

Nonlinear process monitoring based on new reduced Rank-KPCA method

Hajer Lahdhiri¹ · Ilyes Elaissi¹ · Okba Taouali¹ · Mohamed Faouzi Harkat² · Hassani Messaoud¹

Published online: 9 October 2017
© Springer-Verlag GmbH Germany 2017

Abstract Kernel Principal Component Analysis (KPCA) is an efficient multivariate statistical technique used for nonlinear process monitoring. Nevertheless, the conventional KPCA suffers high computational complexity in dealing with large samples. In this paper, a new kernel method based on a novel reduced Rank-KPCA is developed to make up for the drawbacks of KPCA. The basic idea of the proposed novel approach consists at first to construct a reduced Rank-KPCA model that describes properly the system behavior in normal operating conditions from a large amount of training data and after that to monitor the system on-line. The principle of the proposed Reduced Rank-KPCA is to eliminate the dependencies of variables in the feature space and to retain a reduced data from the original one. The proposed monitoring method is applied to fault detection in a numerical example, Continuous Stirred Tank Reactor and air quality-monitoring network AIRLOR and is compared with conventional KPCA and Moving Window KPCA methods.

Keywords Reduced Rank-KPCA · Nonlinear process monitoring · Fault detection

1 Introduction

The modern manufacturing industries require higher product quality and safety operations. In order to ensure the suitable functioning system and to minimize downtime in case of failure, it is important to detect process faults as early as possible (Baklouti et al. 2016). Several process monitoring-based Multivariate Statistical Process (MSP) methods were developed thanks to their efficiencies and simplicities (Botre et al. 2016; Mansouri et al. 2016). In fact, these methods are based on the analysis of process history data in order to model the relationships between variables (Chetouani 2008). Principal Components Analysis (PCA) (Tharrault et al. 2008; Jaffel et al. 2014) and Partial Least Squares (PLS) (Li et al. 2010; Taouali et al. 2015) are the two commonly used methods in the field of process fault detection and diagnosis. More recently Independent Component (ICA) (Lee et al. 2004a, b; Zhao et al. 2008) have been proposed for this purpose.

The PCA method (Harkat et al. 2006) is a statistical method widely used of analysis and dimensional reduction. It is used to identify the dependency structure between the observations in order to transform the input space into a smaller dimensional space while retaining the maximum variance of the input data. The PCA technique produces new uncorrelated variables called principal components (PC) with each component is linear combinations of original variables. However, the majority of chemical and biological processes data have nonlinear relationships. In fact, PCA only defines a linear projection of data. Hence, it is incapable to analyze and represent the data with nonlinear characteristics. This limitation and nonlinearity problem have motivated various researchers to develop nonlinear extensions, such as nonlinear PCA that combined principal curve and Neural Network (NN) (Harkat et al.

✉ Okba Taouali
taoualiok@gmail.com

¹ Research Laboratory of Automation, Signal Processing and Image (LARATSI), National School of Engineering Monastir, Rue Ibn ELJazzar, 5019 Monastir, Tunisia

² Department of Electronics, Faculty of Engineering Annaba, Badji Mokhtar, BP. 12, 23000 Annaba, Algeria

2010; Dong and McAvoy 1996) and Kernel Principal Components Analysis (KPCA) that define a nonlinear generalization of the conventional linear PCA (Schölkopf et al. 1998; Mika et al. 1998).

The KPCA method is a simple and interesting technique developed by Schölkopf to faithfully model the nonlinear relationships between process data. Using the concept of kernel tricks (Aronszajn 1950; Aizerman et al. 1964), KPCA can efficiently project the input data with linearly inseparable structure onto a higher dimensional feature space in which the data becomes linearly separable and then perform the conventional PCA in the feature space. KPCA requires only solving an eigenvalue problem without any nonlinear optimization compared to other nonlinear PCA techniques based on NN. Moreover, KPCA was applied in several applications and showed a good results in terms of analysis, modeling and fault detection accuracies: face recognition (Kim et al. 2002) speech recognition (Amaro et al. 2005) nonlinear process monitoring and fault diagnosis (Lee et al. 2004a, b; Fazai et al. 2016; Cho et al. 2005; Kazor et al. 2016; Sheriff et al. 2017).

However, it suffers from two major drawbacks. Firstly, the KPCA is not a sparse model (the computation time of projections in feature space and memory increase with the number of training data). Secondly, the conventional KPCA relies on a time invariant model thus not adapted for monitoring time-varying industrial processes. To overcome these disadvantages, in this paper we proposed a novel method to reduce the order of KPCA model and to guarantee a good fault detection.

This article is organized as follows. In Sect. 2, a brief presentation KPCA and its fault detection index. Section 3 presents the new proposed approach called Reduced Rank-KPCA and its principle. In Sect. 4, we develop two online algorithms for dealing with fault detection of nonlinear dynamic process. Section 5 demonstrates the efficiency of the proposed method compared to classical methods through three examples. Conclusions are given at the end of the paper.

2 Previews work

2.1 Review of kernel principal component analysis (KPCA) description

Kernel Principal Component Analysis (KPCA) is among the most popular dimensional reduction and analysis method (Li and Zhang 2013). It extends the linear PCA to deal with nonlinear modes. The main objective of Kernel PCA is to model process data with non-linear structure. It consists to transform the nonlinear part of input data space E into a linear part in a new high dimensional feature space

denoted H and to perform PCA in that space. The feature space H is nonlinearly transformed from the input space E with a non-linear mapping function ϕ . The mapping of sample $x \in E$ in the feature space H can be written as:

$$\begin{aligned} \phi : E \subset \mathbb{R}^m &\rightarrow H \subset \mathbb{R}^H \\ x &\rightarrow \phi(x) \end{aligned} \quad (1)$$

Let us consider $X = [x_1, \dots, x_i, \dots, x_N]^T$ the training data matrix scaled to zero mean and unit variance. Where $x_i \in E \subset \mathbb{R}^m$ is a data vector, N is the number of observation samples and m is the number of process variables.

The monitoring phase based on linear PCA approach requires the selection of principal components that maximize the variance in the data set. This is done by the Eigen-decomposition of the covariance matrix. Similarly, this approach was generalized in Kernel PCA approach by scholkopf (Schölkopf et al. 1998). The covariance matrix C_ϕ in the feature space H is given by:

$$C_\phi = \frac{1}{N} \sum_{i=1}^N \phi(x_i) \phi(x_i)^T \quad (2)$$

Let $\chi = [\phi(x_1) \dots \phi(x_i) \dots \phi(x_N)]^T \in \mathbb{R}^{N \times h}$ define the data matrix in the feature space H , then C_ϕ can be expressed as:

$$C_\phi = \frac{1}{N} \chi^T \chi \quad (3)$$

The principal components of the mapped data $\phi(x_1) \dots \phi(x_i) \dots \phi(x_N)$ are computed by solving the eigenvalue decomposition of C_ϕ such that:

$$\lambda_j \mu_j = C_\phi \mu_j \quad \text{with } j = 1, \dots, h \quad (4)$$

with μ_j is j th eigenvector and λ_j is the associated j th eigenvalue. For $\lambda_j \neq 0$ there exist coefficients $\alpha_{i,j}$; $i = 1 \dots N$ such all eigenvectors μ_j can be considered as a linear combination of $[\phi(x_1) \phi(x_2) \dots \phi(x_N)]$ and can be expressed by:

$$\mu_j = \sum_{i=1}^N \alpha_{i,j} \phi(x_i) \quad (5)$$

However, in practice, the mapping function ϕ is not defined and then the covariance matrix C_ϕ in the feature space cannot be calculated implicitly. Thus, instead of solving eigenvalue problem directly on C_ϕ , we apply the kernel trick used firstly for Support Vector Machine (SVM) (Vapnik 1999). The inner product given in Eq. (2) may be calculated by a kernel function $\mathbf{k}(\cdot, \cdot)$ that satisfy Mercer's theorem (Mercer 1909) as follow:

$$\langle \phi(x), \phi(x') \rangle_H = \mathbf{k}(x, x') \quad \forall x, x' \in \mathbb{R}^m \quad (6)$$

Let define a kernel matrix K associated to a kernel function \mathbf{k} as:

$$K = \chi\chi^T = \begin{bmatrix} \mathbf{k}(x_1, x_1) & \cdots & \mathbf{k}(x_1, x_N) \\ \vdots & \ddots & \vdots \\ \mathbf{k}(x_N, x_1) & \cdots & \mathbf{k}(x_N, x_N) \end{bmatrix} \in \mathbb{R}^{N \times N} \tag{7}$$

Applying the kernel matrix may reduce the problem of eigenvalue decomposition of C_ϕ . Hence, Eigen-decomposition of the kernel matrix K is equivalent to performing PCA in \mathbb{R}^H , so that:

$$NAV = KV \tag{8}$$

With: A is the diagonal matrix of eigenvalues λ_j arranged in descending order.

$$A = \text{diag}(\lambda_1 \dots \lambda_j \dots \lambda_N) \tag{8.a}$$

And V is the matrix of their corresponding eigenvectors.

$$V = [\alpha_1 \dots \alpha_j \dots \alpha_N] \tag{8.b}$$

Since the principal components are orthonormal, it is required to guarantee the normality of μ_j in Eq. (4), such that:

$$\langle \mu_j, \mu_j \rangle_H = 1; \quad j = 1 \dots n \tag{9}$$

With n the number of the first non-zero Eigen-values.

