2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

Fault Detection of the Tennessee Eastman Process using Online Reduced Kernel PCA*

Radhia Fazai[1], Majdi Mansouri[2], Okba Taouali[1], Mohamed-Faouzi Harkat[3], Nassreddine Bouguila[1], and Mohamed Nounou[3]

Abstract— **In this paper, we propose an online reduced kernel principal component analysis (KPCA) method for process monitoring. The developed method consists in updating the KPCA model depending on the dictionary which contains linearly independent kernel functions and then using this new reduced KPCA model for process monitoring. The process monitoring performances are studied using Tennessee Eastman Process (TEP). The results demonstrate the effectiveness of the developed online KPCA technique compared to the classical online KPCA method.**

I. INTRODUCTION

Effective operation of various engineering systems requires tight monitoring of some of their key process variables. However, most industrial processes present a large number of highly correlated data, which deals to huge amounts of process data to be analyzed. Multivariate statistical methods such as principal component analysis (PCA) [1], partial least squares [2] have been widely applied for process monitoring. Among these methods, PCA is the most popular one. PCA decomposes the measurement space into a principal component space that contains the normal variability and a residual space that contains the process noise. However, PCA method assumes that the relationships among process variables are linear, which limit their application if these relationships are nonlinear.

To overcome the problem posed by nonlinear data, many nonlinear PCA methods have been developed [3], [4], [5], [6]. Among these methods, kernel PCA (KPCA) [7] stands out because of its simplicity and elegance. The main idea of KPCA is to map measurements from their original space to a higher dimensional feature space via a non nonlinear map where PCA is applied. KPCA has been applied successfully for process monitoring. Similar to the conventional PCA, fault detection with KPCA is done using the SPE and T^2 statistics in the feature space [8], [9], [10], [11].

In the KPCA method, the size of the kernel matrix is given by the number of the training samples. Every time a new sample x is collected, N kernel functions have to be evaluated to form the kernel vector $k(x)$. Therefore, using KPCA model for process monitoring imposes a high computational cost when the training data set is large.

*Qatar National Research Fund (a member of Qatar Foundation)

[1]Laboratory of Automatic Signal and Image Processing, National School of Engineers of Monastir, University of Monastir, Tunisia,

[2]Corresponding author, Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar, Tel:+974.4423.0608, Fax: +974.4423.0065, E-mail: majdi.mansouri@qatar.tamu.edu,

[3]Chemical Engineering Program, Texas A&M University at Qatar, Doha, Qatar.

To overcome the problems stated above, a new method referred as online reduced KPCA for process monitoring is proposed. The method aims to reduce the number of kernel functions. The reduced set of kernel functions is called dictionary. A new criterion based on squared prediction error evaluation determines the relevance of discarding or adding the kernel function to the dictionary $\mathcal{D}_r$. The resulting dictionary contains linearly independent kernel functions.

The proposed method is performed in the online phase and takes into consideration the dynamic behaviors of the systems by changing the model structure. In the proposed online reduced KPCA method, the size of dictionary is variable and depends on the evaluated criterion.

The proposed method forms the dictionary in an online way and then it updates the old KPCA model and uses it for process monitoring. The new method reduces significantly the computation time required to detect faults in nonlinear systems while conserving the structure of the data in the feature space.

The performance of the proposed method is evaluated using a Tennessee Eastman Process (TEP).

The simulation results show that the proposed method provides better detection performances when compared to the online KPCA method.

This paper is organized as follows: Section 2 presents the proposed online reduced KPCA method. Then, in Section 3, the fault detection performance is studied using a TEP. At the end, the conclusions are presented in Section 4.

II. ONLINE REDUCED KPCA

In literature, there are many online kernel methods for fault detection such as online KPCA[12] and SVD-RKPCA [13]. The online KPCA method is performed in two phases. The offline phase consists in determining the initial KPCA model and the SPE index and its control limit. In the online phase, we update only the SPE index. The number of kernel functions in the online KPCA method grows linearly with the number of training data. The large number of kernel functions makes the kernel algorithm unsuitable for online and real time applications. On one hand, it makes the algorithm in danger of over-fitting when the number of kernel functions is too large. On the other hand, the computational complexity grows superlinearly with the number of kernel functions, which leads to a large training time for the update of the algorithm. To solve this problem, a dictionary containing the most relevant independent kernel functions is defined. For this purpose, an approximation criterion is developed to

build the new reduced KPCA model and use it for process monitoring.

