People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research

University M'Hamed BOUGARA – Boumerdes

Institute of Electrical and Electronic Engineering

Department of Power and Control

Final Year Project Report Presented in Partial Fulfilment of
the Requirements for the Degree of

MASTER

In **Control**

Option: **Control**

Title:

Reduced Kernel PCA based Approach for Fault Detection in Complex Systems

Presented by:

BENNAI Sabrina.

Supervisor:

Dr.KOUADRI Abdelmalek.

Mr.BENCHEIKH Fares (Co-Supervisor).

Registration Number:………/2019

ABSTRACT

Multivariate statistical methods have been widely applied to complex systems for fault detection. While methods based on principal component analysis (PCA) are popular, more recently kernel PCA (KPCA) has been utilized to better model nonlinear process data.

This report proposes a new method for fault detection using a reduced kernel principal component analysis (RKPCA) to cope with the computational problem introduced by KPCA. The proposed RKPCA method consists on reducing the number of observations in a data matrix using the dissimilarities between the pairs of its observations.

PCA, KPCA and the suggested approach RKPCA are carried out using the cement rotary kiln system. The Hotelling's T^2, Q in addition to the new proposed index called the combined statistic φ are used as fault indicators. The two methods PCA and KPCA are compared to the proposed approach in terms of False Alarms Rate (FAR), Missed Alarms Rate (MDR), Detection Time Delay (DTD), the cost function (J) and the Execution Time (ET).

The obtained results demonstrate the effectiveness of the proposed technique in reducing the computational time from 1h37min when KPCA is used to 9min30s.Moreover, it has effectively detected the different types of faults when using the φ index.

Keywords: Fault detection; Principal Component analysis (PCA); Kernel Principal Component analysis (KPCA); Reduced Kernel Principal Component analysis (RKPCA); the combined index.

First, I thank the Almighty God, thank you for the guidance, strength, power of mind and protection. All of these, I offer to you.

This study is wholeheartedly dedicated to:

*My **beloved Parents**, who have been my source of inspiration and gave me strength when I thought of giving up, who continually provide their moral, spiritual and emotional support*

*My brothers **Nadir** and **Fouad***
*My beloved sister **Kahina** for her infinite support,*
*My sister and brother in law **Dihia** and **Farid,***
*My dear nephews **Aksel** and **Gaya**.*

*To all my **friends** and **classmates** who shared their words of advice and encouragement to finish this study.*

AKNOLEGEMENT

I would like to express my deep and sincere gratitude to my supervisor, Dr. ***Abdelmalek KOUADRI*** *for providing me invaluable guidance throughout this work.*
His dynamism, vision and motivation have deeply inspired me. It was a great privilege and honor to work and study under his guidance.

I would like to thank Mr. ***Fares BENCHEIKH*** *for his help.*
I really appreciate all his contributions of time and ideas to make this experience productive and stimulating.

I would also like to acknowledge honorary all the teachers of the Institute of Electrical and Electronics Engineering of the University of M'Hamed Bougara- Boumerdes for their support and help during the last five years. I have learned a lot from their knowledge and moralities.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

AEM: Abnormal Event Management

CPV: Cumulative Percent Variance

DTD: Detection Time Delay

FAR: False Alarm Rate

FDD: Fault Detection and Diagnosis

FDI: Fault Detection and Identification

KPCA: Kernel Principal Component Analysis

MDR: Missed Alarm Rate

MSPC: Multivariate Statistical Process Control

NN: Neural Network

NOC: Normal Operation Condition

PC's: Principal Components

PCA: Principal Component Analysis

PLS: Partial Least Square

RKPCA: Reduced Kernel Principal Component Analysis

SPC: Statistical Process Control

SPE: The Sum Squared of Error

GENERAL INTRODUCTION

The industrial world has had very rapid and important development thus the manufacturing processes are becoming increasingly complex. For the improvement of reliability, availability, and safety, it is necessary to ensure efficient process monitoring techniques. This is achieved by applying advanced methods of supervision and fault diagnosis methods. Several fault detection techniques have been developed in literature [1, 2]. Venkatasubramanian *et al*. (2003) have proposed a classification of these methods based on the methodology used to construct fault diagnostic models from data, drawing a distinction between statistical methods and neural networks. The work has listed the fault detection and diagnosis methods in three different categories: (i) quantitative model-based schemes, (ii) qualitative models and search strategies and (iii) process data-based methods [3].

In industrial chemical and biological processes, data-based approaches are seen as the most cost effective approach to dealing with the complex systems and have seen explosive growth over the last few decades [4, 5]. They are referred to as statistical process control (SPC), and conventional SPC charts such as Shewhart control charts, cumulative sum (CUSUM) control charts, and exponentially weighted moving average (EWMA) control charts have been widely used. Such SPC charts are well established for monitoring univariate processes, but they do not function well for multivariable processes. Therefore, multivariate statistical process control (MSPC) techniques have been developed in order to extract useful information from process data and utilize it for process monitoring [5-8]. Principal component analysis (PCA) [9, 10] is among the most popular statistical methods used for modeling and faults detection problems, it is capable of compressing high-dimensional data with little loss of information by projecting the data onto a lower-dimensional subspace defined by a new set of derived variables (principal components (PCs)) [11]. Various fault detection techniques have been developed and utilized in practice. The main indices used with PCA methods are Hotelling's T^2, and the sum of squared residuals, SPE, or Q statistic. In some complex processes with particularly nonlinear characteristics, PCA performs poorly due to its assumption that the process data are linear [12]. To address the nonlinearity problem, Kramer (1992) proposed a nonlinear PCA based on auto-associative neural network (NN) [12]. Dong & McAvoy (1996) suggested a nonlinear PCA that combined principal curve and NN [13]. Alternative nonlinear PCA methods developed by Hiden *et al*. (1999), Cheng and Chiu (2005), Kruger *et al*. (2005) and Maulud *et al*. (2006) have also been proposed to solve the nonlinear process monitoring problem [14-17]. However, most of the existing non- linear PCA approaches are based on neural network, which has to solve a nonlinear optimization problem. A new nonlinear PCA technique for tackling the nonlinear problem, called kernel PCA (KPCA), has been in development in recent years (18-20). The main idea of KPCA is first to map the input space into a feature space via a nonlinear map, which makes data structure more linear, and then to extract principal components in the feature space [18-20]. Similar to the PCA, the Hotelling's T^2 statistic and the Q statistic are two indices commonly used in KPCA-based process monitoring [21].