Using Eqs. (5) and (9) leads to:

$$\begin{aligned} \langle \mu_j, \mu_j \rangle_H &= \sum_{i,k}^N \alpha_{i,j} \alpha_{k,j} \langle \phi(x_i), \phi(x_k) \rangle_H \\ &= \sum_{i,k}^N \alpha_{i,j} \alpha_{k,j} K_{i,k} \\ &= \langle \alpha_j, K \alpha_j \rangle_H \\ &= \lambda_j \langle \alpha_j, \alpha_j \rangle_H \end{aligned} \tag{10}$$

where: $K_{i,k} = \mathbf{k}(x_i, x_k)$. The corresponding eigenvectors α_j must be scaled as:

$$\langle \alpha_j, \alpha_j \rangle_H = \|\alpha_j\|^2 = \frac{1}{\lambda_j} \quad j = 1, \dots, n \tag{11}$$

Many kernel functions have been proposed in literature i.e.:

- Laplacian kernel:

$$\mathbf{k}(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|}{\sigma}\right) \tag{12}$$

- Gaussian kernel [Radial Basis Function (RBF)]:

$$\mathbf{k}(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{13}$$

- Polynomial function:

$$\mathbf{k}(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d \tag{14}$$

where d is a positive integer.

- Sigmoid function

$$\mathbf{k}(x_i, x_j) = \tanh(\beta_0 \langle x_i, x_j \rangle + \beta_1) \tag{15}$$

where β_0 and β_1 are fixed by the user to satisfy Mercer’s theorem.

Finally, in Eq. (2) we have implicitly assumed that $\sum_{i=1}^N \phi(x_i) = 0$. Generally, this is not the case, this leads to the normalization of the kernel matrix, where we replacing K by the Gram matrix G as follows:

$$G = K - 1_N K - K 1_N + 1_N K 1_N \tag{16}$$

with: $1_N = \frac{1}{N} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{N \times N}$.

2.2 Number of principal components

Determining the number of retained principal components (ℓ) is the important step of modeling based on KPCA. The cumulative percent variance (CPV) has been proposed to compute the retain PC (ℓ) (Zhang et al. 2012) and (Jaffel et al. 2016a, b). The cumulative percent variance (CPV) is the sum of the first ℓ eigenvalues divided by their total variations. It can be expressed as:

$$CPV(\ell) = \frac{\sum_{j=1}^{\ell} \lambda_j}{\sum_{j=1}^m \lambda_j} 100\% \tag{17}$$

The number ℓ of retained PCs is chosen if the CPV is higher to 95%.

2.3 Fault detection

Like in PCA approach, the squared prediction error (SPE) is usually used for fault detection using KPCA (Nomikos and MacGregor 1995; Lahdhiri et al. 2017). However, the conventional KPCA does not provide any approach of data reconstruction in the feature space. Thus, the computation SPE index is difficult in the KPCA method. (Lee et al. 2004a, b; Choi et al. 2005) proposed a simple expression to calculate SPE in the feature space H , which is shown as follows:

$$SPE = \|\mathbf{k}(x, x) - k_{x_i}^T \hat{P} \hat{\Lambda}^{-1} \hat{P}^T k_{x_i}\| \tag{18}$$

where, $\hat{P} = [\alpha_1, \dots, \alpha_\ell]$ is the matrix of the first ℓ principal eigenvectors of $K, \hat{\Lambda} = \text{diag}[\lambda_1, \dots, \lambda_\ell]$ is the diagonal matrix of the first ℓ eigenvalues of K and $k_{x_i} = [\mathbf{k}(x_1, x_i), \dots, \mathbf{k}(x_N, x_i), \dots, \mathbf{k}(x_N, x_i)]^T; i = 1 \dots N$.

The confidence limit for SPE index can be calculated using the χ^2 -distribution and is given by:

$$SPE_{\text{lim}} \approx \delta_\alpha^2 \quad (19)$$

where δ_α^2 is the control limit expressed by:

$$\delta_\alpha^2 = g\chi_{h,\alpha}^2 \quad (20)$$

with: $g = \frac{b}{2a}$ and $h = \frac{2a^2}{b}$, where a is the estimated mean and b is the variance of the SPE.

3 The proposed reduced Rank-KPCA approach

3.1 Principle

For fault detection based on conventional KPCA approach, we would like to identify the best KPCA monitoring model. For a given test observation $x_t \in \mathbb{R}^m$, the calculation of its projections in the feature space requires computation of kernel vector k_{x_t} in respect of all N training data. Therefore, the amount of training data tends to be large, which leads to computational complexity of the KPCA model. Considering this, the choice of reduced training data can be made randomly, but this does not insure that the reduced data will represent an optimal model of the system behavior. Recently, several solutions have been developed to solve this problem and to search a reduced data set that represent adequately the system (Taouali et al. 2016; Jaffel et al. 2016a, b; Honeine 2012).

This section will present a new reduced Rank-KPCA method for monitoring nonlinear dynamic system. The key idea of this proposed method is to remove the dependencies of variables in the feature space and to retain a reduced data from the original one $X_r = [x_1, \dots, x_i, \dots, x_r]^T \in \mathbb{R}^{r \times m}$, where r is the number of retained observations. The procedure of this novel monitoring process includes two phases: offline reduced Rank-KPCA model identification and online fault detection procedure. The first phase is to identify the reduced reference model that describes the normal operating condition. Its principle is to retain the observations that generate independent linear combinations in the feature space and reveal the useful information. After that, the built model is performed on-line in order to monitor the system.

3.2 Reduced Rank-KPCA: model identification

In order to identify the reduced Rank-KPCA model, we save the most useful new observation in term of information about the system in a reduced training data matrix X_r . Let consider that the system operate in the normal condition during N_0 instants. At first time, the initial reduced data matrix is expressed as: $X_r = [x_1] \in \mathbb{R}^{1 \times m}$.

At each instant t , a new observation x_t is collected, its kernel vector k_{x_t} is calculated and the kernel matrix is updated by adding a column and a row to the previous one, as:

$$K_r^t = \begin{bmatrix} K_r^{t-1} & k_{x_t} \\ k_{x_t}^T & \mathbf{k}(x_t, x_t) \end{bmatrix} \in \mathbb{R}^{r \times r} \quad (21)$$

The rank of the updated kernel matrix is calculated, its value leading to either case: the reduced data matrix is incremented by adding the new observation or it is left unchanged.

- **Case 1:**

$$\text{rank}(K_r^t) = r \quad (22)$$

The kernel matrix has a full rank, which describes the independencies between the projection data in the feature space. In this case, the new observation is added to the reduced data matrix.

- **Case 2:**

$$\text{rank}(K_r^t) < r \quad (23)$$

The kernel matrix has not a full rank, which describes the dependencies between the projection data in the feature space. In this case, the reduced data matrix left unchanged and we return the kernel matrix to its previous state.

Once, all the observations were evaluated, we obtain the reduced data matrix $X_r \in \mathbb{R}^{r \times m}$ and we construct the reduced kernel matrix $K_r \in \mathbb{R}^{r \times r}$, such that

$$K_r = \begin{bmatrix} \mathbf{k}(x_1, x_1) & \cdots & \mathbf{k}(x_1, x_r) \\ \vdots & \ddots & \vdots \\ \mathbf{k}(x_r, x_1) & \cdots & \mathbf{k}(x_r, x_r) \end{bmatrix} \in \mathbb{R}^{r \times r} \quad (24)$$

Then, we estimate the initial reduced Rank-KPCA model (eigenvalues and eigenvectors).

4 The proposed monitoring strategies based on reduced Rank-KPCA

The monitoring strategies based on the novel reduced Rank-KPCA, are presented by two phases: offline learning and online monitoring. The off-line phase is useful for the two proposed online monitoring processes.

4.1 Off-line learning

Consider the system in normal operating condition during N_0 instants.

Step 1 Get a new observation $x_t \in \mathbb{R}^m$, scale it and calculate its kernel vector.

Step 2 Update the kernel matrix K_r .

Step 3 Test if K_r has a full rank, the new observation contain useful information. Then, go to step 4. Otherwise, cancel the updating of the kernel matrix and return to step 1.

Step 4 Update the reduced data matrix.

When, all the N_0 observations were evaluated, we obtain the reduced training data $X_r \in \mathbb{R}^{r \times m}$.

Step 5 Compute the mean and variance of the reduced training data.

Step 6 Construct the reduced kernel matrix and scale it to obtain the Gram matrix G_r .

Step 7 Solve the eigenvalue problem and determine the number ℓ of retained kernel principal components.

Step 8 Compute the monitoring index SPE and its confidence limit.

4.2 Online monitoring with fixed reduced Rank-KPCA model

Step 1 Obtain a new observation x_{t+1} and scale it with the mean and variance computed at off-line training.

Step 2 Calculate the kernel vector $k_t = \mathbf{k}(x_t, x_i) \in \mathbb{R}^{1 \times N}$ and scale it.

Step 3 Compute the fault detection index SPE for x_{t+1} , if the control limit is exceeded, a fault is declared.

Step 4 Return to step 1.

4.3 Online monitoring based on reduced Rank-KPCA and moving window

The idea of this proposed monitoring process is to improve the effectiveness of the new Reduced Rank-KPCA method for dealing with fault detection of nonlinear dynamic process. Its principle is to update the built reduced model with moving window only if a new observation contains new pertinent information about the process and satisfies the condition (22) is available. A detailed algorithm steps are shown as follows:

Step 1 Obtain a new observation x_{k+1} and scale it with the mean and variance computed at off-line training.

Step 2 Calculate the kernel vector $k_t = \mathbf{k}(x_t, x_i) \in \mathbb{R}^{1 \times N}$ and scale it.

Step 3 Compute the fault detection index SPE for x_{t+1} , if the control limit is exceeded, a fault is declared so turn to step 1, else go to step 4.

Step 4 If the condition (22) is not satisfied, it is not required to update the reduced Rank-KPCA model, so turn to step 1, else go to next step.

Step 5 Update the reduced data matrix X_r by rejecting the oldest observation and adding the new observation.

Step 6 Actualize the reduced Rank-KPCA model (update the number of PC).

Step 7 Update the control limit of the SPE statistic.

Step 4 Return to step 1.