A. Approximation Criterion

The (linear) approximation criterion [14], [15], [16] determines the relevance of discarding or accepting the kernel function $\phi(x_k)$ of the measurement x_k at time k in the dictionary $\mathcal{D}_r$. where $\mathcal{D}_r = \{\phi(x_{w_1}), \phi(x_{w_2}), ..., \phi(x_{w_r})\}$, with r is the number of reduced data set and $\{w_1, w_2, ..., w_r\} \subset \{1, 2, ..., N\}$, is the dictionary. The dictionary $\mathcal{D}_r$ contains the most relevant r kernel functions among the N available.

The approximation criterion consists of comparing $\phi(x_k)$ to its projection onto the principal component space spanned by the r elements of the dictionary.

$\phi(x_k)$ is included in the dictionary if the following condition is satisfied :

$$\varepsilon_k = min_\beta \|\phi(x_k) - \sum_{j=1}^{r} \beta_j \phi(x_{w_j})\|_H^2 \geq \nu \qquad (1)$$

where ν is a positive threshold parameter that controls the level of sparsity.

Note that equation (1) ensures the linear independence of the elements of the dictionary $\mathcal{D}_r$.

Simplifying the left part of equation (1), we get:

$$\begin{array}{ll} \varepsilon_k & = min_\beta \sum_{j,i=1}^{r} \beta_j \beta_i k(x_{w_j}, x_{w_i}) \\ & -2\sum_{j=1}^{r} \beta_j k(x_{w_j}, x_k) + k(x_k, x_k) \\ & = min_\beta \beta^T K_r \beta - 2\beta^T k^r(x_k) + k(x_k, x_k) \end{array} \qquad (2)$$

where $K_r \in \mathbf{R}^{r \times r}$ is the reduced Gram matrix, $k^r(x_k)$ is:

$$k^r(x_k) = [k(x_{w_1}, x_k)\, k(x_{w_2}, x_k) \cdots (x_{w_r}, x_k)]^T \qquad (3)$$

and β is the vector of optimal coefficients taking partial derivatives of equation (2) with respect to $\beta_1, \beta_2, ..., \beta_r$, respectively, and let them be zero, which leads to:

$$\beta_k = K_r^{-1} k^r(x_k), \quad k = 1, ..., N \qquad (4)$$

By substituting equation (4) into equation (2), the expression of ε_k can be written as:

$$\varepsilon_k = k(x_k, x_k) - k^r(x_k)^T \beta_k \qquad (5)$$

The resulting dictionary $\mathcal{D}_r$, called $\nu-$approximate, satisfies the following relation [15], [16]:

$$\min_{i=1,...,r} \min_{\beta_1,...,\beta_N} \|\phi(x_{w_i}) - \sum_{\substack{j=1 \\ i \neq j}}^{r} \beta_j \phi(x_{w_j})\| \geq \sqrt{\nu} \qquad (6)$$

B. Online monitoring using reduced KPCA

In this section, online monitoring using reduced KPCA algorithm is derived. The parameters of the algorithm are updated only when the dictionary changes. This includes the update of the kernel matrix, its inverse, the number of principal components, the SPE index and its control limit.

For a new observation x_k at instant k, an approximation criterion and evaluation of SPE index determine whether the dictionary remains unchanged or should be updated. Indeed, the dictionary is updated if the new kernel function $\phi(x_k)$ is added to the dictionary since it cannot be well approximated by a linear combination of the dictionary elements and when the observation x_k is not faulty.

For normal process change ($SPE_k^r < SPE_{\alpha,k-1}^r$), two cases may arise depending on the condition as in equation (1):

Case 1: unchanged dictionary ($\epsilon_k < \nu$)

In this case, the new kernel function $\phi(x_k)$ is not added to the dictionary which remains unchanged.