Although capable of capturing nonlinear relationships between variables with a high degree of accuracy, kernel principal component analysis (KPCA) suffers from a high computational cost and requires the storage of the symmetric kernel matrix (computation time increases with the number of samples) [4].
Existing studies for reducing the computational cost of kernel PCA have been proposed including Approximate and iterative techniques [22]. Approximate approaches suggested by Lopez-Paz *et al*. (2014) construct a low rank approximator of the kernel matrix, and use its eigensystems as an alternative. Due to the low-rank structure, the approximator can be easily stored and manipulated. The major limitation of approximate approaches is that there always exists a non-vanishing gap between their solution and that found by eigendecomposing K directly [23]. Iterative approaches proposed by Kim *et al*. (2005) use partial information of K in each round to estimate the top eigenvectors, and thus do not need to keep the entire matrix in memory. With appropriate initialization, the solution of iterative approaches will converge to the groundtruth asymptotically. However, there is no guarantee of the convergence rate or the global convergence property for general initial conditions [24].

In this project, as an alternative and effective method to detect abnormalities in processes with nonlinear nature and to deal with the problem of need storage and computation time when using Kernel principal component (KPCA), a new nonlinear method is suggested. We propose a new reduced version of KPCA which consists on approximating the information matrix by a set of observations using the distance between observations such that only one observation is preserved in case of similarity. The Hotelling's T^2 and Q statistics are used as faults indicators. In addition, we propose a new index called the combined index or simply φ statistic which incorporates the T^2 and Q statistics in a balanced way [25, 26].

The proposed method (RKPCA) has been tested on a cement manufacturing plant and compared to PCA and KPCA in terms of False Alarm Rate (FAR),Missed Alarm Rate (MDR),Detection Time Delay (DTD), the Cost function (J) and the Execution Time (ET).

This report is organized as follow: Chapter I presents the theoretical background required to know about fault detection and diagnosis field. Chapter II the mathematical foundation of PCA, KPCA and the proposed method RKPCA are explored and their application as multivariate statistical process control (MSPC) techniques in complex processes are demonstrated. Chapter III provides the detailed description of the cement plant followed by the experimental setup. The obtained results are discussed and concluding remarks are drawn.

Chapter I.
Fault detection and diagnosis

I.1. Introduction

Fault detection and diagnosis is an important problem in process engineering. It is the central component of abnormal event management (AEM) which deals with the timely detection, diagnosis and correction of abnormal conditions of faults in a process [3]. Over the years, the increasing need for efficiency and product quality has encouraged the development of many fault diagnosis methods and which can be classified into three categories: Quantitative Model based methods, Qualitative model based methods, and process history based methods [27].

In this chapter, we first address the definitions and nomenclature used in the area of process fault diagnosis. In the next section, we propose a list of ten desirable characteristics that one would like a diagnostic system to possess. In section 4, we discuss the transformations of data that take place during the process of diagnostic decision making. In section 5, a classification of fault diagnosis methods is provided.

I.2. Definitions of fault detection and diagnosis

A fault is generally defined as a departure of an observed variable or calculated parameter from an accepted range [28]. More specifically, a fault is an unpermitted deviation of at least one characteristic property of a variable from an acceptable behavior. This means that a fault may lead to the malfunction or failure of the system [29].

The time dependency of faults may show up as (Fig. 1).

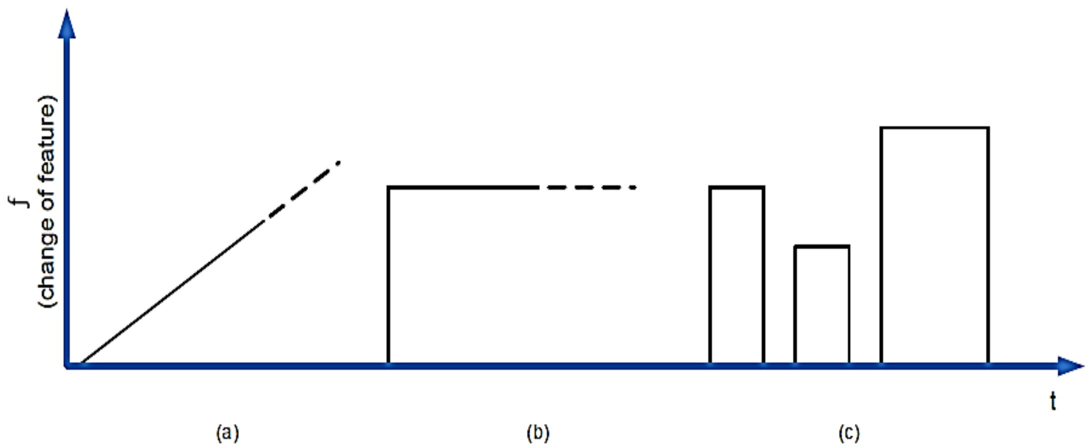

Figure 1. Time dependency of faults: (a) Incipient; (b) Abrupt; (c) Intermittent [29].

- Abrupt fault (stepwise): The cause of the fault and/or the effects remains continuing until corrected.
- Incipient fault (drift-like): The effect on process remains constant until corrected.
- Intermittent fault (with interrupts): The effect disappears and reappears in time.

Faults can also be classified as additive and multiplicative faults (Fig.2). Additive faults influence a variable by the addition of fault features. These types of faults can appear as offsets in a process metric from a normal or desired value. Multiplicative faults affect the variable as a product and often manifest themselves as parameter changes with the process [29].

Faults may include:

- **Structural changes**

Structural changes results from changes in the process itself. They occur due to hard failures in the process equipment. Structural malfunctions result in a change in the information flow between various variables. An example of a structural failure would be failure of a controller. Other examples include a stuck valve, a broken or leaking pipe and so on [3].

- **Malfunctioning sensors and actuators**

Errors that usually occur with actuators and sensors are due to a fixed failure, a constant bias (positive or negative). The occurrence of a failure in one of the instruments could cause the plant state variables to deviate beyond acceptable limits. It is the purpose of diagnosis to quickly detect any instrument fault which could seriously degrade the performance of the control system [3].

- **Gross parameter changes in a model**

Parameter failures happen when there is a disturbance entering from the environment through one or more independent variables. An example of such a malfunction is a change in the concentration of the reactant from its normal or steady state value in a reactor feed. Here, the concentration is an independent variable [3].

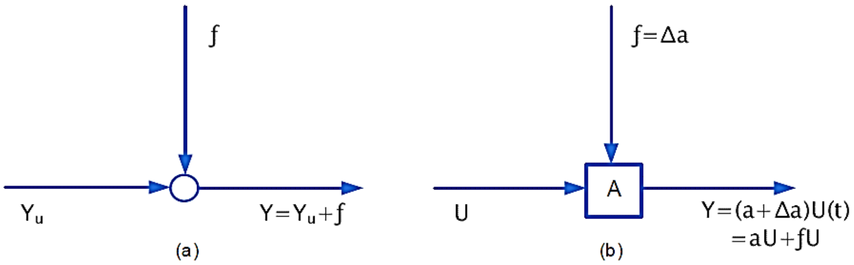

Figure 2. Basic fault models: (a) Additive and (b) Multiplicative faults [3].