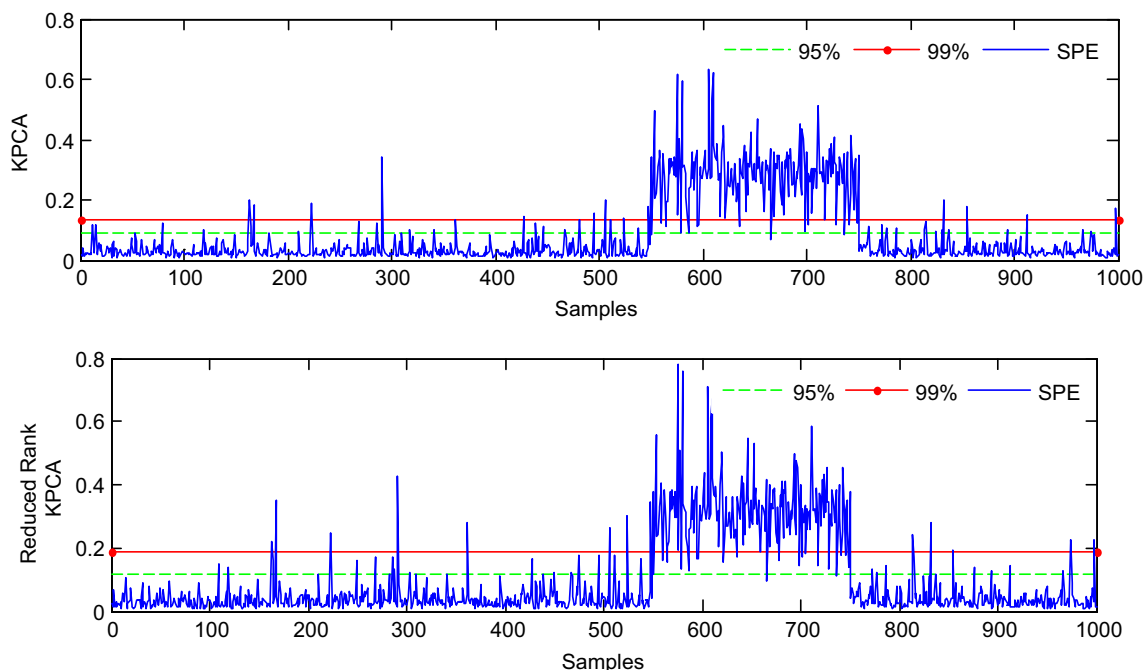


Fig. 1 Evolution of the SPE index with Fault 1 using KPCA and Reduced Rank-KPCA

5 Simulation results

The utility and validity of our proposed monitoring strategies were demonstrated using the statistic SPE index. First, we applied the proposed approach to fault detection

in a numerical example. Second, the well Known simulated CSTR process and the air quality monitoring network are used to test the effectiveness of the proposed technique and compare it to the classical monitoring techniques KPCA and MWKPCA (Liu et al. 2009).

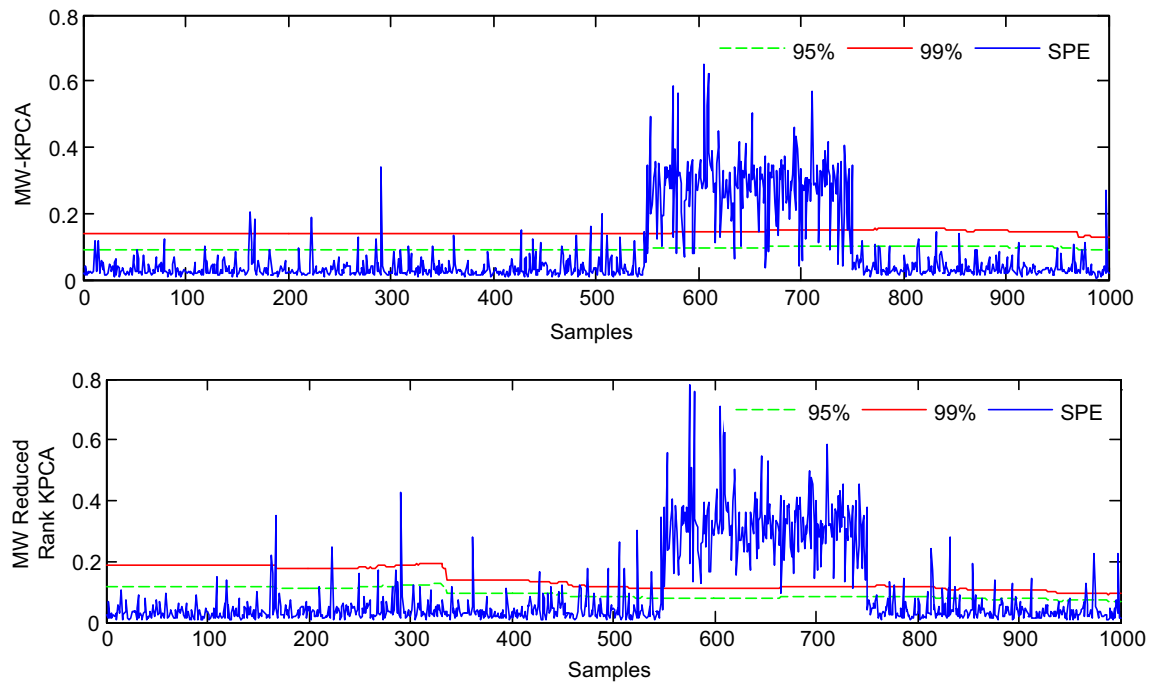


Fig. 2 Evolution of the SPE index with Fault 1 using MWKPCA and MW Reduced Rank-KPCA

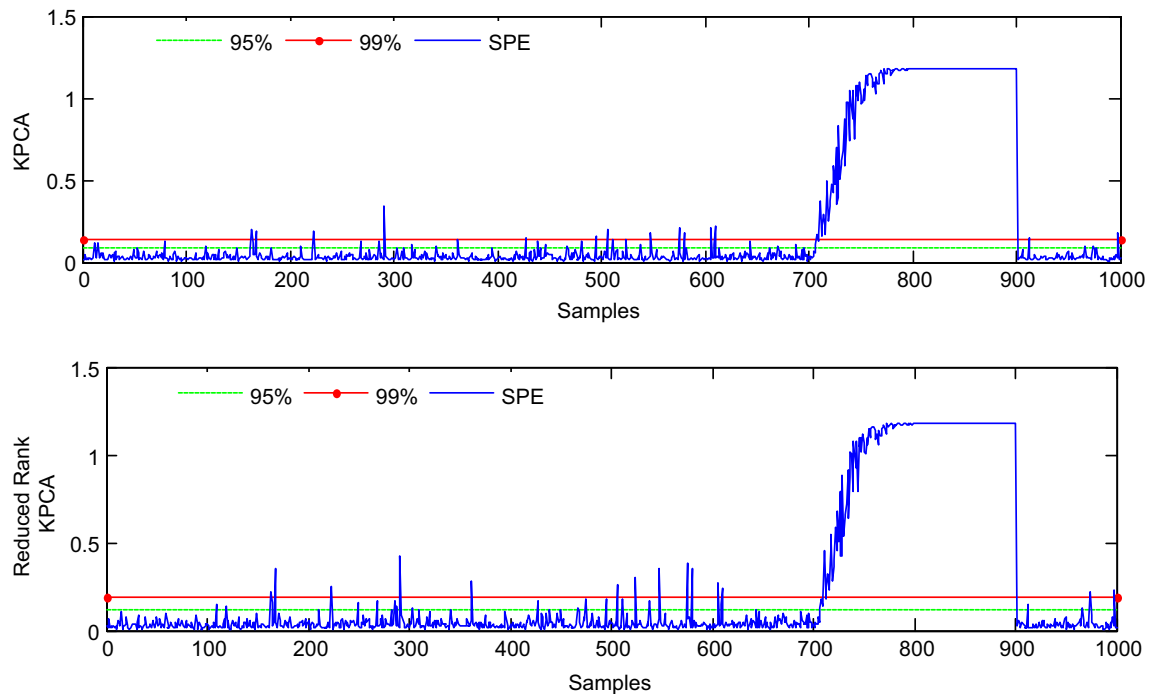


Fig. 3 Evolution of the SPE index with Fault 2 using KPCA and Reduced Rank-KPCA

5.1 Application to numerical example

5.1.1 Process description

We study the applicability of the proposed reduced Rank-KPCA approach by applying it for a simple numerical example (Kallas et al. 2014). This example contains four variables $s_i, i = 1 \dots 4$.

The following equations detail the nonlinear redundancy interrelationships between s_1, s_2, s_3 and s_4 . Where $s_1, s_2 \in [0 \dots 1]$ were uniformly distributed signals.

$$\begin{cases} s_3 - s_1 s_2 = 0 \\ s_4 - s_1^2 = 0 \end{cases} \tag{25}$$

The observation vector $x(k)$ is given by: $x(k) = [s_1(k) s_2(k) s_3(k) s_4(k)]^T$.