Using the dictionary $\mathcal{D}_r$, the expression of the k^{th} eigenvector V_k^r can be written as:

$$V_k^r = \sum_{i=1}^{r} \alpha_{k,i}^r \phi(x_{w_i}), \quad k = 1, ..., N \qquad (7)$$

The kernel vector $k^r(x_k)$ in this case is given by:

$$k^r(x_k) = \left[\ k(x_{w_1}, x_k) \quad k(x_{w_2}, x_k) \quad . \quad . \quad . \quad k(x_{w_r}, x_k)\ \right]^T \qquad (8)$$

Case 2: changed dictionary ($\epsilon_k \geq \nu$)

In this case, the kernel function $\phi(x_k)$ is added to the dictionary $\mathcal{D}_r$. So, the new dictionary becomes:
$\mathcal{D}_r = \left[\ \phi(x_{w_1}) \quad \phi(x_{w_2}) \quad . \quad . \quad . \quad \phi(x_{w_r}) \quad \phi(x_{w_{r+1}})\ \right]^T$.
The j^{th} eigenvector at instant k is given by:

$$V_k^r = \sum_{i=1}^{r+1} \alpha_{k,i}^r \phi(x_{w_i}), \quad k = 1, ..., N \qquad (9)$$

where $w_{r+1} = k$.

In this case, the reduced kernel matrix $K_{r+1} \in \mathbf{R^{r+1 \times r+1}}$ is updated as follows:

$$K_{r+1} = \begin{bmatrix} K_r & k^r(x_k) \\ k^r(x_k)^T & k(x_k, x_k) \end{bmatrix} \qquad (10)$$

The SPE^r for an observation x_k using the dictionary $\mathcal{D}_r$ is defined as:

$$SPE^r = k(x_k, x_k) - (k^r(x_k))^T C^r (k^r(x_k)) \qquad (11)$$

where

$$\begin{cases} k^r(x_k) = \left[\ k(x_{w_1}, x_k) \quad k(x_{w_2}, x_k) \quad . \quad . \quad . \quad k(x_{w_{r+1}}, x_k)\ \right]^T \\ C^r = P^r(\Lambda^r)^{-1}(P^r)^T \\ \Lambda^r = diag(\lambda_{\ell_1} ... \lambda_{\ell_{r+1}}) \\ P^r = \left[\ \bar{\alpha^r}_{\ell_1} \quad \bar{\alpha^r}_{\ell_2} \quad . \quad . \quad . \quad \bar{\alpha^r}_{\ell_{r+1}}\ \right] \end{cases} \qquad (12)$$

where, ℓ_{r+1} is the number of principal components using the dictionary $\mathcal{D}_r$. Note that when introducing a new observation in the dictionary, to update the vector of the coefficients β, the inverse of the kernel matrix K_{r+1}^{-1} is calculated. But, the calculation of this inverse can increase the computational complexity. To overcome this problem, the inverse of the kernel matrix is updated iteratively using the Woodbury matrix identity:

$$\begin{aligned} K_{r+1}^{-1} &= \begin{bmatrix} K_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\varepsilon_k} \begin{bmatrix} -K_r^{-1} k^r(x_k) \\ 1 \end{bmatrix} \begin{bmatrix} -k^r(x_k)^T K_r^{-1} & 1 \end{bmatrix} \\ &= \begin{bmatrix} K_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\varepsilon_k} \begin{bmatrix} -\beta_k \\ 1 \end{bmatrix} \begin{bmatrix} -\beta_k^T & 1 \end{bmatrix} \end{aligned} \tag{13}$$

Thus, the vector of the coefficients β is updated as follows:

$$\beta_k = K_{r+1}^{-1} k^r(x_k). \tag{14}$$

In case of abnormal process changes ($SPE_k^r > SPE_{\alpha,k-1}^r$), the dictionary D_r remains unchanged,

$$\mathcal{D}_r = \begin{bmatrix} \phi(x_{w_1}) & \phi(x_{w_2}) & . & . & . & \phi(x_{w_r}) \end{bmatrix}^T.$$

The algorithm which studies the developed online reduced KPCA is presented in Algorithm 1.

Algorithm 1: Online reduced KPCA algorithm

Inputs: $N \times m$ data matrix X.