Within the context of supervision, supervisory functions exist to indicate undesired or not-permitted process states and to take action in reaction to these indications [29]. We can distinguish between the following functions:

- Monitoring: Variables are measured and compared to alarm limits. This information is presented to an operator.
- Automatic Protection: Actions are taken in reaction to a dangerous state (includes interlocks).
- Fault Detection and Diagnosis: Data is measured; features calculated; symptoms detected; diagnosis and decision on actions made.

The obvious advantages of the first two functions are simplicity and reliability. The disadvantage is the delay in reacting to sudden or gradually increasing faults. It is also not possible to perform in depth diagnosis. Therefore, a method having the features of the third function is necessary. This will allow deep insight into the process behavior.

I.3. Desirable attributes of a fault detection and diagnosis system

In order to compare methodologies or to assess whether a fault detection and diagnosis system is successful, it is useful to identify a set of desirable characteristics that a diagnostic system should possess [3].

I.3.1. Quick detection and diagnosis

Fault detection and diagnosis system should respond to a failure or a malfunction as quickly as possible. However, systems designed for rapid detection are sensitive to high frequency influences, which makes the system sensitive to noise leading to frequent false alarms in even normal operations [3].

I.3.2. Isolability

This characteristic shows the capability of a diagnostic system to distinguish multiple failures; these failures sometimes overlap with modeling uncertainties in terms of residuals. For example, if a diagnostic system, such as analytical-based or rule-based methods, requires multiple models in the procedures, rejecting modeling uncertainties will frequently occur and isolability will gradually decline for each step of multiple modeling processes. Therefore, there is a trade-off between isolability and the rejection of modeling uncertainties in an appropriate diagnostic system [3].

I.3.3. Robustness

A diagnostic system satisfying robust feature means its performance should be insensitive to the effect of various noise and modeling uncertainties [3].

I.3.4. Novelty identifiability

A fault detection and diagnosis system should be able to decide whether a process is in a normal or malfunction operation and, if an abnormal condition occurs, whether the causes are from known or novel unknown malfunction [3].

I.3.5. Adaptability

It is also desirable to have extendable systems. This would allow processes to change due to changes in the external inputs, structural changes and also changes in the environmental conditions. Thus the diagnostic system should be adaptable to changes [3].

I.3.6. Classification error estimate

The evaluation of a diagnostic system can be performed through this feature in terms of accuracy and reliability to encourage the confidence of users. The feature of error estimate shows the efficiency of diagnostic decisions [3].

I.3.7. Explanation facility

Fault detection and diagnosis system should explain where and how faults occur in a system. This feature is significantly required for on-line decision applications of diagnostic classifiers. It is useful for building operators to examine a system according to the explanations or recommendations of a fault detection and diagnosis approach [3].

I.3.8. Modelling requirements

Number of modeling methods of FDD should be as minimal as possible for quick and easy implementations; otherwise, the method will be not suitable to apply in real-time applications. Rules-based and process historical methods generally require no modeling process, whereas the detailed first-principle model based method may hardly satisfy this feature [3].

I.3.9. Multiple fault identifiability

The ability to identify multiple faults is an important but a difficult requirement. It is a difficult problem due to the interacting nature of most faults. Naturally, the combination of several faults typically occurs in nonlinear or larger systems, leading to the difficult separation of individual fault [3].

I.3.10. Storage and computational requirements

This criterion is specifically required for the fast real-time implementation of diagnostic classifiers. Then, the systems should be reasonably balanced between high storage capacities and less computational complexity.

I.4. Transformation of measurements in a diagnostic system

One can view the diagnostic decision making process as a series of transformations or mappings on process measurements [3].

Figure 3 shows the various transformations that a process data go through before the final diagnostic decision is made.

- Measurement space: is a space of measurements $x_1, x_2, \dots, x_N$ with no a priori problem knowledge relating these measurements. These are the input to the diagnostic system.
- Feature space: is a space of points $y = (y_1, y_2, \dots, y_i)$ where y_i is the i^{th} feature obtained as a function of the measurements. These measurements are analyzed and combined using a priori process knowledge to extract useful features about the process behavior to aid diagnosis. The mapping from the feature space to decision space is usually achieved by either using a discriminant function or in some cases using simple threshold functions.
- Decision space: The decision space is a space of points $d = (d_1, d_2, \dots, d_k)$ where k is the number of decision variables, obtained by suitable transformations of the feature space.
- Class space: is a set of integers $c = (c_1, c_2, \dots, c_j)$ where j is the number of failure classes and normal class of data to any of which a given measurements pattern may belong.

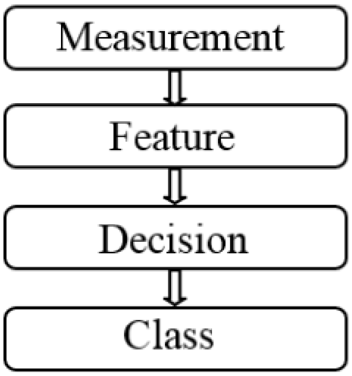

Figure 3. Transformations in a diagnostic system [3].

I.5. Classification of fault detection and diagnosis methods

There is great quantity of literature on fault diagnosis systems ranging from analytical methods to artificial intelligence and statistical approaches. The classification of these fault diagnosis methods very often is not consistent. This is mainly because researchers are often focused on a particular branch [29].

The major difference in fault diagnosis methods is the knowledge used for formulating the diagnostics. At the limits, diagnostics can be based on *a priori* knowledge (e.g., models based entirely on first principles) or driven completely empirically (e.g., by black-box models). Both approaches use models and both use data, but the approach to formulating the diagnostics differs fundamentally [30]. The model-based approach is usually developed based on some fundamental understanding of the physics of the process. It can be classified into quantitative model-based and Qualitative model-based. In quantitative models this understanding is expressed in terms of mathematical functional relationships between the inputs and outputs of the system. In contrast, in qualitative model equations these relationships are expressed in terms of qualitative functions centered on different units in a process [3].

Purely process history approaches (i.e., methods based on black-box models) use no *a priori* knowledge of the process but, instead, derive behavioral models only from measurement data from the process itself. In this latter case, the models may not have any direct physical significance [30].There are different ways in which this data can be transformed and presented as a priori knowledge to a diagnostic system. This is known as the feature extraction process from the process history data, and is done to facilitate later diagnosis. This extraction process can mainly proceed as either quantitative or qualitative feature extraction (Fig. 4) [3].