5.1.2 Fault detection results

Thousand samples were generated from this process. The 500 first samples were used to construct the reduced model and the 500 last samples are used to test the fault detection

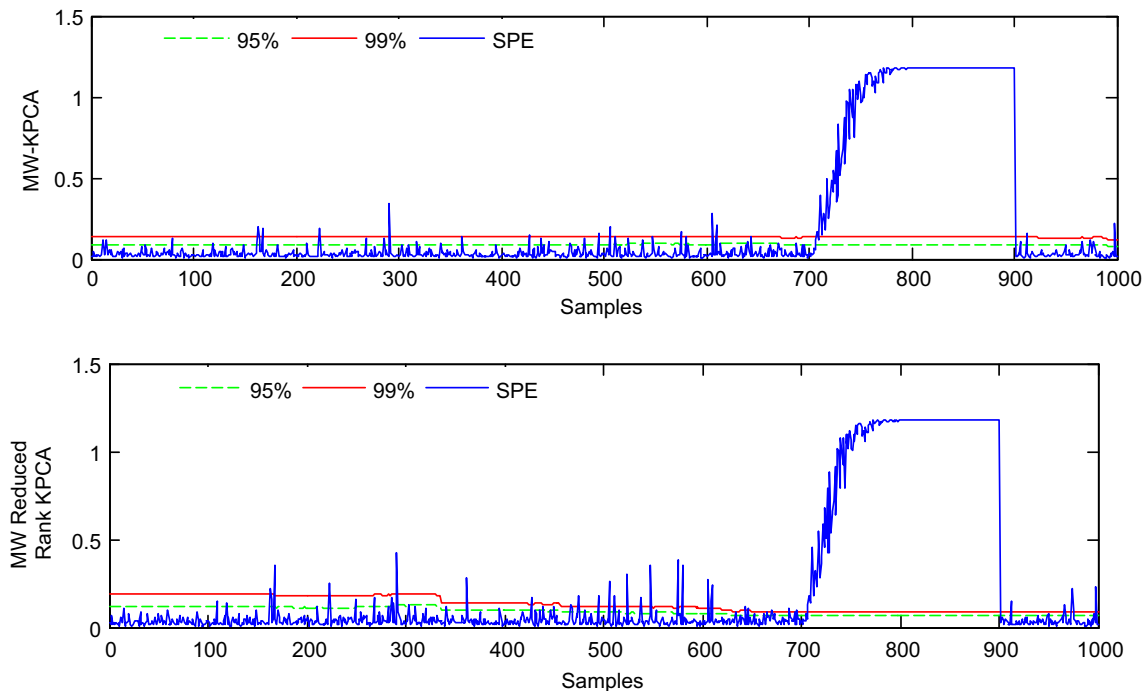


Fig. 4 Evolution of the SPE index with Fault 2 using MWKPCA and MW Reduced Rank-KPCA

Table 1 Performances of (KPCA, Reduced Rank-KPCA, MWKPCA and MW Reduced Rank-KPCA) for online monitoring

	Approach	FAR 95%	FAR 99%	GDR 95%	GDR 99%	CT (s)
Normal Operating condition	KPCA	5.3	1.7	94.7	98.3	3.58
	Reduced Rank-KPCA	4.2	1.7	95.8	98.3	1.39
	MWKPCA	5.4	1.6	92.5	96.4	81.15
	MW reduced Rank-KPCA	9.3	5	90.7	95	3.16
Fault 1	KPCA	5.38	1.62	98.01	92.04	3.74
	Reduced Rank-KPCA	4.5	1.62	98.5	90.05	1.23
	MWKPCA	1.625	2.75	92.04	86.57	81.15
	MW reduced Rank-KPCA	7.25	3.62	100	99	2.74
Fault 2	KPCA	2	0.125	96.52	95.52	4.093
	Reduced Rank-KPCA	3.5	1.5	96.52	94.53	1.328
	MWKPCA	5.25	1.63	96.52	95.52	80.62
	MW reduced Rank-KPCA	7.5	3.5	97.02	96.52	2.641

techniques. In this study, we have used the Gaussian kernel [Radial Basis Function (RBF)]. The kernel parameter σ is determined using the cross validation. The number of selected observations using the new reduced Rank-KPCA is equal to $r = 167$ from 500 observations. The confidence limits of the statistical SPE index are set to 95 and 99%, respectively. Next, the detection performance of developed technique is assessed and compared to the classical methods, through two types of faults:

- Fault 1: A step bias of s_1 by adding 30% than its range of variation. The fault is introduced between the samples 550 and 750.
- Fault 2: A ramp change of s_4 by adding $0.015 \times (k - 699)$ was introduced between the samples 700 and 900.

The evolution of the SPE index for conventional KPCA, MWKPCA and the proposed monitoring techniques with

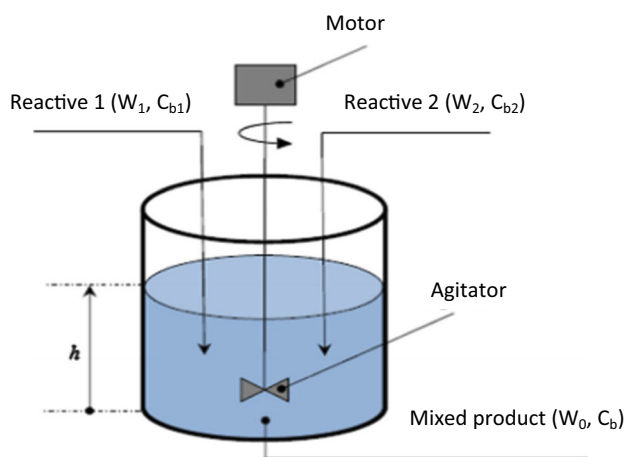


Fig. 5 Schematic diagram of the reactor

Table 2 Sensor Fault description

Case	Affected variable	Fault description	Location
Fault 1	w_2	Bias 20%	700–900
Fault 2	C_{b2}	Bias 15%	500–700
Fault 3	C_{b1}	Bias 16%	650–850

Fault 1 and Fault 2 is shown in Figs. 1, 2, 3 and 4. It is clear that the injected faults are successfully detected in both thresholds (95 and 99%). In Table 1, we summarize the detection performances in terms of False Alarm Rate (FAR), Good Detection Rate (GDR) and the average Computation Time (CT). We notice that the evaluated GDR and FAR using KPCA, MWKPCA monitoring and Reduced Rank-KPCA methods still comparable in normal operating conditions and for both Fault1 and Fault2. However, the two developed techniques based on Reduced Rank-KPCA provide a smaller CT when compared to the KPCA and MWKPCA. We can show also from Table 1, that the developed MW Reduced Rank-KPCA provides better results when compared to the developed Reduced Rank-KPCA in terms of FAR and GDR, and both of them out performances the classical monitoring methods: KPCA and MWKPCA.

5.2 Proposed techniques and its applications to CSTR process

5.2.1 Process description

The non-isothermal Continuous Stirred Tank Reactors (CSTR) is a dynamic nonlinear process, which has been mostly used for comparison of several process monitoring and fault detection strategies such as VMWKPCA (Fazai et al. 2016), SVD-RKPCA (Jaffel et al. 2016a, b) and RKPCA (Taouali et al. 2016).

The CSTR process is composed of two feed streams that are the stream of reactant 1 and 2 which are mixed to furnish a final product. The schematic diagram of the reactor is configured in Fig. 5. The dynamic behavior of the process is described by the following equations:

$$\begin{cases} \frac{dh(t)}{dt} = w_1(t) + w_2(t) - 0.2\sqrt{h(t)} \\ \frac{dC_b(t)}{dt} = (C_{b1}(t) - C_b) \frac{w_1}{h(t)} + (C_{b2}(t) - C_b) \frac{w_2}{h(t)} - \frac{k_1 C_b(t)}{(1 + k_2 C_b(t))^2} \end{cases} \quad (26)$$

where h , C_b and w_0 are respectively the height, the concentration and the feed of reacting mixture, C_{b1} and C_{b2} are the concentration of reactant 1 and reactant 2,

Table 3 Performances of (KPCA, Reduced Rank-KPCA, MWKPCA and MW Reduced Rank-KPCA) in normal operating conditions

	Approach	FAR 95%	FAR 99%	GDR 95%	GDR 99%	CT (s)
Normal Operating condition	KPCA	4.9	2.5	95.1	97.5	2.18
	Reduced Rank-KPCA	5.3	2.2	94.7	97.8	1.31
	MWKPCA	1.61	12.3	83.9	87.9	33.74
	MW Reduced Rank-KPCA	8.2	4.1	91.8	95.9	2.30

respectively, w_1 and w_2 are the flow rate of reactant 1 and reactant 2, respectively, k_1 and k_2 are the reaction rates assumed to be constants ($k_1 = k_2 = 1$).

In this study, we considered five monitored variables w_1, w_2, C_b, C_{b1} and C_{b2} . The observation vector $x(k)$ defined by:

$$x(k) = [w_1(k) \ w_2(k) \ C_b(k) \ C_{b1}(k) \ C_{b2}(k)]^T. \quad (27)$$

5.2.2 Simulation results

During a normal operation of the process, 1000 simulated observations are used. The first 300 observations were used as a training dataset in order to construct the initial reduced model, and the 700 last samples were used as a test dataset, to illustrate the performances of the new proposed method, three types of faults are considered (Table 2).