1) Initialization: $k = 1$, $r = 1$, SPE_k^r index, the control limit $SPE_{\alpha,k}^r$, the initial eigenvalue $\lambda_{\ell_{r,k}} = 1$, $w_k = 1$, $\ell_{r,k} = 1$, $K_k = k(x_k, x_k)$, $\beta_k = 1$;
2) $k \leftarrow k + 1$, ;
3) Obtain a new normal sample x_k and scale it;
4) Determine $k^r(x_k) = \begin{bmatrix} k(x_{w_1}, x_k) & . & . & . & k(x_{w_r}, x_k) \end{bmatrix}^T$ and scale it;
5) Compute the monitoring index SPE_k^r for the new data x_k;
   While $SPE_k^r < SPE_{\alpha,k-1}^r$, then
   a) $\beta_k \leftarrow K_r^{-1} k^r(x_k)$;
   b) $\varepsilon_k \leftarrow k(x_k, x_k) - k^r(x_k)^T \beta_k$;
      If $\epsilon_k \geq \nu$
      i) $r \leftarrow r + 1$ and $w_r \leftarrow k$,
      ii) Update the kernel matrix K_{r+1} and its inverse K_{r+1}^{-1} then scale them,
      iii) Update $\ell_{r,k}$, V_k^r, $SPE_{\alpha,k-1}^r$
      Else, dictionary remain unchanged
   Go to 2.

III. Fault Detection of the Tennessee Eastman Process using Online Reduced KPCA

The Tennessee Eastman Process (TEP) is a benchmark simulation model of a complex industrial chemical process. It was proposed by Downs and Vogel [25]. The TEP contains eight components: G and H are the main products, A, C, D and E are reactants, F is a byproduct and B is an inert component. The reactions that take place in the reactor are:

$$\begin{aligned} A(g) + C(g) + D(g) &\rightarrow G(liq) \\ A(g) + C(g) + E(g) &\rightarrow H(liq) \\ A(g) + E(g) &\rightarrow F(liq) \\ 3D(g) &\rightarrow 2F(liq) \end{aligned} \tag{15}$$

This process contains five major units: a reactor, a stripper, a condenser, a recycle compressor, and a separator. The TEP contains 41 measured variables (19 composition measurements and 22 continuous process measurements) and 12 manipulated variables. Details of TEP including the reactions are widely detailed in literature [17] (Figure 1).

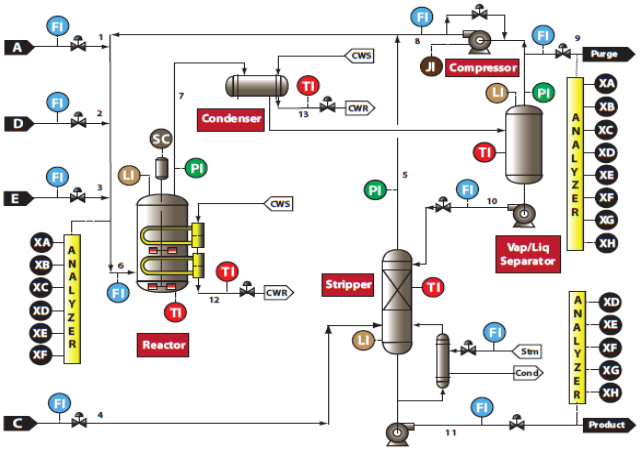

Fig. 1. Tennessee Eastman Process flow diagram

20 variables are considered for TEP monitoring and these variables are listed in Table I. In this study, 1000 observations are generated for process monitoring. The kernel parameter of the RBF function is set to 32. The number of kernel principal components is chosen using the CPV method, such that the first eigenvalues represent about 95% of the total variance. The confidence level is equal to 95%.

TABLE I
List of the monitoring variables in the TEP.

Variables	Measured variables	Unit
x_1	A feed(stream 1)	km^3/h
x_2	D feed(stream 1)	kg/h
x_3	E feed(stream 1)	kg/h
x_4	Total feed(stream 4)	km^3/h
x_5	Recycle flow(stream 8)	km^3/h
x_6	Reactor feed rate(stream 6)	km^3/h
x_7	Reactor pressure	kPa
x_8	Reactor level	%
x_9	Reactor temperature	^0C
x_{10}	Purge rate (stream 9)	km^3/h
x_{11}	Product separator temperature	^0C
x_{12}	Product separator level	%
x_{13}	Product separator pressure	kPa
x_{14}	Product separator underflow (stream 10)	m^3/h
x_{15}	Stripper level	%
x_{16}	Stripper pressure	kPa
x_{17}	Stripper underflow(stream 11)	m^3/h
x_{18}	Stripper temperature	^0C
x_{19}	Stripper stream flow	kg/h
x_{20}	Compressor work	kW

The developed online reduced KPCA algorithm is applied to TEP to determine the reduced number of observations for monitoring. Figures 2 and 3 present, respectively, the variation of the reduced number of observations as well as the number of principal components under normal operation

using the proposed method. We can show from these figures that the number of observations is reduced to 48 ($\nu = 0.18$).