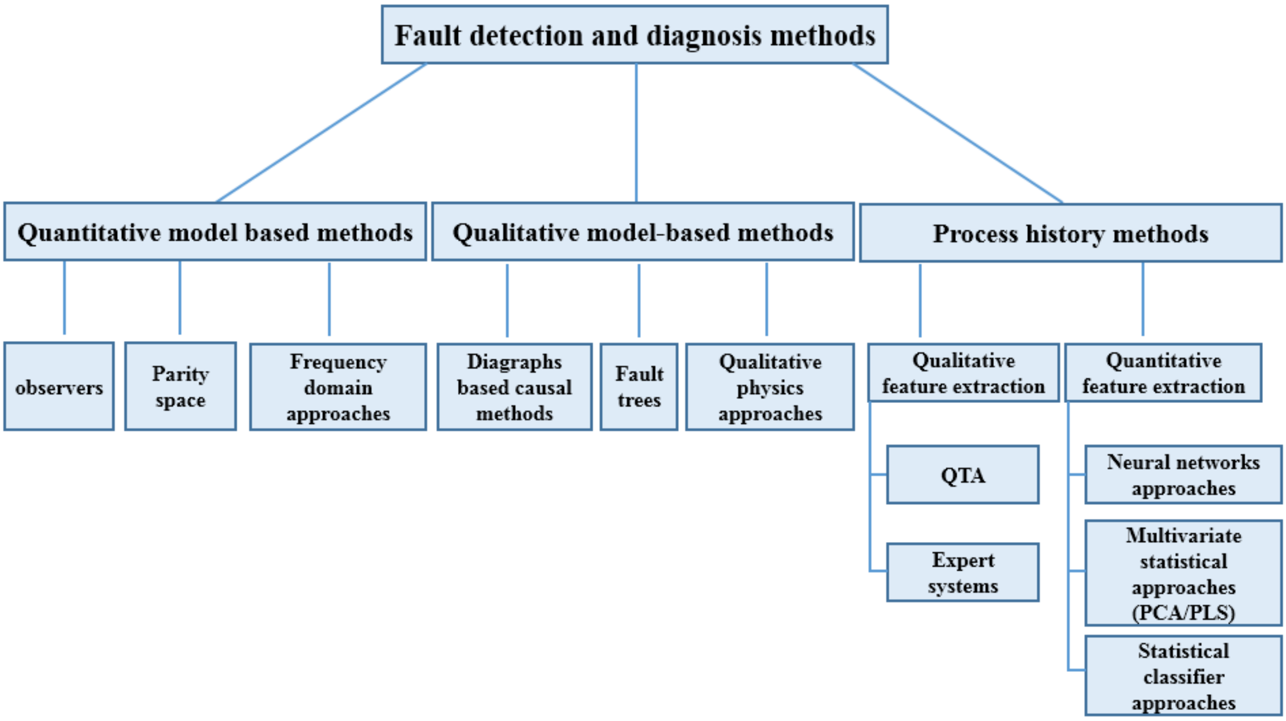

Figure 4. Classification of fault detection and diagnosis methods.

I.5.1. Quantitative model based methods

Model-based methods rely on analytical redundancy by using explicit mathematical models of the monitored process, plant, or system to detect and diagnose faults [30]. The essence of this concept is to check for consistency between the actual outputs of the monitored system and the outputs obtained from a (redundant. i.e. not physical) analytical mathematical model. Therefore, any inconsistency expressed as residuals, can be used for detection and isolation purposes. These residuals should be close to zero when no fault occurs but show ‘significant’ values when the underlying system changes [3].

The quantitative model-based approaches have been based on using general input-output and state space models to generate residuals. These approaches can be classified into observers, parity space and frequency domain approaches.

I.5.1.1. Observers

The main concern of observer-based FDI is the generation of a set of residuals which detect and uniquely identify different faults. The method develops a set of observers, each one of which is sensitive to a subset of faults while insensitive to the remaining faults and the unknown inputs [3, 30].

I.5.1.2. Parity space

The essence of this method is to check the parity (consistency) of the plant models with sensor outputs (measurements) and known process inputs. Under ideal steady state operating conditions, the so-called residual or the value of the parity equations is zero [3].

I.5.1.3. Frequency domain approaches

Residuals are also generated in the frequency domains via factorization of the transfer function of the monitored system.

Quantitative model based methods are often used because of the fact that:

- Models are based on sound physical or engineering principles.
- They provide the most accurate estimators of output when they are well formulated.
- Detailed models based on first principles can model both normal and “faulty” operation; therefore, “faulty” operation can be easily distinguished from normal operation.
- The transients in a dynamic system can only be modeled with detailed physical models [30].

Whereas weaknesses of fault detection and diagnosis methods based on quantitative model-based include:

- They can be complex and computationally intensive.
- The effort required to develop a model is significant.
- These models generally require many inputs to describe the system, some for which values may not be readily available [30].

I.5.2. Qualitative Model-Based Methods

Fault detection and diagnostics based on qualitative modeling techniques represent another broad category that is based on *a priori* knowledge of the system. Unlike quantitative modeling techniques in which knowledge of the system is expressed in terms of quantitative mathematical relationships, qualitative models use qualitative relationships or knowledge bases to draw conclusions regarding the state of a system and its components (e.g., whether operations are “faulty” or “normal”) [3].

Qualitative model-based methods can be classified into:

I.5.2.1. Digraphs based causal models

Causal graphs provide a good way to represent physical cause-effect relations between different process variables that are of interest for fault diagnosis. In the causal directed graph models, the nodes denote the variables, while the directed edges between the nodes represent the causal relations between these variables, through which faults can propagate. The Signed Directed Graph (SDG) method, the simplest causal directed graph method, uses pure qualitative information, which can give rise to ambiguous fault diagnosis [31].

I.5.2.2. Fault Trees

Fault tree analysis (FTA) describes all possible causes of a specified system state in terms of the state of the components within the system. This will be achieved with the use of coherent and non-coherent fault trees. A coherent fault tree is constructed from AND and OR logic, therefore only considers component failed states. The non-coherent method expands this allowing the use of NOT logic which implies that the existence of component failed states and working states are both taken into account [32].

I.5.2.3. Qualitative Physics Approaches

The detailed physical models are based on detailed knowledge of the physical relationships and characteristics of all components in a system. Using this detailed knowledge for mechanical systems, a set of detailed mathematical equations based on mass, momentum, and energy balances along with heat and mass transfer relations are developed and solved. Detailed models can simulate both normal and “faulty” operational states of the system (although modeling of faulty states is not required by all methods).qualitative physics approach is represented in mainly two approaches. The first approach is to derive qualitative equations from the differential equations termed as confluence equations [33]. Considerable work has been done in this area of qualitative modeling of systems and representation of causal knowledge [34].The other approach in qualitative physics is the derivation of qualitative behavior from the Ordinary Differential Equations (ODEs). These qualitative behaviors for different failures can be used as a knowledge source [35].