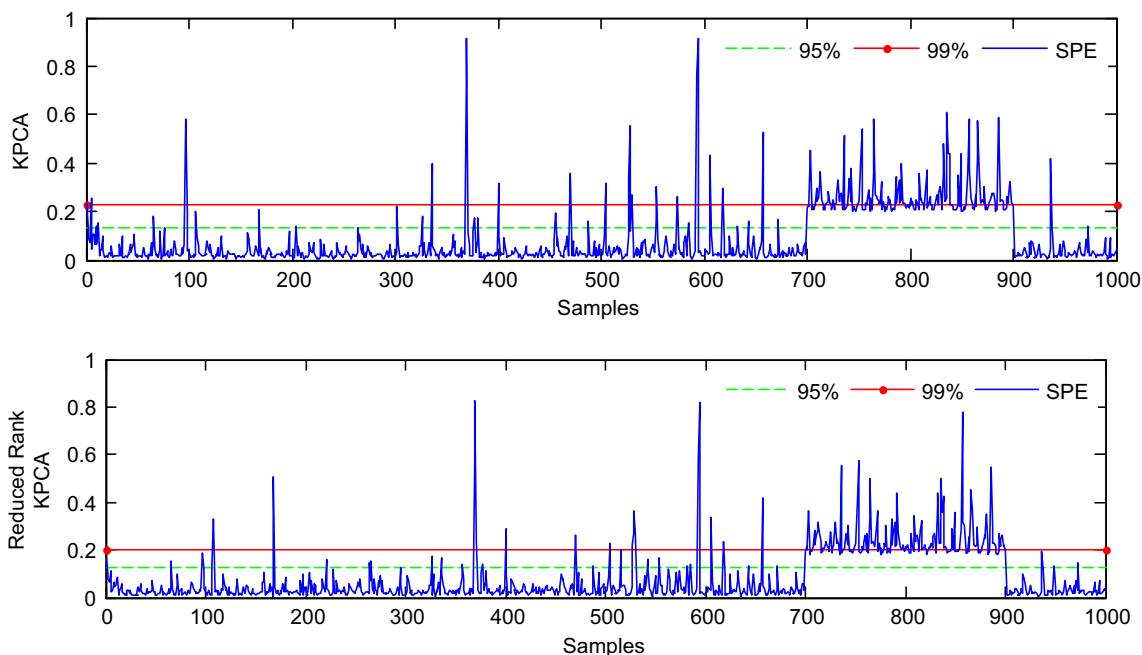


Fig. 6 Evolution of the SPE index with Fault 1 using KPCA and Reduced Rank-KPCA

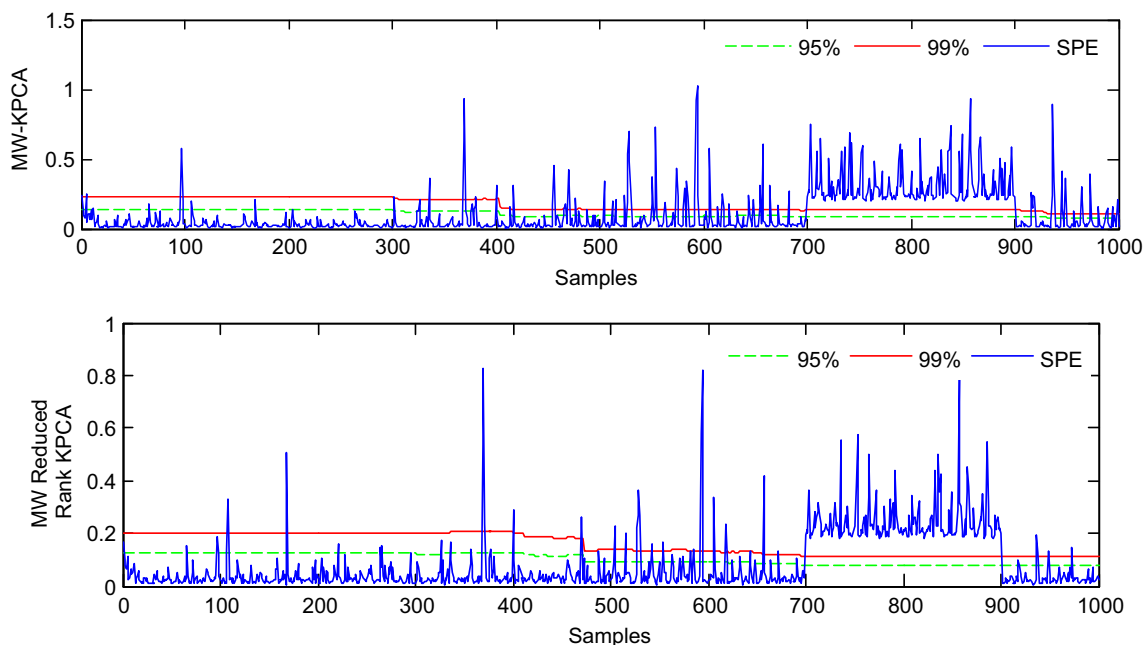


Fig. 7 Evolution of the SPE index with Fault 1 using MWKPCA and MW Reduced Rank-KPCA

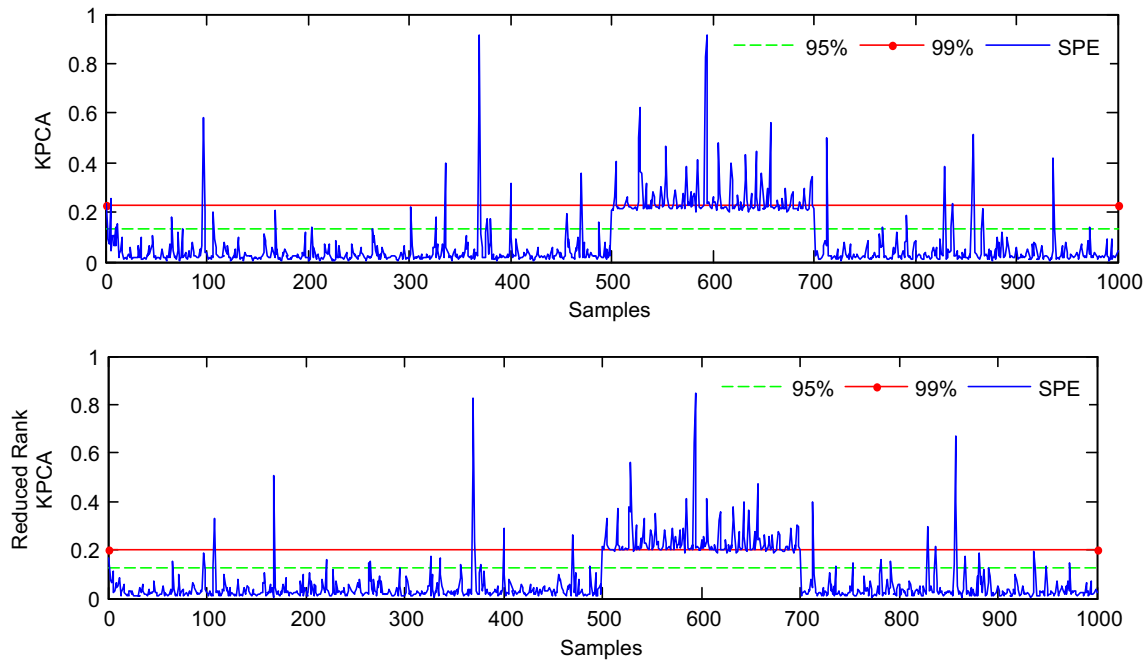


Fig. 8 Evolution of the SPE index with Fault 2 using KPCA and Reduced Rank-KPCA

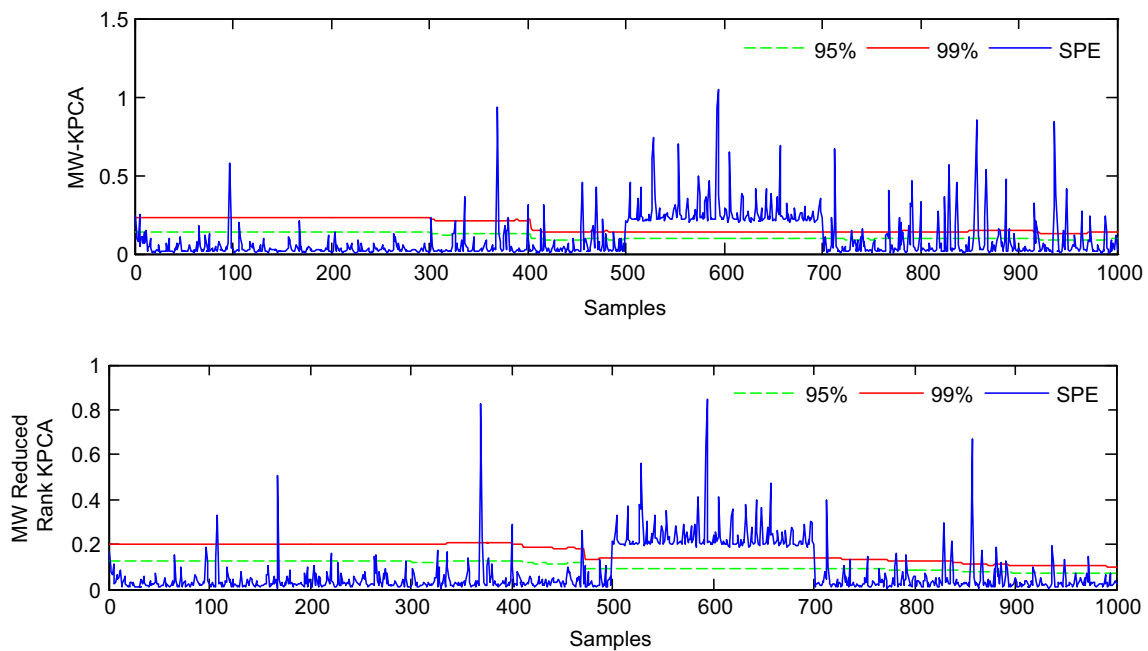


Fig. 9 Evolution of the SPE index with Fault 2 using MWKPCA and MW Reduced Rank-KPCA

Table 3 shows a comparison of the average computation time (CT), FAR and GDR for the proposed methods and conventional ones in normal operating conditions. According to Table 3, conventional KPCA and reduced Rank-KPCA have similar performances in term of FAR and GDR but we remark that the proposed Reduced Rank-KPCA has a reduced CT. It is also noted that the proposed MW Reduced Rank-KPCA has a higher GDR and a

reduced FAR at the confidence level 99% compared to MWKPCA.