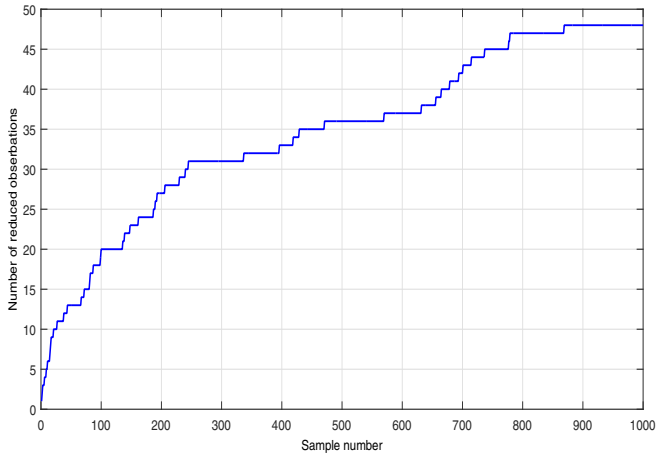

Fig. 2. Variation of the reduced number of observations.

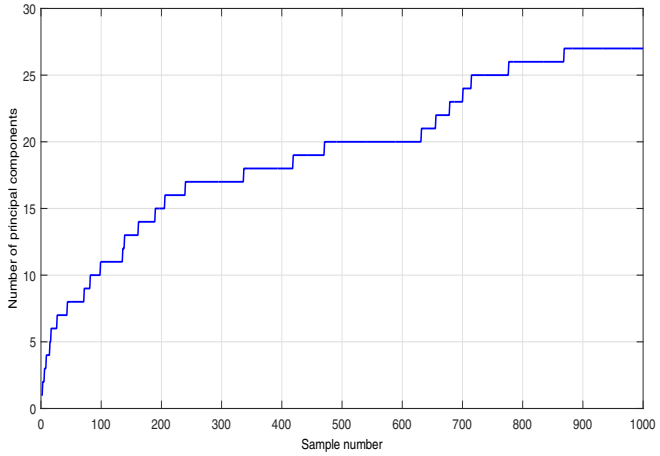

Fig. 3. Variation of the number of principal components.

The monitoring results under normal operation using both online reduced KPCA and online KPCA methods are presented in Figures 4 and 5. We can see from these figures that the online reduced KPCA method (Figure 13) provides further reduction in false alarm rate compared to the online KPCA (Figure 4), which can be explained by the fact that the online KPCA model is not able to model the dynamic TEP efficiently.

To further compare the performance of the online reduced KPCA and the online KPCA methods, the elapsed time at each iteration as well as the false alarm rates are computed. It should be noted, from Table II and Figure 6, that the online KPCA method returns higher elapsed time and false alarm rate compared to the online reduced KPCA. The elapsed time of the proposed method is reduced to 97.49% using a number of kernel functions which is equivalent to 4.9% of the number of initial data. Table II shows also that the online reduced KPCA method has less false alarm rate (0.5%) compared to the online KPCA method (3.4%).

Next, a bias fault is injected to the variable x_2 with an

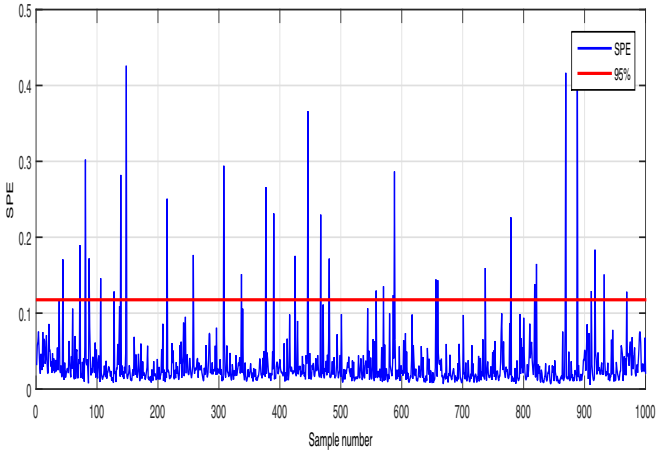

Fig. 4. Time evolution of SPE index under normal operation.