The main advantages of qualitative model-based methods include:

- They are well suited for data-rich environments and noncritical processes.
- These methods are simple to develop and apply.
- Their reasoning is transparent, and they provide the ability to reason even under uncertainty.
- They possess the ability to provide explanations for the suggested diagnoses because the method relies on cause-effect relationships.
- Some methods provide the ability to perform FDD without precise knowledge of the system and exact numerical values for inputs and parameters [30, 35, 36].

Whereas, Weaknesses of FDD based on qualitative models include:

- The methods are specific to a system or a process.
- Although these methods are easy to develop, it is difficult to ensure that all rules are always applicable and to find a complete set of rules, especially when the system is complex.
- As new rules are added to extend the existing rules or accommodate special circumstances, the simplicity is lost.
- These models, to a large extent, depend on the expertise and knowledge of the developer [30, 35, 36].

I.5.3. Process History (Data-driven) Methods

In process history based methods, only the availability of large amount of historical process data is needed. This data can be transformed and presented as a priori knowledge to a diagnostic system using different ways. And this is known as feature extraction. This extraction process can be either qualitative or quantitative in nature. Two of the major methods that extract qualitative history information are the expert systems and trend modelling methods. Methods that extract quantitative information can be broadly classified as non-statistical or statistical methods [37].

I.5.3.1. Qualitative feature extraction

Two of the major methods that extract qualitative history information are expert systems and Qualitative Trend Analysis [37].

- **Expert systems**

An expert system is generally a very specialized system that solves problems in a narrow domain of expertise. The main components in an expert system development include: knowledge acquisition, choice of knowledge representation, the coding of knowledge in a knowledge base, the development of inference procedures for diagnostic reasoning and the development of input – output interfaces. The main advantages in the development of expert systems for diagnostic problem-solving are ease of development, transparent reasoning, the ability to reason under uncertainty and the ability to provide explanations for the solutions provided [35, 37].

- **Qualitative trend analysis (QTA)**

Trend analysis and prediction are important components of process monitoring and supervisory control. Trend modeling can be used to explain the various important events that happen in a process, to diagnosis malfunctions and to predict future states. From a procedural perspective, in order to obtain a signal trend not too susceptible to momentary variations due to noise, some kind of filtering needs to be employed [35].

I.5.3.2. Quantitative feature extraction

Methods that extract quantitative information can be broadly classified as Non-statistical or statistical methods. Neural networks are an important class of non-statistical classifiers. Principal component analysis (PCA)/partial least squares (PLS) and statistical pattern classifiers form a major component of the statistical feature extraction methods [35, 37].

- **Multivariate statistical approaches:**

Multivariate statistical techniques are powerful tools capable of compressing data and reducing its dimensionality so that essential information is retained and easier to analyze than the original huge data set; and they are able to handle noise and correlation to extract true information effectively. Multivariate statistical process control methods, such as Principal Component Analysis (PCA) and Partial Least Squares (PLS), have been used in process monitoring problems. These are based on transforming a set of highly correlated variables to a set of uncorrelated variables [38, 39]. Principal component analysis (PCA) probably the most popular among these techniques [9, 40]. PCA is capable of compressing high-dimensional data with little loss of information by projecting the data onto a lower-dimensional subspace defined by a new set of derived variables (principal components (PCs)) [11].

- **Statistical classifier approaches**

Fault diagnosis is essentially a classification problem and hence can be cast in a classical statistical pattern recognition framework. Fault diagnosis can be considered as a problem of combining, over time, the instantaneous estimates of the classifier using knowledge about the statistical properties of the failure modes of the system [35, 41, 42].

- **Neural network approach**

Neural networks have been proposed for classification and function approximation problems. In general, neural networks that have been used for fault diagnosis can be classified along two dimensions: (i) the architecture of the network such as sigmoidal and radial basis (ii) The learning strategy such as supervised and unsupervised learning [37].

Fault detection and diagnosis methods based on process history are well suited to problems for which theoretical models of behavior are poorly developed or inadequate to explain observed performance and where training data are plentiful or inexpensive to create or collect. This approach provides black-box models, which are easy to develop and do not require an understanding of the physics of the system being modeled with a generally manageable computational requirement.

Beside all the advantages listed earlier, the most significant drawbacks is that most of the models cannot be used to extrapolate beyond the range of the training data and a large amount of training data is needed, representing both normal and faulty operation. The models are specific to the system for which they are trained and rarely can be used on other systems. Process data-based methods are suitable where no other methods exist. Some are applicable for virtually any kind of pattern recognition problems [30, 35].

I.6. A comparison of various approaches

In the previous section, we have reviewed the three conceptually different frameworks for process fault diagnosis (Quantitative model-based, Qualitative model-based and Process history). In this section, we provide a comparative evaluation of these different frameworks against a common set of desirable characteristics for a diagnostic system which are proposed in section 3. The evaluations are summarized in Table 1.

Table 01. Comparison of diagnosis methods [37].

	Observer	Digraphs	Expert systems	QTA	PCA	Neural networks
Quick detection and diagnosis	√	?	√	√	√	√
Isolability	√	x	√	√	√	√
Robustness	√	√	x	√	√	√
Novelty identifiability	?	√	x	?	√	√
Classification error	x	x	x	x	x	x
Adaptability	x	√	√	?	x	x
Explanation facility	x	√	√	√	x	x
Modelling requirement	?	√	√	√	√	√
Storage and computation	√	?	√	√	√	√
Multiple fault identifiability	√	√	x	x	x	x

√**: suitable; x: not suitable; ? : not assessed**

I.7. Statistical Process Monitoring (SPM)

Statistical Process monitoring is an analytical decision making tool which allows to see when a process is working correctly and when it is not. Variation is present in any process, deciding when the variation is natural and when it needs correction is the key to quality control.

Control chart is a major tool of SPM. It is a graphical representation for the process quality characteristics, that has been measured or computed from an acquired data, and which determines whether the process is under statistical control (healthy state) or not.

Control chart consists of two parts (Fig.5): the first is a series of measurements plotted in time order. Whereas the second is the control limits which is represented usually by three horizontal lines: the center line (CL) (typically, the mean); the upper control limit (UCL) and the lower control limit (LCL). They are used to define the range within which any variability is considered to be due to chance causes (natural variability) only. Consequently, point that lies beyond the control limits is treated because of an assignable cause (fault) [43].

The control limits are defined based on the empirical distribution of the monitored quality. Therefore, the control limits are determined with a given confidence level; i.e., we select Control Limits to ensures that $1 - \alpha$ percent of the points expressing chance causes are lying within the control limits.

