*IOSR Journal of Computer Engineering (IOSR-JCE)*
*e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 23, Issue 6, Ser. I (Nov. –Dec. 2021), PP 13-26*
*www.iosrjournals.org*

# Process monitoring based a Novel Reduced Rank KPLS approach

Okba Taouali
[1]*National School of Engineering Monastir,Tunisia*
[2]*Department of Computer Engineering, Faculty of Computers and Information Technology, University of Tabuk, Tabuk, Saudi Arabia*

***Abstract:*** *This article proposes a novel reduced fault detection method based on the data-driven approaches for process monitoring. An overview on machine learning methods took place to illustrate the efficiencies for Fault Detection (FD). The concept, the original ideas, the reduced steps, data storage and implementation conditions are discussed. The performance from the application viewpoint of the proposed Reduced Rank KPLS technique and all discussed methods are validated using an industrial benchmark*

***Key Word****: Machine learning, KPCA, RKPCA, Reduced Rank-KPCA, KPLS, RKPLS, Reduced Rank-KPLS, Fault detection.*

---

Date of Submission: 04-11-2021 Date of Acceptance: 18-11-2021

---

## I. Introduction

The industrial world in the decades has had very an important increasingly emergence of Fault Detection (FD) and fault diagnosis methods. For the improvement of safety and reliability, it is essential to establish efficient system monitoring techniques. Multivariate statistical process control (MSPC) is a collection of such techniques that are usually based on the analysis of measurements system. In the field of monitoring process, PCA approach is one of the important utilized data driven-based method [1], [2]. This method is applied generally to identify the dependency structure between the data measurement for transform the input space into a reduced dimensional space while selecting the maximum variance of the input data measurement. For linear process monitoring, PCA is an efficient MSPC technique for FD and fault isolation. In addition to, the Partial Least Squares (PLS) as an input output technique has illustrate a very important performance for FD [3]. This method can construct a linear mathematical model with linear Latent Variables (LVs) using the relations between inputs/outputs sets: The PLS has been extensively used in diagnosis and supervising for linear process. The majority of industrial processes data have nonlinear relationships. An extended version of PCA, developed in recent years for tackling non-linear problem, named Kernel PCA (KPCA) has been used [4]. The Kernel PLS (KPLS), an extended version of PLS, has been developed for tackling non-linear problem [5]. The main idea of KPCA approach is to first map the input space into a feature space, via nonlinear mapping, and then to compute Principal Components (PCs) in that feature space. The KPLS method can be used to test the data in the input space that are transformed using a non-linearly function into a space entitled feature space with increasing dimensionality, using nonlinear function. Some problems are presented using kernel methods for industrial process monitoring. Since the standard kernel methods perform a projection using an implicit function to an input space, then it has obstacles in distinguishing quality-related and quality unrelated faults. The number of the chosen variables determined by kernel methods can be larger than the linear versions. As a consequence, the computation time and the memory size for the storage may increase due to the number of selected measurement. For this objective, all the debated data-driven methods presents a solution to solve the FD problems and make easy this task. The contribution of this paper is to present an overview of an extension of kernel methods using a reduced observations and to provide a reference for further MSPC methods for monitoring industrial processes. All the debated approach will be tested to an industrial system entitled Tennessee Eastman (TEP) using the Squared Prediction Error (SPE) index to illustrate their efficiencies.

This article is structured as follows: in Section 2, a related works is presented. A kernel method is given in section 3. The discussed methods formulation and Algorithms are given in section 4. The proposed RR-KPLS is detailed in section 5. The fault detection method using SPE index are developed in section 6. Section 7 presents simulation results that illustrate efficiency of all discussed methods using TEP. The conclusion concludes this paper.

## II. Related works

Diagnostic methods are hot topics especially for the modern technology. However, the FD is important and essential as regards finding any sensor fault that might occur in the system. In literature, several methods have been developed for years [6], [1]. The methods based on data driven principle are becoming more applied in the process industry. The recent technology delivers an important quantities of variables, which are stored and measured using control processes. Among the FD techniques, the historical data methods can describe the various modes of operation of the system. This historical data can be divided into two significant methods: the neural networks [7] and the pattern recognition method [8]. These methods present many advantages. Firstly, a low sensitivity to observations

noise is presented. Then their ability to store information in a compact manner is given [9], [10]. Yet, these methods require in this case a heavy learning. Nonetheless, the Multivariable Statistical Process Control (MSPC) is a set of many techniques, which are based on the analysis of measurements system. This method is based on statistical projection methods like essentially: PCA, PLS and Independent Component Analysis (ICA). The ICA method has been suggested in [11], which the goal of this method is to break the data into linear combinations of statistically independent components. The PCA method developed in [12], consists to project the correlated variables across the new variables that are not correlated. However, the PLS method focus to extract data from system to model the relationship between them [13]. On the one hand, the PCA method consists to capture variations affecting the input data with a variance descending order, on the other hand the PLS method finds an optimum pair of latent variables in the input data space related to the output ones. In literature, the PCA and PLS methods are among the most widely used MSPC techniques in modeling, monitoring and diagnosis. These two methods have shown a good performance as a data driven method. However, complex industrial processes are in nature nonlinear. In this context, several extended techniques of the nonlinear PCA and PLS have been proposed in the literature. The kernel methods such as KPCA, in [14], and KPLS, in [15]- [16], have been developed and have emerged to mitigate the traditional disadvantages of PCA, and PLS respectively methods in process monitoring. In [17], the kernel methods can also be used to handle the original input data that are non-linearly transformed into a feature space. The authors in [18] presented the fault detection phase based on the KPCA which can compute principal components (PCs) in space of high dimension by using of nonlinear kernel functions and integral operator. In [19], the authors presented the fault detection phase based on the KPLS which can find the Latent Variables (LVs) that present a nonlinear correlation with the response variables and improve model performance. The principal advantages of KPLS method is that it used in the same time the input and the output set Some problems can be present for industrial process monitoring based on the kernel techniques. As a consequence, the computation time may increase due to the availability of samples, for the storage of kernel matrix during the identification step. Several extended kernel methods, for process monitoring, have been also proposed in the literature. In this paper, a general review of statistical learning techniques in Statistical Process Control, mentioning only a few kernel methods took place. To improve the detection phase, the authors in [20] proposed an optimized parameters of kernel function for process monitoring. Indeed, the kernel parameters depend on the kind of kernel function. Then the different kinds of kernel function have different numbers, characteristics and scope of kernel parameters. Different approaches using kernel methods for monitoring have been developed. Among the developed methods, Taouali et al. [21] were interested in approaches based on Reduced KPCA (RKPCA). This proposed approach deals the problem of need storage and computation time. The suggested RKPCA technique [21] aims to approximate the selected principal components given by the classical KPCA technique using an observation set which have the largest variances with the selected principal components. In fact, the RKPCA is to determine only the observations data that approximate correctly the selected principal components. In [22], Lahdhiri et al. proposed a new methodology based on reduced Rank-KPCA (RRKPCA) method for monitoring nonlinear system. The main principle of the RRKPCA method is to retain the observations that generate independent linear combinations in the feature space and reveal the useful information. On the other side, the authors in [23] suggested a Reduced KPLS (RKPLS) method. The proposed RKPLS method consists to select only the data set of observations that approximate the choosing very important components to construct a kernel matrix with reduced size. The RKPLS method is mainly based on a reduced Gram matrix, in this case the training time decreases rapidly with a reduced number of observations. In this paper, we present the review of Statistical method that contain a kernel method model for monitoring system.