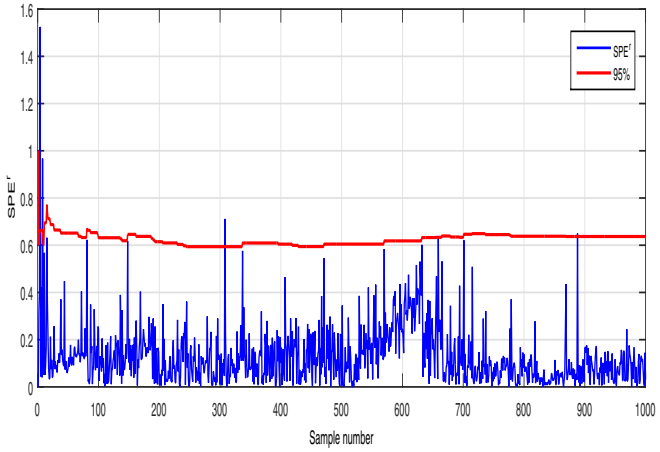

Fig. 5. Time evolution of SPE^r under normal operation.

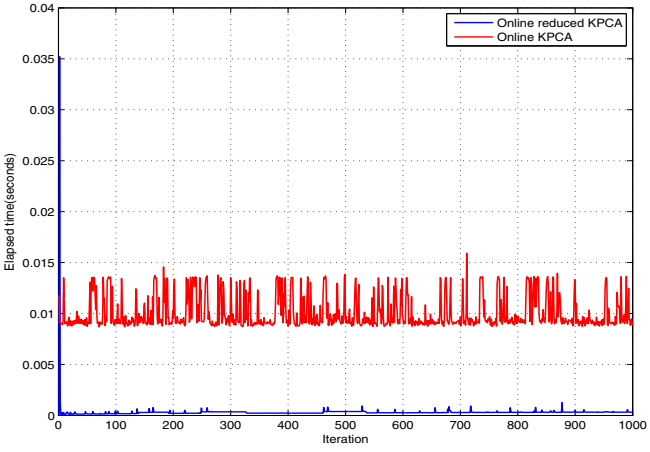

Fig. 6. Variation of the elapsed time

TABLE II
SUMMARY OF FALSE ALARM RATES AND ELAPSED TIME

	Online reduced KPCA	Online KPCA
FAR (%)	0.3	3.4
ET (seconds)	0.471750	18.798056

amplitude equal to 43% of the range of the variation of this variable between the instant 300 and 500. In this case, the variation of the reduced number of observations as well as the number of principal components using the proposed method are presented respectively in Figures 7 and 8.

