. H. Hiden, M. Willis, M. Tham and G. Montague, *Comput. Chem. Eng.*, **23**, 413 (1999).
27. B. Schölkopf, A. Smola and K.-R. Müller, *Neural Computation*, **10**, 1299 (1998).
28. P. Cui, J. Li and G. Wang, *Expert Systems with Applications*, **34**, 1210 (2008).
29. Z. Ge and Z. Song, *Control Eng. Pract.*, **16**, 1427 (2008).
30. Z. Ge, C. Yang and Z. Song, *Chem. Eng. Sci.*, **64**, 2245 (2009).
31. V. H. Nguyen and J.-C. Golinval, *Eng. Struct.*, **32**, 3683 (2010).
32. Z. Ge and Z. Song, *Expert Syst. Appl.*, **38**, 9821 (2011).
33. Y. Yang and J. M. Lee, *J. Process Control*, **23**, 852 (2013).
34. Z. Ge, F. Gao and Z. Song, *J. Process Control*, **21**, 949 (2011).
35. D. M. Tax and R. P. Duin, *Machine Learning*, **54**, 45 (2004).
36. J. J. Downs and E. F. Vogel, *Comput. Chem. Eng.*, **17**, 245 (1993).
37. P. R. Lyman and C. Georgakis, *Comput. Chem. Eng.*, **19**, 321 (1995).