## III. Kernel Methods

Systems have really a nonlinear structure. Because of the limitation of the standard statistics methods for the nonlinear system, several methods have been developed. According to the trend and popular methods, the kernel technique has received a lot of attention. The kernel methods are characterized by the Gram matrix

[24]. Furthermore, the main idea of the kernel technical is to transform $x_i$ i=1,...,N the input process variable data into a special feature space H using a nonlinear function $\phi$ , as indicated in Eq. (1):

$$\phi : x_i \rightarrow H \qquad (1)$$
$$x_i \rightarrow \phi(x_i) \in H$$

Let us consider X the training observations matrix equal to zero mean and variance is equal to unit. The X is an observation vector, N is observation number of measurement and M is systems variables. Then the kernel model is determined by the eigenvalue decomposition of the covariance matrix of data $C_\phi$ in the feature space H.

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

In this case, the $\Psi = [\phi(x_1) \ldots \phi(x_i) \ldots \phi(x_N)]^T \in \mathbb{R}^{N \times h}$ can be used as matrix contains data in the feature space H. The $Co_\phi$ matrix can be written as:

$$Co_\phi = \frac{1}{N}\Psi^T \Psi \qquad (3)$$

The PCs of the mapped observation data are determined by solving the eigenvalue problem decomposition of $C_\phi$ , as indicated by Eq.(4).

$$\lambda_j \gamma_j = Co_\phi \gamma_j \qquad j = 1, \ldots, h \qquad (4)$$

where $\lambda_j$ is the associated $j^{th}$ eigenvalue and $\gamma_j$ is the eigenvector of $Co_\phi$ that corresponding then to non-zero eigenvalue

All solution using $\gamma_j$ with $\lambda_j \neq 0$ lies in the span of $[\Phi(x_1) \ldots \Phi(x_i) \ldots \Phi(x_N)]^T$ can be determined an existing set $\alpha_{i,j}$ ; $i = 1 \ldots N$ There exist parameters can be expressed by:

$$\gamma_j = \sum_{i=1}^{N} \alpha_{i,j} \phi_i \qquad (5)$$

The mapping function $\phi$ in practice is not known. However, the covariance matrix of transformed data isn't defined and determined implicitly. As a consequence, the inner product presented by Eq. (2) can be used the Mercer's theorem [25] to determine a kernel function $k(.,.)$ as follows:

$$\left\langle \phi(t), \phi(t') \right\rangle_H = k(t, t') \qquad \forall\, t, t' \in \mathbb{R}^m \qquad (6)$$

However, The Kernel matrix $K_g \in \mathbb{R}^{N \times N}$ linked to a kernel equation k is presented by the Eq.(7).

$$K_g = \Psi\Psi^T = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_N) \\ \vdots & \ddots & \vdots \\ k(x_N, x_1) & \cdots & k(x_N, x_N) \end{bmatrix} \in \mathbf{R}^{N \times N} \qquad (7)$$

The eigenvalue problem decomposition $Co_\phi$ can be reduced using a kernel matrix Kg principle. Hence, eigenvalue decomposition of the kernel matrix Kg can be presented as following:

$$N\,\Omega\Theta = K\,\Theta \tag{8}$$

where $\Omega = diag\,(\lambda_1, \lambda_2, \lambda_3 \ldots \lambda_j \ldots \lambda_N)$ is eigenvalues $\lambda_j$ diagonal matrix arranged in descending order. The $\Theta = \left[\alpha_1, \alpha_2, \alpha_3 \ldots \alpha_j \ldots \alpha_N\right]$ is the matrix of their corresponding eigenvectors.

Many kernel functions are usually used. The following different kernel functions are:

Polynomial kernel : $k\,(f\,,e) = < f\,,e\,>^p$

Sigmoid kernel : $k\,(f\,,e) = \tanh(\beta_0 \left\langle f\,,e\right\rangle + \delta)$

Radial basis kernel : $k\,(f\,,e) = \exp(-\dfrac{\left\|f\,-e\right\|^2}{2\sigma^2})$

where $p$, $\beta_0$ and $\delta$ are choosing using an optimization technique.

Generally, the normalization and the centralization of the kernel measurement matrix Kr is determined using Eq. (9), where we are substituting Kr by the Gram matrix (Gg).

$$Gg = K_g - 1_N K_g - K_g 1_N + 1_N K_g 1_N \tag{9}$$

where $1_N$ denotes a unit vector with the length is equal to N.

## IV. Related works of static methods

In this part, we present the theory of each studied methods. The KPCA, RKPCA, RRKPCA, KPLS and RKPLS methods have been applied for sensor fault detection and more generally for systems monitoring.

### 4.1 KPCA method

Since 1980s, PCA method has been successfully used in numerous field such as image processing, data information compression, features extraction and processes monitoring. Due to its efficiency and simplicity in processing huge amount of systems data, the PCA is widely used in practice [26], [27], [28]. The KPCA approach is the nonlinear extension of PCA.