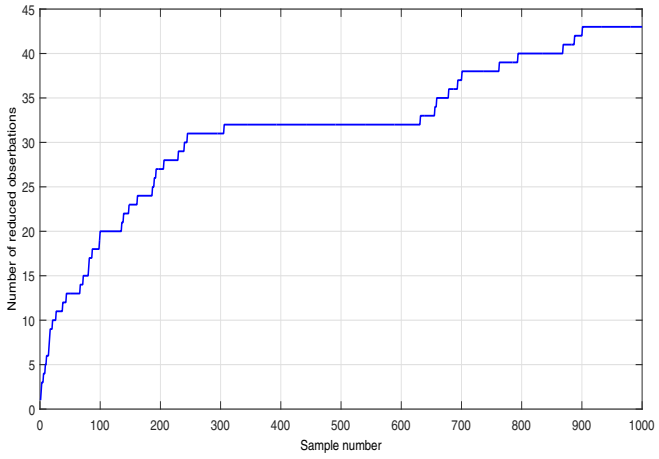

Fig. 7. Variation of the reduced number of observations

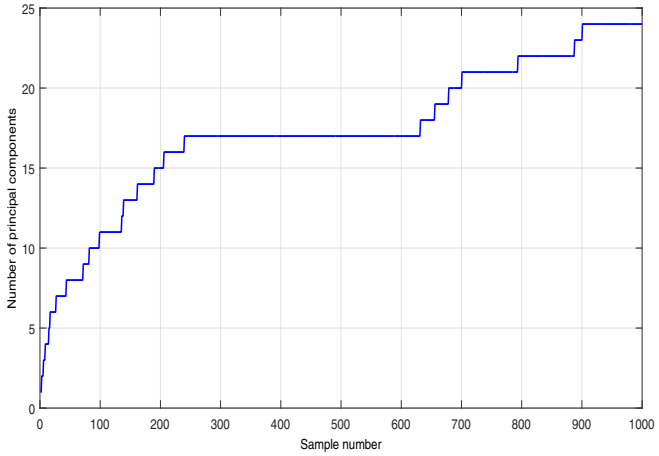

Fig. 8. Variation of the number of principal components

In the presence of a bias fault in x_2, we can show from Figures 9 and 10 the monitoring results of TEP using the online KPCA (SPE) and the proposed online reduced KPCA (SPE^r) techniques. Both online KPCA and online reduced KPCA methods are able to detect the fault but with high false alarm and miss detection rates when using the online KPCA (Figure 9).

Table III shows the detection performances in terms of FAR, GDR and ET for both KPCA and online reduced KPCA methods. The results show that the ET required to perform fault detection with the online reduced KPCA method decreased because of the reduction of the size of the dictionary. The proposed online reduced KPCA provides a reduction of FAR (1.47%) and an increase of GDR (97.01%).

IV. CONCLUSIONS

In this paper, we proposed an online reduced KPCA method for chemical process monitoring. A comparison between the new online reduced KPCA method and the online

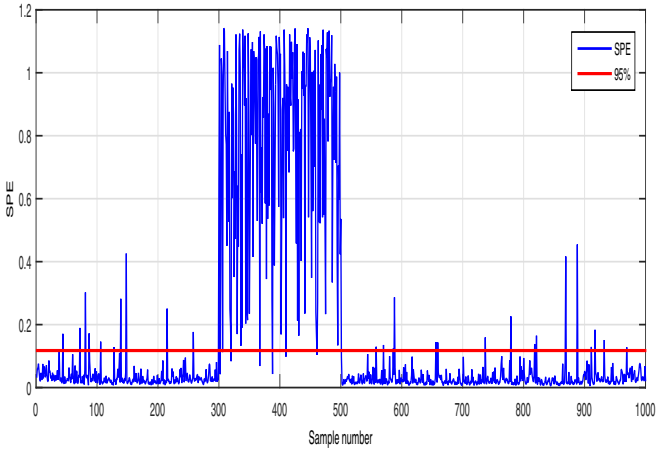

Fig. 9. Time evolution of SPE index with a fault on the variable x_2.

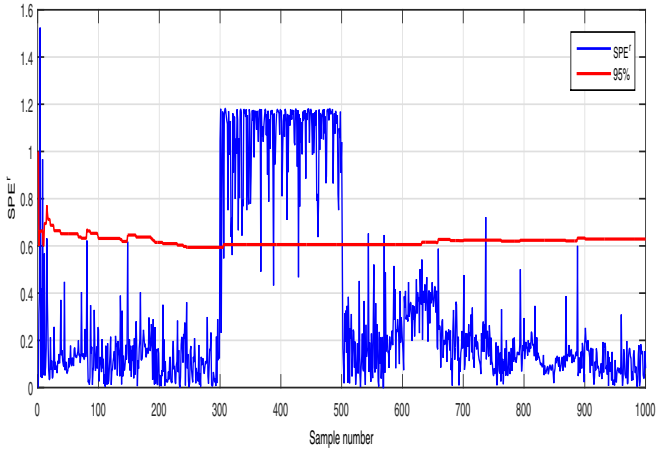

Fig. 10. Time evolution of SPE^r with a fault on the variable x_2

TABLE III
SUMMARY OF GOOD DETECTION RATES, FALSE ALARM RATES, AND ELAPSED TIME FOR TEP DATA

	Online reduced KPCA	Online KPCA
FAR (%)	1.47	6.54
GDR (%)	97.01	96.52
ET (seconds)	0.426162	30.338076

KPCA method is done using a Tennessee Eastman Process (TEP) in order to evaluate the performance of the proposed method. The simulation results show that the developed online reduced KPCA method outperforms the online KPCA method in terms of good detection rate, false alarm rate and elapsed time.