Figures 6, 7, 8 and 9 show the monitoring results of conventional KPCA, Reduced Rank-KPCA, MWKPCA and MW Reduced Rank-KPCA using the SPE index in the cases of Fault 1 and Fault 2. We can show that, the faults are detected at time in the both thresholds (95 and 99%). The performances of the four monitoring methods are

Table 4 Performances of (KPCA, reduced Rank-KPCA, MWKPCA and MW reduced Rank-KPCA) for online monitoring

	Approach	FAR 95%	FAR 99%	GDR 95%	GDR 99%	CT (s)
Fault 1	KPCA	4.87	2.5	100	60	2.59
	Reduced Rank-KPCA	5	2.13	100	69.2	1.14
	MWKPCA	12.5	9.12	100	100	24.64
	MW Reduced Rank-KPCA	7	3.25	100	100	1.91
Fault 2	KPCA	4.13	1.75	100	35.82	2.53
	Reduced Rank-KPCA	4.5	1.5	100	73.13	1.19
	MWKPCA	11.88	7.13	100	100	25.77
	Reduced Rank-KPCA	6.13	3	100	100	2.19
Fault 3	KPCA	5.125	2.6	95.5	5.97	2.25
	Reduced Rank-KPCA	5.5	2.25	96.51	10	1.07
	MWKPCA	13.38	9	99.5	93.03	24.73
	MW Reduced Rank-KPCA	7.37	3.25	100	85.07	2.11

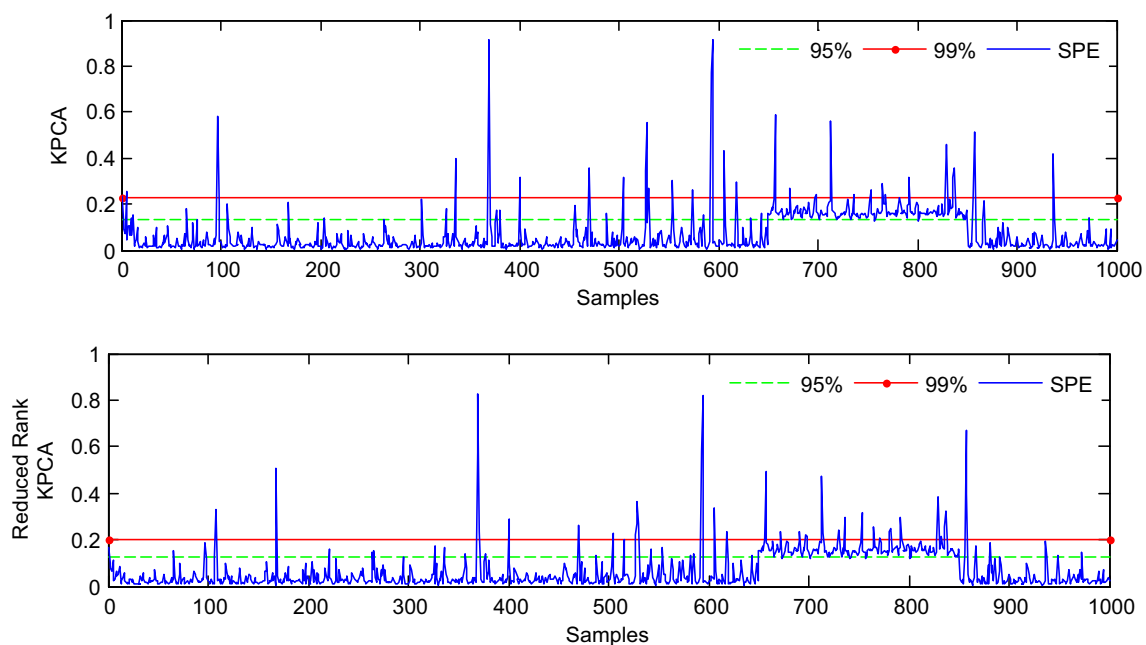


Fig. 10 Evolution of the SPE index with Fault 3 using KPCA and Reduced Rank-KPCA

summarized in Table 4. We note that the proposed Reduced Rank-KPCA and MW Reduced Rank-KPCA at the confidence level 99% have a higher GDR and less CT compared to conventional KPCA and MWKPCA.

Figure 10 shows the monitoring results of conventional KPCA and Reduced Rank-KPCA using the SPE index in the case of Fault 3. We notice also that the conventional KPCA and the Reduced Rank-KPCA are not able to detect the fault in the threshold 99%. Figure 11 show the ability of the MWKPCA and MW Reduced Rank-KPCA to detect this fault. MW Reduced Rank-KPCA has a less computation time and a reduced false alarm rate than MWKPCA (see Table 4, Fig. 7).

5.3 Proposed techniques and its application to an air quality monitoring network

5.3.1 Process description

The air quality monitoring network AIRLOR, operating in Lorraine, France. The AIRLOR consists of twenty stations placed in several sites: rural, peri-rural and urban. Each station served to the acquisition of some pollutants in the air like nitrogen oxides (NO and NO₂), ozone (O₃), carbon monoxide (CO) and Sulphur dioxide (SO₂). In this study, six stations are served to the registering of the additional metrological parameters. The main objective is to detect defects of the sensors, which measure ozone concentration

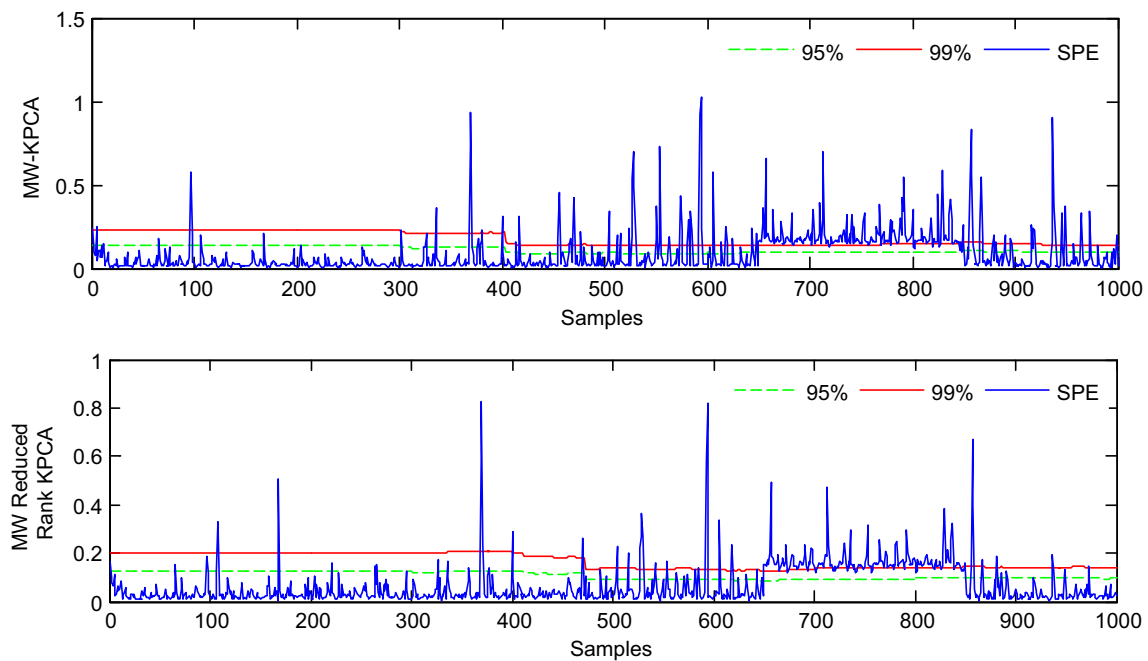


Fig. 11 Evolution of the SPE index with Fault 3 using MWKPCA and MW Reduced Rank-KPCA

O₃ and nitrogen oxides NO and NO₂. The phenomenon of photochemical pollution presents a dynamic nonlinear behavior (Harkat et al. 2006).

The observation vector $x(k)$ contains 18 monitored variables v_1 to v_{18} , which corresponding to ozone concentration, nitrogen oxide and nitrogen dioxide, respectively, of each station.

5.3.2 Simulation results

Only 1000 simulated observations are used. The first 500 observations were used in the training phase to elaborate the reduced Rank-KPCA model and the last observations 501–1000 were used to test the proposed fault detection methods. In this part, we simulated two faults, which are presented in Table 5.

$$x(k) = \begin{bmatrix} \underbrace{v_1(k) \ v_2(k) \ v_3(k)}_{O_3(k)NO(k)NO_2(k):station1} & \cdots & \underbrace{v_{10}(k) \ v_{11}(k) \ v_{12}(k)}_{O_3(k)NO(k)NO_2(k):station4} & \cdots & \underbrace{v_{16}(k) \ v_{17}(k) \ v_{18}(k)}_{O_3(k)NO(k)NO_2(k):station6} \end{bmatrix}^T \tag{28}$$

Table 5 Sensor Fault description

Case	Affected variable	Fault description	Fault location
Fault 1	$v_{18}(NO_2)$	Bias 65%	750–950
Fault 2	$v_2(NO)$	Bias 35%	520–750

Table 6 summarizes the monitoring performances of the four methods when the process is operating in normal condition. We note that the proposed monitoring methods have a less computation time compared to the others.