Considering a transformed data matrix X presented by m variables and N measurement of observations taken under normal operating condition as

$$X = \left[X_1, X_2 \cdots ,,\ldots, X_n\right]^T \in \square^{N \times m} \tag{10}$$

Using nonlinear mapping $\phi$, a measured input observation is projected into feature space H with an hyper-dimensional order as developed in the previous section.

For the KPCA, Determining the number of retained PCs is an important step for monitoring system. The Cumulative Percent Variance (CPV) is utilize to determine the retain PCs [29]. The CPV is a measure that represent a percent variance explained by the $\ell$ first PCs. The number $\ell$ satisfies the criterion PCV and it is given by Eq.11.

$$\ell = \arg(PCV \geq 95) \tag{11}$$

where

$$CPV\,(\ell) = \frac{\sum_{i=1}^{\ell} \lambda_i}{\sum_{i=1}^{m} \lambda_i} 100 \tag{12}$$

Based on the concept of kernel, the KPCA approach can adequately map the input observations with a linearly relationship onto feature space H characterized by a higher dimensional, afterwards perform the linear PCA in the feature Hilbert space. An algorithm for applying classical KPCA is given in the following steps.

1- Considering an initial standardized block of learning measurement data, determine optimized kernel parameter;
2- Calculate the mean, the standard deviation of different variables system and standardize the matrix of measurement data;
3- build the Gram matrix and calculate it;
4- KPCA model is obtained;
5- Determine control threshold of the monitoring statistic;
6- Collect a new observation $x_t \in \mathbb{R}^m$ and compute it with the mean and variance in the offline phase;
7- Calculate the kernel vector $k_{x_t}; t = 1 \ldots N$ and scale it;
8- Perform monitoring statistic. If the control threshold is overtake, a sensor fault will be declared;
9- feedback to step 5;

**4.2 RKPCA method**

The monitoring techniques based on reduced model suggested by Taouali et al. [21] consists on approximating the retained a pertinent information about the system given by the KPCA method. We select, for the reduced model, a reduced number of measurements $x_b^{(k)} \in \{x^i\}_{\{i=1,\ldots,N\}}$ between the N available measurement variables of the data information matrix. Indeed, the selected data observations can be considered as a new data measurement matrix. The RKPCA technique approaches each vector $\{\gamma_j\}_{\{j=1,\ldots,\ell\}}$ by a transformed input measurement $\phi(x_b^k) \in \{\phi(x^{(i)})\}_{i=1,\ldots,N}$ which have a high projection value in the direction of $\gamma_j$ . Then, the RKPCA selects among the set of $\{\phi(x^{(i)})\}_{i=1,\ldots,N}$ the closest vectors the closest vectors, $\phi(x_b^{(j)})$ to $\gamma_j$ for each principal component. We project all the transformed $\{\phi(x^{(i)})\}_{i=1,\ldots,N}$ on the principal component $\gamma_j$ and we choose the measurements $\{\phi(x^{(i)})\}_{i=1,\ldots,N}$ that satisfied Eq.13.

$$\begin{cases} \phi(x_b^{(k)})_k = \underset{i=1,\ldots,N}{Max}\ \tilde{\phi}(x^{(i)})_k \\ \qquad and \\ \phi(x_b^{(k)})_{i \neq k} < \zeta \end{cases} \tag{13}$$

where $\zeta$ is a given threshold and $\tilde{\phi}(x^{(i)})_k$ is the $k^{th}$ component projection of $\phi(x^{(i)})$ on $\gamma_j$ . Furthermore, the downsized Gram matrix $K_r^{\ell}$ joined to a kernel function k can be present as show in Equation. (14):

$$K_r^{\ell} = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_\ell) \\ \vdots & & \vdots \\ k(x_\ell, x_\ell) & \cdots & k(x_\ell, x_\ell) \end{bmatrix} \in \mathbb{R}^{\ell x \ell} \tag{14}$$

**4.3 RRKPCA method**

For the KPCA model, the important amount of training data can lead a computational complexity. Therefore, the key principle of the RRKPCA approach is to eliminate the dependencies between variables in the feature Hilbert space and to select a reduced measurement from the original one [22].
The monitoring RRKPCA with fixed model can be defined as following. The first step is to recognize and determine the downsized reference model that presents the normal operating behavior. In order to recognize the reference RR-KPCA model, we memorize the important useful new observations in terms of process

information in a downsized training observations matrix. After that, the built model is performed on-line in order to detect faults. The initial reduced data matrix is expressed as:

$$X_r = [x_1] \in \mathbb{R}^{1 \times m} \tag{15}$$

A new measurement $x_t$ is available at each instant t. Then its transformed kernel vector $k_{x_t}$ is computed and also the kernel transformed matrix is changed by adding a column vector and a row data vector to the last one (Eq.16).

$$K_{rd}^{t} = \begin{bmatrix} K_{rd}^{t-1} & k_{x_t} \\ k_{x_t}^{T} & k(x_t, x_t) \end{bmatrix} \in \mathbb{R}^{rd \times rd} \tag{16}$$

In this case, the rank of the new kernel matrix $K_{rd}^{t}$ is determined, its value leading to either case:

- $rank(K_{rd}^{t}) = r$

  In this case the kernel matrix $K_{rd}^{t}$ has a full rank then the new measurements is added to the reduced data matrix. This matrix presents the independences between the projection observations in the feature Hilbert space.

- $rank(K_{rd}^{t}) < r$

  In this case the kernel matrix $K_{rd}^{t}$ has not a full rank then the reduced measurements matrix stay unchanged and afterwards we come back the kernel matrix $K_{rd}^{t}$ to its last state. This matrix describes the dependencies between the projection data in the feature Hilbert space. Finally, when all the measurements were evaluated, we obtain respectively the downsized data matrix and the reduced kernel matrix, $X_r$ and $K_r$, such that

$$K_r = \begin{bmatrix} k(x_1, x_1) \cdots k(x_1, x_r) \\ \vdots \qquad\qquad \vdots \\ k(x_r, x_1) \cdots k(x_r, x_r) \end{bmatrix} \in \mathbb{R}^{r \times r} \tag{17}$$

**4.4 KPLS method**

The PLS method builds linear multivariable regression model. This method extracts often a set a measurement vectors named LVs from the initial input/output measurements space. according to the related work, the system monitoring based on the PLS approach have been most discussed [13], [30], [31]. The KPLS is the nonlinear extension of PLS.