ACKNOWLEDGMENT

This work was supported by Qatar National Research Fund (a member of Qatar Foundation) under the NPRP grant NPRP9 − 330 − 2 − 140.

REFERENCES

[1] M.-F. Harkat, M. Mansouri, M. Nounou, and H. Nounou, "Enhanced data validation strategy of air quality monitoring network," *Environmental Research*, vol. 160, pp. 183–194, 2018.

[2] S. J. Zhao, J. Zhang, and Y. M. Xu, "Performance monitoring of processes with multiple operating modes through multiple pls models," *Journal of process Control*, vol. 16, no. 7, pp. 763–772, 2006.

[3] R. Baklouti, M. Mansouri, M. Nounou, H. Nounou, and A. B. Hamida, "Iterated robust kernel fuzzy principal component analysis and application to fault detection," *Journal of Computational Science*, vol. 15, pp. 34–49, 2016.

[4] M. Mansouri, M. Nounou, H. Nounou, and K. Nazmul, "Kernel pca-based glrt for nonlinear fault detection of chemical processes," *Journal of Loss Prevention in the Process Industries*, vol. 26, no. 1, pp. 129–139, 2016.

[5] I. B. Khediri, M. Limam, and C. Weihs, "Variable window adaptive kernel principal component analysis for nonlinear nonstationary process monitoring," *Computers & Industrial Engineering*, vol. 61, no. 3, pp. 437–446, 2011.

[6] Q. Jiang and X. Yan, "Weighted kernel principal component analysis based on probability density estimation and moving window and its application in nonlinear chemical process monitoring," *Chemometrics and Intelligent Laboratory Systems*, vol. 127, pp. 121–131, 2013.

[7] J. Fan and Y. Wang, "Fault detection and diagnosis of non-linear non-gaussian dynamic processes using kernel dynamic independent component analysis," *Information Sciences*, vol. 259, pp. 369–379, 2014.

[8] J.-M. Lee, C. Yoo, and I.-B. Lee, "Fault detection of batch processes using multiway kernel principal component analysis," *Computers & Chemical Engineering*, vol. 28, no. 9, pp. 1837–1847, 2004.

[9] S. W. Choi, C. Lee, J.-M. Lee, J. H. Park, and I.-B. Lee, "Fault detection and identification of nonlinear processes based on kernel pca," *Chemometrics and Intelligent Laboratory Systems*, vol. 75, no. 1, pp. 55–67, 2005.

[10] V. H. Nguyen and J.-C. Golinval, "Fault detection based on kernel principal component analysis," *Engineering Structures*, vol. 32, no. 11, pp. 3683–3691, 2010.

[11] R. Fazai, O. Taouali, M. F. Harkat, and N. Bouguila, "A new fault detection method for nonlinear process monitoring," *International Journal of Advanced Manufacturing Technology*, vol. 87, no. 9-12, pp. 3425–3436, 2016.

[12] J.-M. Lee, C. Yoo, S. W. Choi, P. A. Vanrolleghem, and I.-B. Lee, "Nonlinear process monitoring using kernel principal component analysis," *Chemical Engineering Science*, vol. 59, no. 1, pp. 223–234, 2004.

[13] I. Jaffel, O. Taouali, M. F. Harkat, and H. Messaoud, "Kernel principal component analysis with reduced complexity for nonlinear dynamic process monitoring," *The International Journal of Advanced Manufacturing Technology*, vol. 88, no. 9-12, pp. 3265–3279, 2017.

[14] P. Honeine, "Online kernel principal component analysis: A reduced-order model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1814–1826, 2012.

[15] ——, "Approximation errors of online sparsification criteria," *IEEE Transactions on Signal Processing*, vol. 63, no. 17, pp. 4700–4709, 2015.

[16] ——, "Analyzing sparse dictionaries for online learning with kernels," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6343–6353, 2015.

[17] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Computers & Chemical Engineering*, vol. 17, no. 3, pp. 245–255, 1993.