Figures 12 and 13 show the monitoring results of the four methods using the SPE index in the case of bias fault

Table 6 Performances of (KPCA, Reduced Rank-KPCA, MWKPCA and MWReduced Rank-KPCA) in normal operating condition

	Approach	FAR 95%	FAR 99%	GDR 95%	GDR 99%	CT (s)
Normal Operating condition	KPCA	4.4	1.7	956	983	3.412
	Reduced Rank-KPCA	9.3	4.6	90.7	95.4	1.465
	MWKPCA	7.5	2.4	92.5	97.6	22.92
	MWReduced Rank-KPCA	16.5	9.3	83.5	90.7	2.887

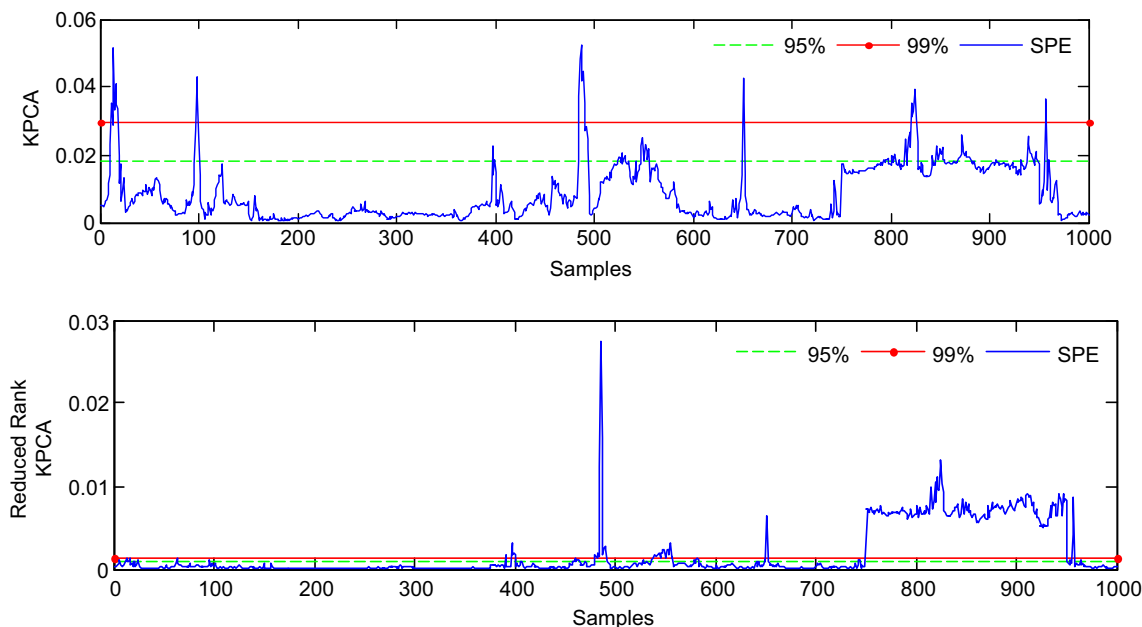


Fig. 12 Evolution of the SPE index with Fault 1 using KPCA and Reduced Rank-KPCA

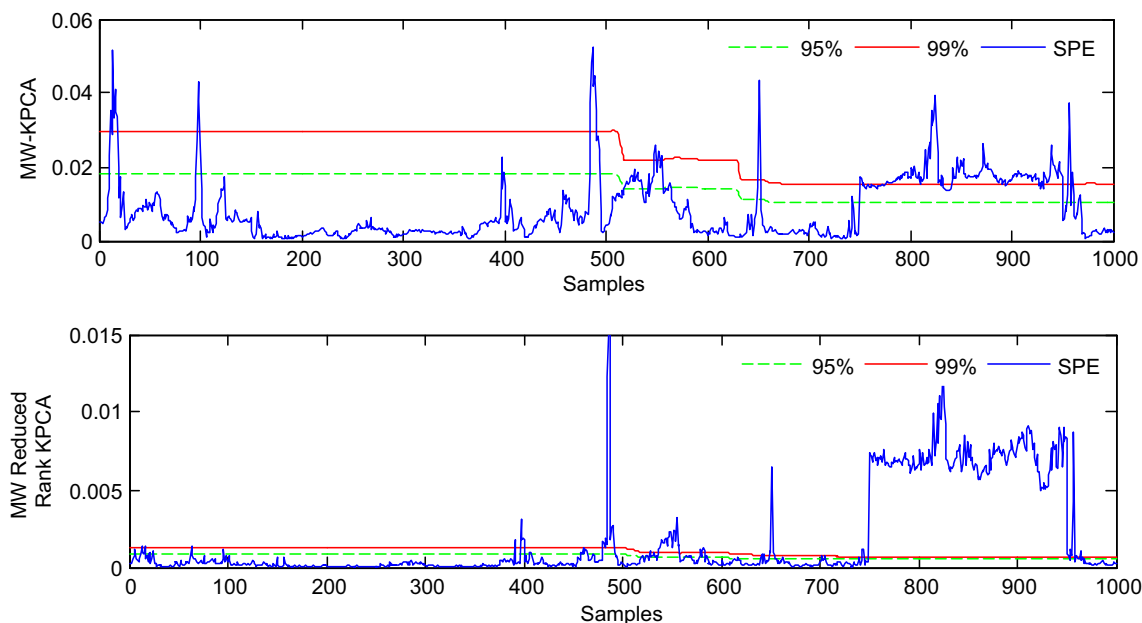


Fig. 13 Evolution of the SPE index with Fault 1 using MWKPCA and MW Reduced Rank-KPCA

at the measured concentration of NO_2 . The fault is detected at time in the both thresholds (95 and 99%), using Reduced Rank-KPCA, MWKPCA and MW Reduced Rank-KPCA. From Table 7, we show that MW Reduced Rank-KPCA has a highly GDR, reduced FAR and significantly reduced CT compared to MWKPCA.

The second fault is a bias, it was introduced into the measured concentration of NO . Figure 14 shows the detection index SPE for this fault. It is impossible to detect this fault using conventional KPCA. However, the

proposed Reduced Rank-KPCA method provides a good detection of the simulated fault at the time in the threshold 95%. As it can be seen in Table 7, the evaluated GDR at the threshold 99% is less than 50%. Since the reduced Rank-KPCA is a monitoring method based on the use of a fixed model so it is powerless to monitoring a dynamic non-linear process. The simulation results of the two adaptive KPCA methods were shown in Fig. 15. The injected fault is clearly detected in both thresholds (95 and

Table 7 Performances of (KPCA, Reduced Rank-KPCA, MWKPCA and MW reduced Rank-KPCA) for online monitoring

	Approach	FAR 95%	FAR 99%	GDR 95%	GDR 99%	CT (s)
Fault 1	KPCA	5.5	2.13	28.36	2.985	3.429
	Reduced Rank-KPCA	10.37	5.5	100	100	1.545
	MWKPCA	36.82	11.94	100	86.07	13.94
	MW Reduced Rank-KPCA	15	8.25	100	100	2.048
Fault 2	KPCA	3.89	2.08	6.95	0.43	3.982
	Reduced Rank-KPCA	7.27	3.12	93.94	44.16	1.879
	MWKPCA	4.80	2.34	20.78	5.194	24.36
	MW Reduced Rank-KPCA	13.51	7.53	100	97.83	1.962

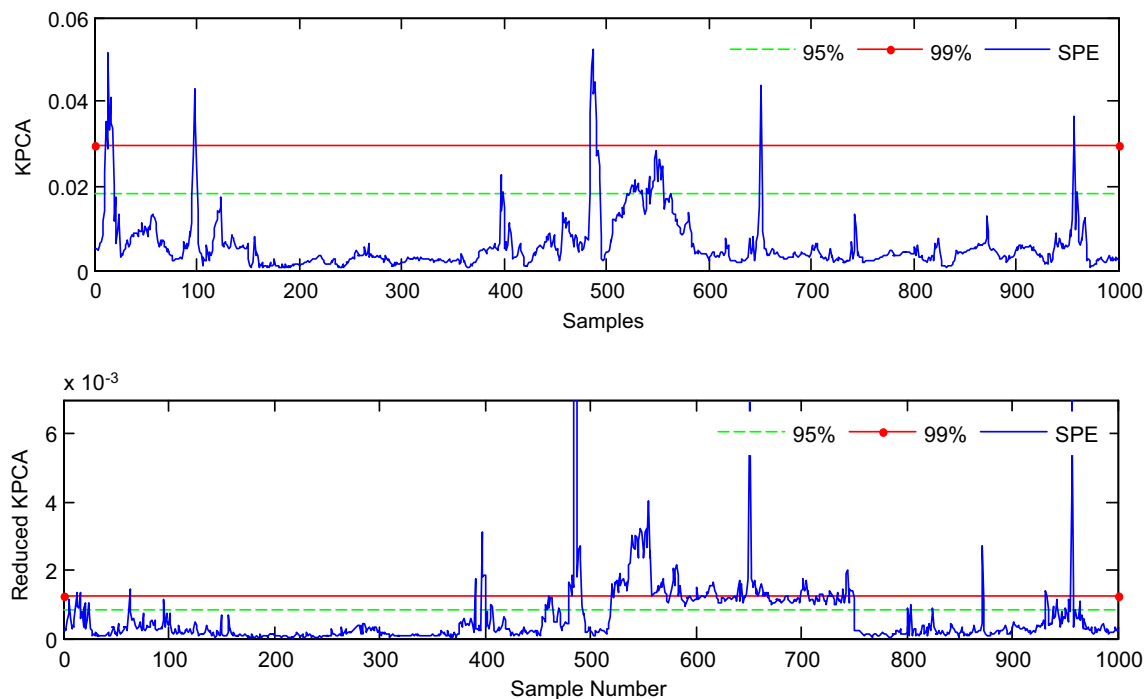


Fig. 14 Evolution of the SPE index with Fault 2 using KPCA and Reduced Rank-KPCA

99%) using the proposed MW Reduced Rank-KPCA. We notice that MWKPCA is unable to detect this fault.

6 Conclusion

In this paper, a new reduced Rank-KPCA method has been proposed. Firstly, we derive an algorithm for choosing the pertinent observations that describe the independence between variables in feature space and after we build the reduced Rank-KPCA model. Then, this proposed method was developed into two monitoring strategies for non-linear processes. The first one was called Reduced Rank-KPCA based on a fixed reduced model and is applied to fault detection in a numerical example, a CSTR benchmark and an air quality-monitoring network. This proposed monitoring method gives more sophisticated results and

better monitoring performance compared with the conventional KPCA.

The computation time is the most important criterion when using Reduced Rank-KPCA model for process monitoring. In offline training, the CT of reduced model construction seems large in comparison to KPCA model construction. However, once the reduced model and the confidence limits of normal operating data have been calculated, the online monitoring procedure of Reduced Rank-KPCA has provided the shortest CT.