Considering the matrix of measurement $X_{in} \in \mathbb{R}^{N \times m}$ of input data containing N observations with m system variables and the matrix of the output $Y_{out} \in \mathbb{R}^{N \times J}$ of outputs comprising N measurements with J quality variables, we have:

$$X_{in} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_N \end{bmatrix} \in \mathbb{R}^{N \times m} \qquad Y_{out} = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_N \end{bmatrix} \in \mathbb{R}^{N \times J}$$

The PLS approach projects the input and output matrices to space with low-dimensional with an L number of LVs. Then the PLS method decomposes the matrices Xin and Yout as:

$$\begin{cases} Xin = TP^T + E \\ Yout = UQ^T + F \end{cases} \quad (18)$$

Where $P = [p_1, \dots, p_l]$ $and$ $Q = [q_1, \dots, q_l]$ $T = [t_1, \dots, t_l]$ and $and$ $U = [u_1, \dots, u_l]$
The corresponding prediction outputs on the learning observations can be determine using:

$$\hat{Y} = KgU\,(T^T\ KgU)^{-1}T^T Yout \quad (19)$$

The corresponding prediction outputs on the validation observations can be determined using:

$$\hat{Y}_t = K_{test}U\,(T^T KgU)^{-1}T^T\ Yout \quad (20)$$

where $K_{test}$ the validation observations Gram matrix.

**4.5 RKPLS method**

The learning data for kernel technique used for monitoring must be stored in a memory. The number of LVs for the industrial systems used for the KPLS approach may be greater than that for the linear case. For this reason, the time cost may increase due to the number of observations for the storage of the symmetric Gram matrix K during the learning step using the KPLS model. The RKPLS method is consider only the observation set that approximates correctly the selected crucial components to obtain a downsized Gram matrix [32].
The main objective of the RKPLS technique is to reduce the cost time and to select a reduced number of observations reduced number of observations about system. The RKPLS approach aims to replace each latent component $\{\xi_j\}_{j=1\dots p}$ by transformed input measurement $\phi(x_L^j) \in \phi(x_{Latent}^j)_{i=1,\dots,N}$ , which have the maximum projection value in the way of $\xi_j$ .

all the available vectors of the transformed measurement $\phi\{x^i\}_{i=1,\dots,N}$ are projected on the latent component $\xi_j$ and we select $x_{Latent}^{(j)} \in \{x^i\}_{i=1,\dots,N}$ that satisfies Eq.(21):

$$\begin{cases} \phi(x_{Latent}^{(j)})_j = \max\limits_{i=1,\dots,N} \tilde{\phi}(x^{(i)})_j \\ \quad and \\ \phi(x_{Latent}^{(j)})_{i\neq j} < \zeta \end{cases} \quad (21)$$

where $\zeta$ is a chosen threshold.

After determining the reduced data set $x_{Latent}^{(j)} \in \{x^i\}_{i=1,\dots,L}$ , a downsized matrix of data can be written as:

$$X_r = \left[x_{latent}^{(1)}, x_{latent}^{(2)} \cdots, \dots, x_{latent}^{(L)}\right]^T \quad (22)$$

The downsized Kernel Matrix Kr associated to a kernel function k, as presented in Eq. (23):

$$K_r = \begin{bmatrix} k(x_1, x_1)\cdots k(x_1, x_{Latent}) \\ k(x_2, x_1)\dots k(x_2, x_{Latent}) \\ \dots\dots \quad \dots \quad \dots \\ k(x_{Latent}, x_1)\cdots k(x_{Latent}, x_{Latent}) \end{bmatrix} \in \mathbb{R}^{Lx\,L} \quad (23)$$

## V. The proposed Reduced Rank-KPLS method

For the KPLS model, the important amount of training data can lead a computational complexity. Therefore, the key principle of the Reduced Rank KPLS approach is to eliminate the dependencies between variables in the feature Hilbert space and to select a reduced measurement from the original one.

The monitoring RR-KPLS with fixed model can be defined as following. The first step is to recognize and determine the downsized reference model that presents the normal operating behavior. In order to recognize the reference RR-KPLS model, we memorize the important useful new observations in terms of process information in a downsized training observations matrix. After that, the built model is performed on-line in order to detect faults. The initial reduced data matrix is expressed as:

$$X_D = [x_1] \in \mathbb{R}^{1x\,m} \quad (24)$$

A new measurement $x_q$ is available at each instant t. Then its transformed kernel vector $k_{x_q}$ is computed and also the kernel transformed matrix is changed by adding a column vector and a row data vector to the last one (Eq.25).

$$K_D^q = \begin{bmatrix} K_r^{q-1} & k_{x_q} \\ k_{x_q}^T & k(x_q, x_q) \end{bmatrix} \in \mathbb{R}^{D \times D} \tag{25}$$

In this case, the rank of the new kernel matrix $K_D^q$ is determined, its value leading to either case:

- $rank(K_D^q) = D$

  In this case the kernel matrix $K_D^q$ has a full rank then the new measurements is added to the reduced data matrix. This matrix presents the independences between the projection observations in the feature Hilbert space.

- $rank(K_D^q) < r$

  In this case the kernel matrix $K_D^q$ has not a full rank then the reduced measurements matrix stay unchanged and afterwards we come back the kernel matrix $K_{rd}^t$ to its last state. This matrix describes the dependencies between the projection data in the feature Hilbert space. Finally, when all the measurements were evaluated, we obtain respectively the downsized data matrix and the reduced kernel matrix, $X_D$ and $K_D$, such that

$$K_D = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_D) \\ \vdots & & \vdots \\ k(x_D, x_1) & \cdots & k(x_D, x_D) \end{bmatrix} \in \mathbb{R}^{D \times D} \tag{26}$$

## VI. Fault detection

The kernel methods for monitoring procedure can usually use the SPE statistic in the feature space. The SPE index is calculate as the norm of the residual vector belongs to the feature Hilbert space, which is shown as follows (Eq. (27))

$$SPE = \left\| k(x,x) - k_{x_t}^T \hat{P} \hat{\Lambda}^{-1} \hat{P}^T k_{x_t} \right\| \tag{27}$$