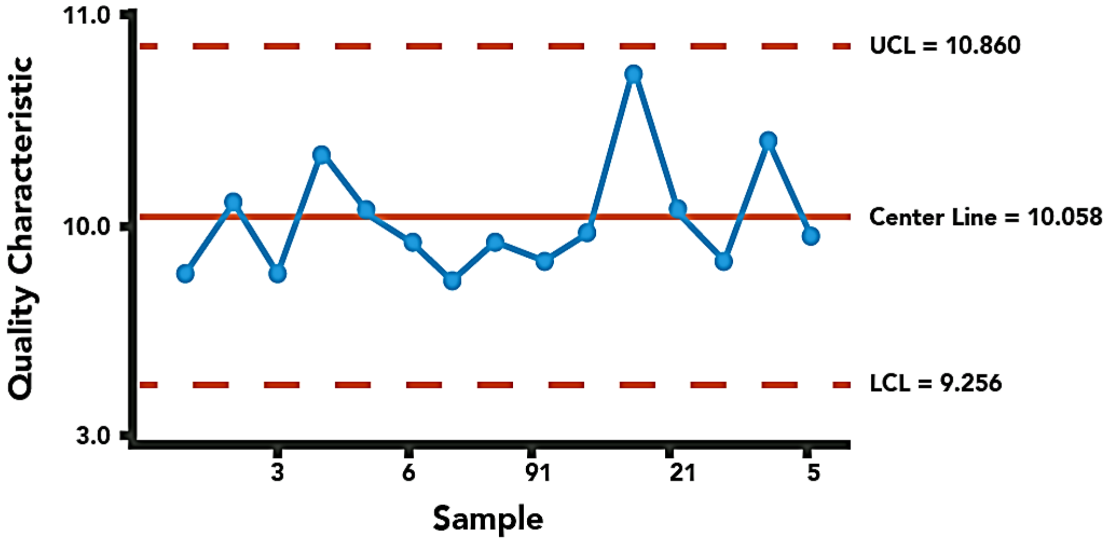

Figure 5. Control chart for SPC.

I.8. Conclusion

The basic aim of this chapter is to give some definitions and terminologies used in the field of fault detection and diagnosis and to review various methods of fault detection from different perspectives. We have also compared these methods against a common set of desirable characteristics for a diagnostic system that we proposed in section 3. A definition of SPC was investigated in order to achieve a background in the field. Due to the vastness of FDD field, a deep investigation for each method is a time expensive process. In the next chapters, we will focus mainly on multivariate statistical approaches.

Chapter II.

Multivariate statistical approaches

II.1. Introduction

Methods based on historical process data are seen as the most cost-effective approaches dealing with the complex systems [4]. Traditionally, data-driven techniques based on statistical process control (SPC) charts such as Shewhart, CUSUM and EWMA charts have been used to monitor processes and improve product quality. However, such univariate control charts show poor fault detection performance when applied to multivariate processes. This shortcoming of univariate control charts has led to the development of many process monitoring schemes that use multivariate statistical methods [26].

When the assumptions of linearity hold, multivariate statistical process control based on the use of principal component analysis can be used very effectively for early detection and analysis of any abnormal plant behavior [4, 10]. However, the PCA identifies only linear structure in a given dataset, as it is nothing but a linear projection [15, 21, 27].

In order to extend this technique to deal with nonlinear structures, many studies have been proposed to define nonlinear extensions of PCA. The kernel principal component analysis (KPCA) is among the most popular nonlinear statistical methods [18, 19, 20] .It can efficiently compute principal components (PCs) in high-dimensional feature spaces by means of integral operators and nonlinear kernel functions. Despite recently reported KPCA-based monitoring applications, KPCA monitoring model requires the storage of the symmetric kernel matrix (computation time may increase with the number of samples)[27]. RKPCA deals with the problem of need storage and computation time by approximating the data matrix by a reduced one.

This chapter is organized as follow: theoretical background of PCA method as a multivariate statistical tool for process monitoring is presented and a simple calculation of the statistic indices The Hotelling T^2, Q and the combined is provided in section 2. In section 3, KPCA for fault detection is presented with its Theorical literature. In section 4, the proposed reduced KPCA method for process monitoring is detailed.

II.2. Principal Component Analysis (PCA)

Principal component analysis (PCA), is a statistical technique that was proposed by Karl Pearson in 1901 and developed by Hotelling in 1947. Its aim is to transform a set of correlated original data to an uncorrelated data set that represents most of the information of the original data (Fig.6) [37].

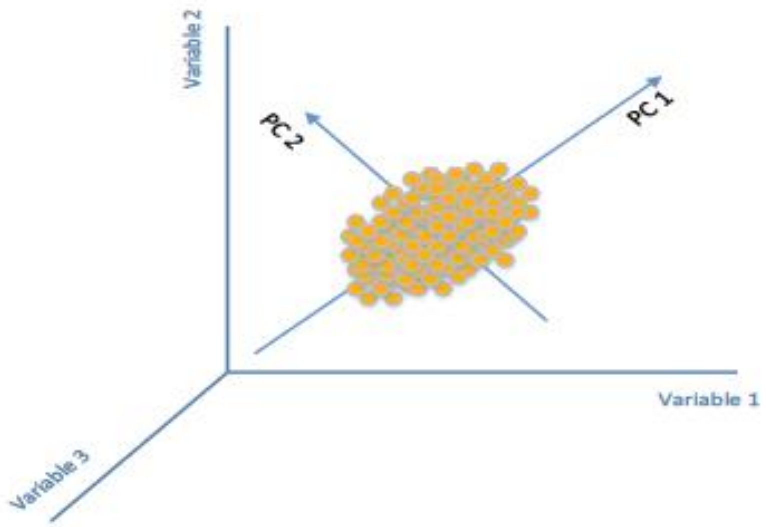

Figure 6. Dimension redaction from 3D to 2D using PCA.

II.2.1. Statistical process modeling using PCA

Let X_0 , denotes the original data matrix with m process physical variables and n samples, $X_0 \epsilon R^{n\times m}$. This data matrix is first normalized to a matrix $X\epsilon R^{n\times m}$ with: (i) zero mean: to simplify the computation of the covariance matrix and (ii) unit variance: to unify the units of different the measurements (Pressure (Pa), Voltage (V), Temperature (°C), etc.) [37, 44] and whose covariance matrix is [30]:

$$\boldsymbol{S} = \frac{\boldsymbol{1}}{\boldsymbol{n-1}} \boldsymbol{X^T X} \quad (2.1)$$

Based on Singular Value Decomposition technique S can be written as:

$$\boldsymbol{S} = \boldsymbol{P \Lambda P^T} \quad (2.2)$$

Where P is an $m \times m$ matrix whose columns are known as principal component loading vectors. And the eigenvalues $(\lambda_1, \lambda_2, \lambda_3 \dots, \lambda_m)$ of S are the elements decreasingly ordered of diagonal matrix Λ.