The training data set of the numerical example is assumed invariant in time. Thus, the efficiency of the proposed monitoring with fixed reduced Rank-KPCA model has been proved. Moreover, most of real industrial processes are time varying, hence the reduced Rank-model should be updated. Seeing that, a second monitoring strategy has been proposed to improve the Reduced Rank-

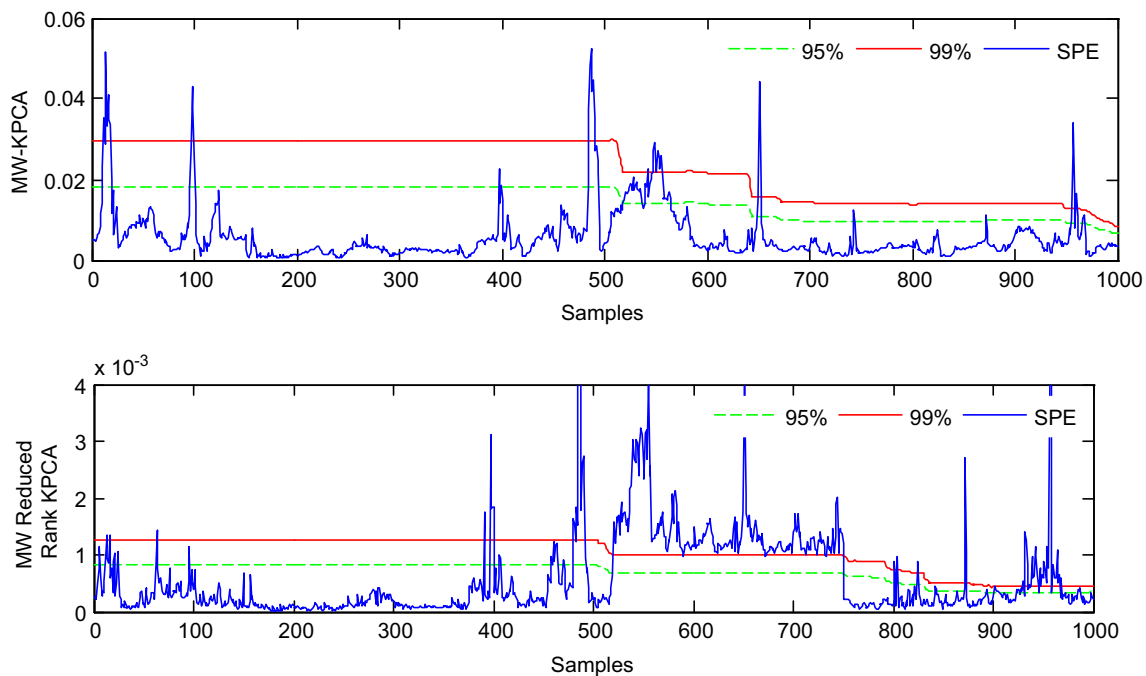


Fig. 15 Evolution of the SPE index with Fault 2 using MWKPCA and MW Reduced Rank-KPCA

KPCA called MW Reduced Rank-KPCA, which aims to update the reduced model using moving window. Applications to the CSTR and Air Quality Monitoring Network have demonstrated that the proposed adaptive monitoring method shows a higher performance for most faults than the MWKPCA method.

References

- Aizerman M, Braverman E, Rozonoer L (1964) Theoretical foundations of the potential function method in pattern recognition learning. *Autom Remote Control* 25:821–837
- Amaro LIMA, Heiga ZEN, Nankaku Y, Tokuda K, Kitamura T, Resende FG (2005) Applying sparse KPCA for feature extraction in speech recognition. *IEICE Trans Inf Syst* 88(3):401–409
- Aronszajn N (1950) Theory of reproducing kernels. *Trans Am Math Soc* 68(3):337–404
- Baklouti R, Mansouri M, Nounou M, Nounou H, Hamida AB (2016) Iterated robust kernel fuzzy principal component analysis and application to fault detection. *J Comput Sci* 15:34–49
- Botre C, Mansouri M, Nounou M, Nounou H, Karim MN (2016) Kernel pls-based glrt method for fault detection of chemical processes. *J Loss Prev Process Ind* 43:212–224
- Chetouani Y (2008) A neural network approach for the real-time detection of faults. *Stoch Environ Res Risk Assess* 22(3):339–349
- Cho JH, Lee JM, Choi SW, Lee D, Lee IB (2005) Fault identification for process monitoring using kernel principal component analysis. *Chem Eng Sci* 60(1):279–288
- Choi SW, Lee C, Lee JM, Park JH, Lee IB (2005) Fault detection and identification of nonlinear processes based on kernel PCA. *Chemom Intell Lab Syst* 75(1):55–67
- Dong D, McAvoy TJ (1996) Nonlinear principal component analysis based on principal curves and neural networks. *Comput Chem Eng* 20(1):65–78
- Fazai R, Taouali O, Harkat MF, Bouguila N (2016) A new fault detection method for nonlinear process monitoring. *Int J Adv Manuf Technol* 87(9–12):3425–3436
- Harkat MF, Mourot G, Ragot J (2006) An improved PCA scheme for sensor FDI: application to an air quality monitoring network. *J Process Control* 16(6):625–634
- Harkat MF, Tharrault Y, Mourot G, Ragot J (2010) Multiple sensor fault detection and isolation of an air quality monitoring network using RBF-NLPCA model. *Int J Adapt Innov Syst* 1(3–4):267–284
- Honeine P (2012) Online kernel principal component analysis: a reduced-order model. *IEEE Trans Pattern Anal Mach Intell* 34(9):1814–1826
- Jaffel I, Taouali O, Elaissi I, Messaoud H (2014) A new online fault detection method based on PCA technique. *IMA J Math Control Inf* 31(4):487–499
- Jaffel I, Taouali O, Harkat MF, Messaoud H (2016a) Kernel principal component analysis with reduced complexity for nonlinear dynamic process monitoring. *Int J Adv Manuf* 88:3265–3279
- Jaffel I, Taouali O, Harkat MF, Messaoud H (2016b) Moving window KPCA with reduced complexity for nonlinear dynamic process monitoring. *ISA Trans* 64:184–192
- Kallas M, Mourot G, Maquin D, Ragot J (2014) Diagnosis of nonlinear systems using kernel principal component analysis. In: *Journal of physics: conference series* (Vol. 570, No. 7, p. 072004). IOP Publishing
- Kazor K, Holloway RW, Cath TY, Hering AS (2016) Comparison of linear and nonlinear dimension reduction techniques for automated process monitoring of a decentralized wastewater treatment facility. *Stoch Environ Res Risk Assess* 30(5):1527–1544
- Kim KI, Jung K, Kim HJ (2002) Face recognition using kernel principal component analysis. *IEEE Signal Process Lett* 9(2):40–42

- Lahdhiri H, Taouali O, Elaissi I, Jaffel I, Harakat MF, Messaoud H (2017) A new fault detection index based on Mahalanobis distance and kernel method. *Int J Adv Manuf Technol* 91:2799–2809
- Lee JM, Yoo C, Lee IB (2004a) Statistical process monitoring with independent component analysis. *J Process Control* 14(5):467–485
- Lee JM, Yoo C, Choi SW, Vanrolleghem PA, Lee IB (2004b) Nonlinear process monitoring using kernel principal component analysis. *Chem Eng Sci* 59(1):223–234
- Li H, Zhang D (2013) Stochastic representation and dimension reduction for non-Gaussian random fields: review and reflection. *Stoch Environ Res Risk Assess* 27(7):1621–1635
- Li G, Qin SJ, Zhou D (2010) Geometric properties of partial least squares for process monitoring. *Automatica* 46(1):204–210
- Liu X, Kruger U, Littler T, Xie L, Wang S (2009) Moving window kernel PCA for adaptive monitoring of nonlinear processes. *Chemom Intell Lab Syst* 96(2):132–143
- Mansouri M, Nounou M, Nounou H, Karim N (2016) Kernel PCA-based GLRT for nonlinear fault detection of chemical processes. *J Loss Prev Process Ind* 40:334–347
- Mercer J (1909) Functions of positive and negative type and their connection with the theory of integral equations. *Philos Trans R Soc Lond Ser A Contain Pap Math Phys Character* 209:415–446
- Mika S, Schölkopf B, Smola AJ, Müller KR, Scholz M, Rätsch G (1998) Kernel PCA and De-noising in feature spaces. In: *NIPS*, vol. 11, pp 536–542
- Nomikos P, MacGregor JF (1995) Multivariate SPC charts for monitoring batch processes. *Technometrics* 37(1):41–59
- Schölkopf B, Smola A, Müller KR (1998) Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput* 10(5):1299–1319
- Sheriff MZ, Mansouri M, Karim MN, Nounou H, Nounou M (2017) Fault detection using multiscale PCA-based moving window GLRT. *J Process Control* 54:47–64
- Taouali O, Elaissi I, Messaoud H (2015) Dimensionality reduction of RKHS model parameters. *ISA Trans* 57:205–210
- Taouali O, Jaffel I, Lahdhiri H, Harkat MF, Messaoud H (2016) New fault detection method based on reduced kernel principal component analysis (RKPCA). *Int J Adv Manuf Technol* 85(5–8):1547–1552
- Tharrault Y, Mourot G, Ragot J, Maquin D (2008) Fault detection and isolation with robust principal component analysis. *Int J Appl Math Comput Sci* 18(4):429–442
- Vapnik V (1999) An overview of statistical learning theory. *IEEE Trans Neural Netw* 10(5):988–999
- Zhang Y, Li S, Teng Y (2012) Dynamic processes monitoring using recursive kernel principal component analysis. *Chem Eng Sci* 72:78–86
- Zhao C, Wang F, Mao Z, Lu N, Jia M (2008) Adaptive monitoring based on independent component analysis for multiphase batch processes with limited modeling data. *Ind Eng Chem Res* 47(9):3104–3113