Where: $k_{x_t} = [k(x_1, x_t), \ldots, k(x_N, x_t)]^T \quad i = 1 \ldots N$, $\hat{P}$ is the matrix of the principal eigenvectors of kernel matrix K and finally $\hat{\Lambda}$ is the diagonal matrix of kernel matrix K

The confidence threshold for SPE index can be computed according to the $\chi^2$-distribution and is written by Eq.28.

$$SPE \approx \delta_\alpha^2 \tag{28}$$

Where $\delta_\alpha^2$ represents the upper control threshold for the SPE with a significance level $\alpha$

The confidence threshold $\delta_\alpha^2$ for the SPE with a significance level can be computed as:

$$\delta_\alpha^2 = g \chi_{h,\alpha}^2 \tag{29}$$

where the confidence level is $(1-\alpha) \times 100\%$ 100%, and g and h are determined using:

$$g = \frac{Variance(SPE)}{2\ mean(SPE)} \text{ and } h = \frac{(2\ mean(SPE))^2}{Variance(SPE)}$$

## VII. Application

The fault detection performances of the kernel methods: KPCA, KPLS, RKPCA, RKPLS, RRKPCA and the proposed RR-KPLS with fixed model are presented in this section. Two fault detection criteria are determined in order to show the performance of the studied methods: The Good Detection Rate (GDR) percent and the False Alarm Rate (FAR) value.

The FAR is considered as

$$FAR = \frac{\text{Volated samples}}{\text{Faultless data}}\% \tag{30}$$

The GDR is calculated as:

$$GDR = \left( \frac{\text{Volated samples}}{\text{Faulty data}} + \frac{\text{Not violated samples}}{\text{Faultless data}} \right)\% \tag{31}$$

### 6.1 Tennessee Eastman process description

The performances of the statistics methods for systems monitoring is studied by applying the TEP system. in Fig.1. The TEP contains two outputs of variables G and H from four reactants: A, C, D and E. The reaction scheme is as follows:

$$\begin{aligned} A(g) + C(g) + D(g) &\rightarrow G(liq) \\ A(g) + C(g) + E(g) &\rightarrow H(liq) \\ A(g) + E(g) &\rightarrow F(liq) \\ 3D(g) &\rightarrow 2F(liq) \end{aligned} \tag{29}$$

The $X_{train}$ matrix presents 22 variables continuously measured from 41 process variables. The 19 variables are considered for the quality matrix $X_{train}$. The variables used to build the data matrix are determined in Table 1. The fault is introduced at 224 observations for the testing data set. Identification, monitoring and modeling represent the subject of several studies and also represent a challenge for the control community for the TEP. Then 21 faults types could be injected as indicated in Table 2.

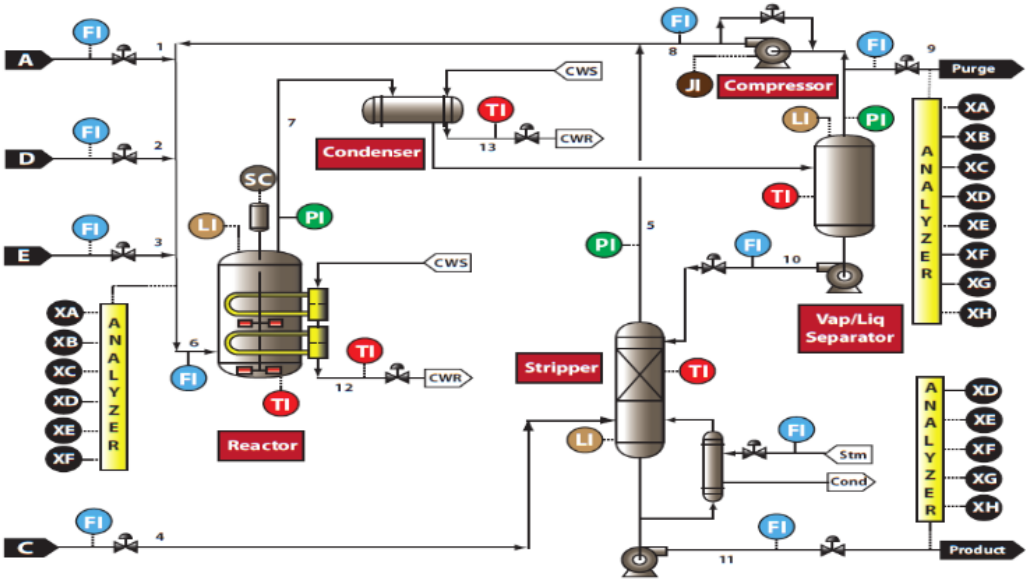



Figure 1: Flow diagram of TEP.

Table 1: Measurement system variables

| Variables | Description | Unit |
|---|---|---|
| $x_1$ | A feed (stream 1) | kscm h |
| $x_2$ | D feed (stream 2) | kg $h^{-1}$ |
| $x_3$ | E feed (stream 3) | kg $h^{-1}$ |
| $x_4$ | A and C feed (stream 4) | kscm h |
| $x_5$ | Recycle flow | kscm h |
| $x_6$ | Reactor feed rate | kscm h |
| $x_7$ | Reactor pressure | kPa |
| $x_8$ | Reactor level | % |
| $x_9$ | Reactor temperature | °C |
| $x_{10}$ | Purge rate | kscm h |
| $x_{11}$ | Separator temperature | °C |
| $x_{12}$ | Separator level | % |
| $x_{13}$ | Separator pressure | kPa |
| $x_{14}$ | Separator underflow | $m^3$ $h^{-1}$ |
| $x_{15}$ | Stripper level | % |
| $x_{16}$ | Stripper pressure | kPa |
| $x_{17}$ | Stripper underflow | $m^3$ $h^{-1}$ |
| $x_{18}$ | Stripper temperature | °C |
| $x_{19}$ | Stripper steam flow | kscm h |
| $x_{20}$ | Compressor work | $m^3$ $h^{-1}$ |
| $x_{21}$ | Reactor water temperature | °C |
| $x_{22}$ | Separator water temperature | °C |