The mean centered and scaled matrix measurement X is then given by:

$$\boldsymbol{X} = \boldsymbol{T.P^T} \quad (2.3)$$

Where $T \epsilon R^{n\times m}$ denotes the principal component scores and can be written as:

$$\boldsymbol{T} = \boldsymbol{X.P} \quad (2.4)$$

By definition, PCA method represents the data by fewer sufficient components. Thus, using $\ell < m$ of the components, one can obtain ℓ -dimensional scores by the following relationship.

$$\widehat{\boldsymbol{T}}_{\boldsymbol{n}} = \boldsymbol{X}.\boldsymbol{P}_{\boldsymbol{\ell}} \quad (\mathbf{2.5})$$

Where P_ℓ contains only the ℓ first columns of P. ℓ Represents the number of retained principal component (PC's).

Once retaining ℓ principal components, the decomposition of the used matrices becomes:

$$\boldsymbol{P} = \left[\widehat{\boldsymbol{P}}_{\boldsymbol{m}\times\boldsymbol{\ell}} \quad \widetilde{\boldsymbol{P}}_{\boldsymbol{m}\times(\boldsymbol{m}-\boldsymbol{\ell})}\right] \quad (\mathbf{2.6})$$

$$\boldsymbol{T} = \left[\widehat{\boldsymbol{T}}_{\boldsymbol{n}\times\boldsymbol{\ell}} \quad \widetilde{\boldsymbol{T}}_{\boldsymbol{n}\times(\boldsymbol{m}-\boldsymbol{\ell})}\right] \quad (\mathbf{2.7})$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} \widehat{\boldsymbol{\Lambda}}_{\boldsymbol{\ell}\times\boldsymbol{\ell}} & \mathbf{0}_{\boldsymbol{\ell}\times(\boldsymbol{m}-\boldsymbol{\ell})} \\ \mathbf{0}_{(\boldsymbol{m}-\boldsymbol{\ell})\times\boldsymbol{\ell}} & \widetilde{\boldsymbol{\Lambda}}_{(\boldsymbol{m}-\boldsymbol{\ell})\times(\boldsymbol{m}-\boldsymbol{\ell})} \end{bmatrix} \quad (\mathbf{2.8})$$

The data matrix X can be decomposed as:

$$\boldsymbol{X} = \boldsymbol{X}\widehat{\boldsymbol{P}}\widehat{\boldsymbol{P}}^{\boldsymbol{T}} + \boldsymbol{X}\widetilde{\boldsymbol{P}}\widetilde{\boldsymbol{P}}^{\boldsymbol{T}} = \boldsymbol{X}\boldsymbol{C} + \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{C}) = \widehat{\boldsymbol{X}} + \boldsymbol{E} \quad (\mathbf{2.9})$$

Where $\hat{X}$ and E are respectively the modeled and the non-modeled variations of X by projection onto the principal component subspace and the residual subspace. $C = \hat{P}\hat{P}^T$ and $I - C = \tilde{P}\tilde{P}^T$ are projection matrices that provides the linear combinations with large and low variations respectively.

When a new measurement $x \epsilon R^{1\times m}$ is observed, it can be decomposed into two parts using PCA model to:

$$\boldsymbol{x} = \widehat{\boldsymbol{x}} + \boldsymbol{e} \quad (\mathbf{2.10})$$

Where $\hat{x} = xC$ and $e = I - C$ are the projection of the new observed sample onto the principal component subspace and the residual subspace respectively [45].

II.2.2. Model dimension selection

Numerous methods exist for selecting the number of PCs when the PCA technique is used for fault detection. Some of the popular methods are: cross validation method [46], parallel analysis [47], revised parallel analysis [48], Kaiser–Guttman method [49] and

Cumulative Percent Variance (CPV) [50]. In the cumulative percent variance (CPV) method, the minimum model dimension that can express a substantial part of the total variance of the data is selected.

The cross-validation method uses part of the training samples for model construction; the remaining samples are compared with the prediction by the model and when the prediction residual sum of squares (PRESS) is less than the residual sum of squares of the previous model, the new component is added to the model.

CPV is the most common used in fault detection [45]. It retains ℓ principal components having their sum of variances greater than a certain percentage of the total variance (usually taken from 70% to 90%). It determines the quality of the constructed model based on the chosen CPV, and it is related to all other approaches since retaining any number of components will result in a certain captured percentage of total variance.

$$\boldsymbol{CPV}(\boldsymbol{\ell}) = \frac{\sum_{i=1}^{i=\ell} \lambda_i}{\sum_{i=1}^{i=m} \lambda_i}\ \% \qquad (\mathbf{2.11})$$

Where λ_i the variance of the score is vector and ℓ is the number of PCs that are retained. When CPV is larger than the selected percentage (taken from 70% to 90%), the corresponding number ℓ of PCs is determined.

In this report herein, the cumulative percent variance (CPV) is utilized to compute the optimum number of retained PCs.

II.2.3. Fault detection using PCA

After obtaining a lower-dimensional representation, new observations are mapped into the associated space and are generally monitored in terms of two statistics: Q and T^2. In this work, φ-Statistic is introduced as a third monitoring index.

The Q statistic measures the goodness-of-fit of the lower dimensional model. Abnormal values of Q can indicate either an unusual observation or that the lower-dimensional model does not account for some of the process variability. The T^2 statistic measures deviations in the lower-dimensional space. Because this space should characterize the major components of the process under normal operation condition (NOC), outliers identified by large T^2 values can indicate abnormal conditions [51].

The φ-Statistic represents the combination of the two previous indices and it monitors the whole measurement space [52]. The three statistics and their corresponding control limits are given as follows:

II.2.3.1. Q -Statistic

The Q statistic, sometimes called the Squared Prediction Error (SPE), represents the variability in the residual subspace (Fig. 07). So it measures the lack of fit of the data to the PCA model [45], it is given by:

$$\boldsymbol{Q = x(I - C)(I - C)^T x^T = \|e\|^2} \qquad (\mathbf{2.12})$$

The distribution of the Q-statistic can be approximated as:

$$\boldsymbol{Q_\alpha = \left[\frac{h_0 C_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2}\right]^{1/h_0}} \qquad (\mathbf{2.13})$$

Where $\theta_i = \sum_{j=\ell+1}^{j=m} \lambda_j^i$, i = 1,2,3 and $h_0 = 1 - \frac{2\theta_1\theta_2}{3\theta_2^2}$

C_α is the critical value for an appropriately chosen confidence level $(1-\alpha)$ Of the standard normal distribution.