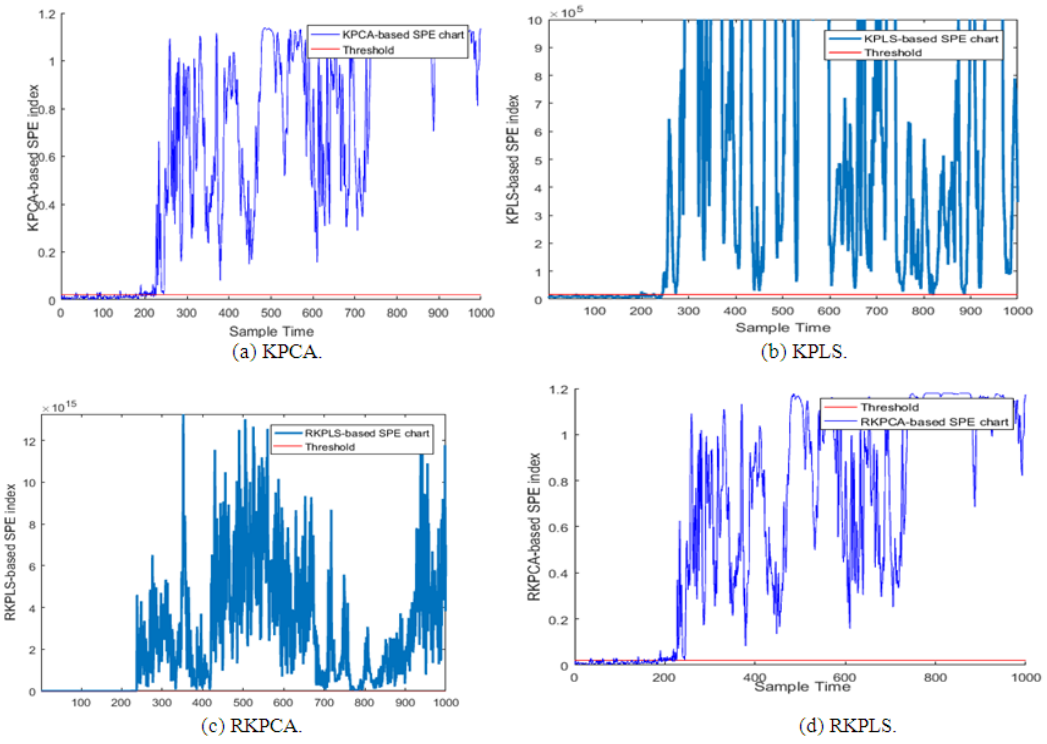
**6.2 Simulation results**

Fig.2 shows the FD performance of the studied methods for IDV1(12), fault test data. For fault IDV1(12), the figures of presented system monitoring are given in Fig.2 in which the reduced methods offer better results than the conventional KPCA and KPLS methods. Moreover, RKPLS, RKPCA and RRKPCA offer not only lower FAR but also correct fault diagnosis information about the properties of the faults. From Fig.2, it can be clearly show that the KPLS and KPCA are higher sensitive in FAR than the reduced approaches. In addition, the results using the SPE index are shown by Fig.2, from which the RR-KPLS provide much better fault detection performance than the standard one.

Table 2: List of monitoring variables in TEP.

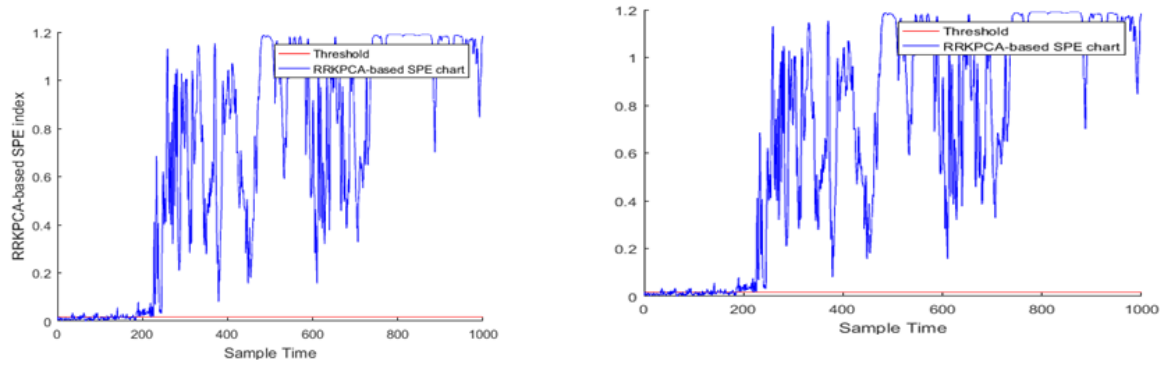
| Fault Number | Process Variable | Type |
|---|---|---|
| IDV1(1) | the A/C feed ratio, a B composition constant | a Step |
| IDV1(2) | the B composition, a A/C ratio constant | a Step |
| IDV1(3) | the D feed temperature | a Step |
| IDV1(4) | the Reactor cooling water inlet temperature | a Step |
| IDV1(5) | the Condenser cooling water inlet temperature | a Step |
| IDV1(6) | the A feed loss | a Step |
| IDV1(7) | the C header pressure loss-reduced availability | a Step |
| IDV1(8) | the A, B, and C feed composition | a Random variation |
| IDV1(9) | the D feed temperature | a Random variation |
| IDV1(10) | the C feed temperature | a Random variation |
| IDV1(11) | the Reactor cooling water inlet temperature | a Random variation |
| IDV1(12) | the Condenser cooling water inlet temperature | Random variation |
| IDV1(13) | the Reaction kinetics | a Slow drift |
| IDV1(14) | the the Reactor cooling water valve | a Sticking |
| IDV1(15) | the Condenser cooling water valve | a Sticking |
| IDV1(16) | an Unknown | an Unknown |
| IDV1(17) | a Unknown | a Unknown |
| IDV1(18) | an Unknown | an Unknown |
| IDV1(19) | Uan nknown | an Unknown |
| IDV1(20) | the Valve fixed at steady state position | a Constant position |
| IDV1(21) | the A/C feed ratio, a B composition constant | a Step |

Table 3 provides the FAR and GDR for the validation data sets of faults to test the studied methods. Table 3 summarize the detailed GDR and FAR.




(a) KPCA. (b) KPLS.

(c) RKPCA. (d) RKPLS.




(e) RRKPCA. (f) RRKPLS.

Figure 2: Process monitoring using KPCA, KPLA, RKPCA, RKPLS, RRKPCA and RR-KPLS in case of IDV1(12).

Table 3: FAR and GDR based on different studied methods.

| Faults | KPCA | | RKPCA | | RRKPCA | | KPLS | | RKPLS | |
|---|---|---|---|---|---|---|---|---|---|---|
| | FAR(%) | GDR (%) | FAR(%) | GDR (%) | FAR(%) | GDR (%) | FAR(%) | GDR (%) | FAR(%) | GDR (%) |
| IDV1(1) | 26.87 | 100 | 8.03 | 100 | 31.69 | 100 | 6.25 | 95.02 | 1.32 | 99.9 |
| IDV1(2) | 23.21 | 99.22 | 14.46 | 98.96 | 34.37 | 99.48 | 49.8 | 20.03 | 29 | 45 |
| IDV1(3) | 33.92 | 60.82 | 29.18 | 34.66 | 41.96 | 77.70 | 62.08 | 50 | 34 | 35.6 |
| IDV1(4) | 24.55 | 47.8 | 6.69 | 21.13 | 23.21 | 55.92 | 22 | 56.66 | 17 | 43.33 |
| IDV1(5) | 24.55 | 64.17 | 6.69 | 96.26 | 23.21 | 68.17 | 1.05 | 19.84 | 0.91 | 98 |
| IDV1(6) | 20.53 | 100 | 4.91 | 100 | 22.32 | 100 | 14.44 | 84.05 | 2.2 | 97.33 |
| IDV1(7) | 24.55 | 76.03 | 6.25 | 61.98 | 26.78 | 85.95 | 0.44 | 77.46 | 0.02 | 99.61 |
| IDV1(8) | 41.07 | 100 | 20.08 | 99.74 | 44.64 | 100 | 0.448 | 96 | 4.05 | 99.23 |
| IDV1(9) | 51.33 | 56.18 | 33.48 | 87.96 | 56.25 | 64.43 | 21.34 | 56.55 | 0.11 | 99.33 |
| IDV1(10) | 21.87 | 92.91 | 4.46 | 83.50 | 25 | 95.10 | 23.11 | 68.8 | 0.02 | 67.64 |
| IDV1(11) | 29.01 | 81.31 | 10.71 | 64.17 | 27.67 | 89.04 | 38.7 | 44.09 | 21.01 | 96.64 |
| IDV1(12) | 31.69 | 100 | 2.53 | 99.87 | 33.92 | 100 | 9.55 | 97.29 | 0.01 | 97.33 |
| IDV1(13) | 21.42 | 97.68 | 5.35 | 98.87 | 19.19 | 98.45 | 38.82 | 90 | 5.31 | 96.64 |
| IDV1(14) | 25.89 | 100 | 7.58 | 100 | 27.23 | 100 | 34.4 | 92.78 | 0.9 | 97.39 |
| IDV1(15) | 33.03 | 53.47 | 18.84 | 60.41 | 37.94 | 69.71 | 22.22 | 47.4 | 1.73 | 89.06 |
| IDV1(16) | 58.03 | 85.30 | 11.07 | 96.31 | 57.58 | 92.91 | 19.98 | 45.2 | 10.05 | 100 |
| IDV1(17) | 34.37 | 98.06 | 11.6 | 97.03 | 36.6 | 98.45 | 43.09 | 56 | 27 | 47 |
| IDV1(18) | 29.91 | 95.37 | 12.16 | 92.78 | 30.80 | 97.55 | 22.67 | 44.4 | 13 | 46 |
| IDV1(19) | 29.64 | 81.05 | 12.05 | 49.74 | 29.46 | 90.20 | 27.09 | 43.34 | 10.71 | 99 |
| IDV1(20) | 22.32 | 88.01 | 8.46 | 84.22 | 21.87 | 96.13 | 19.36 | 98 | 10.71 | 100 |
| IDV1(21) | 40.62 | 75.9 | 24.10 | 33.53 | 50 | 83.50 | 33 | 53 | 27.45 | 55 |

According to Table 3, the fault detection results based on kernel methods in term of FAR and GDR are presented. In this table, each method shows his performance in detection according to conditions.

The reduced methods (RKPCA, RKPLS) give a best performances compared to the classical methods based on SPE index. The results of the proposed method RR-KPLS based on SPE present a better results compared to the classical methods. On the other hand, the reduced rank method (RRKPCA) presents a good performance compared to the classical methods in terms of GDR. In addition, some faults listed are properly undetectable by the given conventional methods. The RKPLS approach gives the best FAR listed in Table 3 over all the other methods. The RKPCA and RRKPCA approach give the best GDR listed in Table 3 over all the other methods. Furthermore, Table 3 shows that the RKPLS based SPE provide good results compared to the other discussed methods.

Furthermore, the next table present the CT for TEP data.

Table 4: Summary of computation time for different studied methods

| Methods | KPCA | RKPCA | RRKPCA | KPLS S | RKPLS | RR-RKPLS |
|---|---|---|---|---|---|---|
| CT(s) | 12.58 | 0.79 | 0.99 | 0.89 | 0.23 | 0.11 |

The reduced techniques based on SPE shows better fault detection performances with CT as illustrated in Table 4. The RKPCA gives the small value of CT compared to the conventional KPCA and RRKPCA. However, the RR-KPLS method present the small value of CT compared to the all discussed methods.

Despite the simplicity of the KPCA method compared to KPLS, the proposed RR-KPLS method presents a more effective fault detection performance.

## VIII. Conclusions

In this article, the basic data-driven approaches and their recent developments of fault detection were firstly reviewed. Then the RR-KPLS method is proposed. Two variants of KPCA technique, i.e. RKPCA and RRKPCA, offer much better FAR and more accurate fault diagnosis information compared to the standard approach. The downsized methods consist essentially to improve the computational time, focused on the important elements and observations.

However, one variants of KPLS, i.e. RKPLS, offer much better FDR, GDR and more accurate fault diagnosis information compared with the standard approach. Indeed, the RR-KPLS solves the problem of Computing Time and the storage of variables. All the discussed and studied techniques were applied on an industrial systemof TEP to complete a comparison study.

## References


[1]. Jafel I, Taouali O, Elaissi E, Messaoud H (2014) A new online fault detection method based on PCA technique. IMA Journal of Mathematical Control and Information 31(4):487-499

[2]. Tharrault Y, Mourot G, Ragot J, Maquin D (2008) Fault detection and isolation with robust principal component analysis. International Journal of Applied Mathematics and Computer Science 18(4):429-442

[3]. Muradore R, Fiorini P (2011) A PLS-based statistical approach for fault detection and isolation of robotic manipulators. IEEE Transactions on Industrial Electronics 59(8):3167-3175

[4]. Taouali O, Elaissi E, Messaoud H (2012) Design and comparative study of online kernel methods identi_cation of nonlinear system in RKHS space. Arti_cial Intelligence Review 37(4):289-300

[5]. Jiao J, Zhao N, Wang G, Yin S (2017) A nonlinear quality-related fault detection approach based on modi_ed kernel partial least squares.ISA transactions 66:275-283

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

[7]. Rumelhart D E (1986) Parallel distributed processing: Explorations in the microstructure of cognition. Parallel distributed processing: Explorations in the microstructure of cognition 1:318-362

[8]. Zhao Y, Xiao F, Wen J, Lu Y, Wang S (2014) A robust pattern recognition-based fault detection and diagnosis (FDD) method for chillers. HVAC&R Research 20(7):798-809

[9]. Chetouani Y (2008) A neural network approach for the real-time detection of faults. Stochastic Environmental Research and Risk Assessment 22(3):339-349

[10]. Zhao Y, Wang S, Xiao F (2013) Pattern recognition-based chillers fault detection method using support vector data description (SVDD). Applied Energy 112:1041-1048

[11]. Kano M, Tanaka S, Hasebe S, Hashimoto I, Ohno H (2003) Monitoring independent components for fault detection. AIChE Journal 49(4):969-976

[12]. Harkat MF, Djelel S, Doghmane N, Benouaret M (2007) Sensor fault detection, isolation and reconstruction using nonlinear principal component analysis. International Journal of Automation and Computing 4(2):149-155

[13]. Taouali O, Elaissi I, Messaoud H (2015) Dimensionality reduction of RKHS model parameters. ISA transactions 57:205{210

[14]. Taouali O, Elaissi I, Messaoud H (2015) Fault detection based on kernel principal component analysis. Engineering Structures 32(11):3683-3691

[15]. Rosipal R, Trejo L J (2001) Kernel partial least squares regression in reproducing kernel Hilbert space. Journal of machine learning research 2(Dec):97-123

[16]. Zhang Y, Du W, Fan Y, Zhang L (2015) Process fault detection using directional kernel partial least squares. Industrial & Engineering Chemistry Research 54(9):2509-2518
[17]. Yin S, Ding S X, Haghani A, Hao H, Zhang P (2012)A comparison study of basic data driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. Journal of process control 22(9):1567-1581
[18]. Lahdhiri H, Taouali O, Elaissi I, Ja_el I, Harakat MF, Messaoud H (2017) A new fault detection index based on Mahalanobis distance and kernel method. The International Journal of Advanced Manufacturing Technology 91(5-8):2799-2809
[19]. Kim K, Lee J, Lee I (2005) A novel multivariate regression approach based on kernel partial least squares with orthogonal signal correction. Chemometrics and intelligent laboratory systems 79(1-2):22-30
[20]. Jia M, Xu H, Liu X, Hao H, Wang N (2012) The optimization of the kind and parameters of kernel function in KPCA for process monitoring. Computers & chemical engineering 46:94-104
[21]. Taouali O, Ja_el I, Lahdhiri H, Harkat M F, Messaoud H (2016) New fault detection method based on reduced kernel principal component analysis (RKPCA). The International Journal of Advanced Manufacturing Technology 85(5-8):1547-1552
[22]. Lahdhiri H, Elaissi I, Taouali O, Harkat M F, Messaoud H (2016) Nonlinear process monitoring based on new reduced Rank-KPCA method. Stochastic Environmental Research and Risk Assessment 32(6):1833-1848
[23]. Said M, ben Abdellafou K, Taouali O, Harkat M F (2019) A new monitoring scheme of an air quality network based on the kernel method. The International Journal of Advanced Manufacturing Technology 103(1-4):153-163
[24]. Schlolkopf B, Smola A, Muller K (1998) Nonlinear component analysis as a kernel eigenvalue problem. Neural computation 10(5):1299-1319
[25]. Mercer J (1909) Functions of positive and negative type and their connection with the theory integral equations. Philosophicol Trinsdictions ofthe Rogyal Society 4-415
[26]. Lahdhiri H, Said M, Abdellafou K, Taouali O, Harkat M F (2019) Supervised process monitoring and fault diagnosis based on machine learning methods. The International Journal of Advanced Manufacturing Technology 102(5-8):2321{-337
[27]. Neffati S, Ben Abdellafou K, Taouali O, Ja_el I, Bouzrara K (2019) An improved machine learning technique based on downsized KPCA for Alzheimer's disease classifcation. International Journal of Imaging Systems and Technology 29(2):121-131
[28]. Said M, Fazai R, Ben Abdellafou K, Taouali O (2018) Decentralized fault detection and isolation using bond graph and PCA methods. The International Journal of Advanced Manufacturing Technology 99(1-4):517-529
[29]. Ja_el I, Taouali O, Harkat M F, Messaoud H (2016) Moving window KPCA with reduced complexity for nonlinear dynamic process monitoring. ISA transactions 64:184-192
[30]. MacGregor J F, Jaeckle C, Kiparissides C, Koutoudi M (1994) Process monitoring and diagnosis by multi block PLS methods. AIChE Journal 40(5):826-838
[31]. Li G, Qin S J, Zhou D (2010) Geometric properties of partial least squares for process monitoring. Automatica 46(1):204-210
[32]. Said M, Ben Abdellafou K, Taouali O (2019) Machine learning technique for data driven fault detection of nonlinear processes. Journal of Intelligent Manufacturing 1-20

Okba Taouali. "Process monitoring based a Novel Reduced Rank KPLS approach." *IOSR Journal of Computer Engineering (IOSR-JCE)*, 23(6), 2021, pp. 13-26.