The threshold Q_α is applied to define the normal variations of the random noise, and any violation of the threshold can indicate that the random noise has significantly changed, hence this is used to detect faults [45].

II.2.3.2. T^2-Statistic

T^2also called the Hotelling's T^2 is another index used for fault detection when PCA is used simultaneously for both dimensional reduction and detecting faults. It represents variability within the principle components subspace and it detects variation that can be greater than what can be explained by the PCA model (Fig.07), and it is defined as:

$$\boldsymbol{T^2 = x\,\widehat{P}\widehat{\Lambda}^{-1}\,\widehat{P}^T\,x^T} \qquad (\mathbf{2.14})$$

The process is considered under normal operation, if T^2 is below an upper control limit defined for a properly chosen significance level α as:

$$\boldsymbol{T_\alpha = \frac{(n^2-1)\ell}{n(n-\ell)} F_\alpha(\ell, n-\ell)} \qquad (\mathbf{2.15})$$

Where $F_\alpha(\ell, n-\ell)$ is the critical value of the Fisher–Snedecor distribution with ℓ and $(n-\ell)$ degrees of freedom [45].

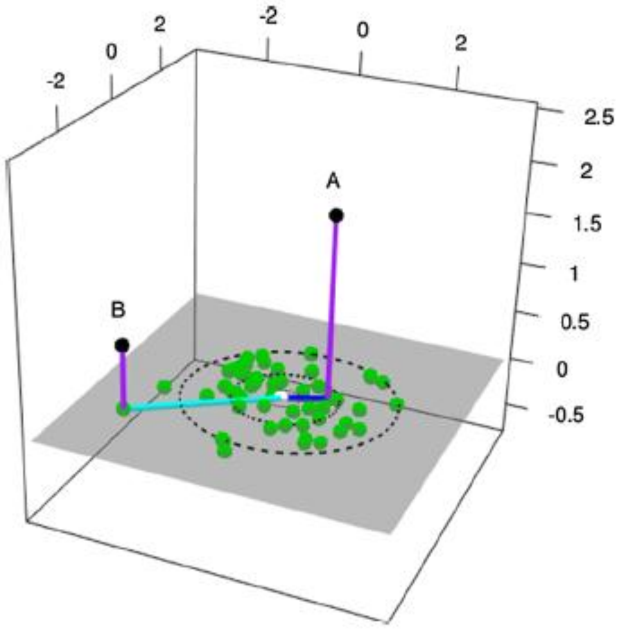

Figure 07. A set of projected observations (green) and two original observations **(A, B)** [21]. **(A)** has a large SPE (squared vertical distance) but small T^2 (squared horizontal distance) relative to **(B).**

II.2.3.3. φ-Statistic

Yue and Qin proposed the use of a combined index for monitoring the principal and residual space simultaneously. Such an index is a combination of the T^2 and SPE indices weighted by their control limits [25]. φ-Statistic gives informations about the variability in the whole measurement space. For a new measurement vector x , φ is defined as:

$$\boldsymbol{\varphi} = \frac{\boldsymbol{T}^2}{\boldsymbol{T}^2{}_{\alpha}} + \frac{\boldsymbol{Q}}{\boldsymbol{Q}_{\alpha}} \quad (2.16)$$

We use the approximate distribution to calculate the confidence limits of the combined index. The distribution of φ can be approximated using $g\chi^2(h)$ where the parameters g and h are given by

$$g = \frac{\frac{\ell}{(T_{\alpha}^2)^4} + \sum_{i=\ell+1}^{N} \frac{\lambda_i^2}{Q_{\alpha}^4}}{(N-1)\left(\frac{\ell}{(T_{\alpha}^2)^2} + \sum_{i=\ell+1}^{N} \frac{\lambda_i}{Q_{\alpha}^2}\right)} \quad (2.17)$$

$$h = \frac{\left(\frac{\ell}{(T_{\alpha}^2)^2} + \sum_{i=\ell+1}^{N} \frac{\lambda_i}{Q_{\alpha}^2}\right)^2}{\frac{\ell}{(T_{\alpha}^2)^4} + \sum_{i=\ell+1}^{N} \frac{\lambda_i^2}{Q_{\alpha}^4}} \quad (2.18)$$

Table 02. PCA-based fault detection algorithm [45].

PCA-based fault detection algorithm
Offline monitoring
Obtain training fault free data set that represent the normal operations.
Scale the data to zero mean and unit variance.
Compute the covariance matrix S using **(2.1)**
Calculate the eigenvectors and eigenvalues of S.
Determine how many principal components to be used. Many techniques can be used in this regard. In this work, the CPV criterion is used **(2.11).**
Calculate $Q_{\alpha}, T_{\alpha}^2, \varphi_{\alpha}$ using equations **(2.13)**, **(2.15)** and **(2.17; 2.18)** respectively.
Online monitoring
Obtain testing data (possibly faulty data).
Scale the data using mean and variance of the training set.
Calculate T^2, Q, φ.
Check for faults: if $T^2 > T_{\alpha}^2$ or $Q > Q_{\alpha}$ or $\varphi > \varphi_{\alpha}$, then declare a fault.

II.2.4. PCA main drawbacks

The selection of the optimal number of principal components (PCs) in fault detection using principal components analysis (PCA) is a critical and sensitive operation because overestimating the number of components will results in a contamination of the extracted informations by adding noise dimensions with no useful informations. This will lead to an important amount of false alarms. In the other hand, underestimating the number of components results in a loss of information [44].

Limitations of the PCA approach include also its lack of exploitation of autocorrelation and its linear nature [37]. The minor principal components would normally represent insignificant variance in the data for the linear case, but this cannot be said with certainty for nonlinear data. To confidently represent a nonlinear data set, more principal components have to be retained. This increases computational requirements. It is also difficult to discern which minor components capture nonlinearity and which represent insignificant variation [12].

In the next section, a nonlinear version of PCA called Kernel principal is presented to dealing with nonlinear systems that cannot be accommodated adequately by linear multivariate methods.

II.3. Kernel Principal Component Analysis (KPCA)

For some complicated cases in industrial processes with particular nonlinear characteristics, the conventional PCA performs poorly due to its assumption that the process data are linear. To overcome this limitation, several nonlinear extensions of PCA were reported. The kernel principal component analysis is considered among the most popular nonlinear statistical methods [53] developed recently [18, 26].

Using a nonlinear mapping, KPCA first maps the original process data into a high dimensional feature space where the data structure is more likely to be linear (Fig. 08) [54]. Linear PCA is then conducted on this feature space, and the resulting principal components are able to capture nonlinearities in the original data space.

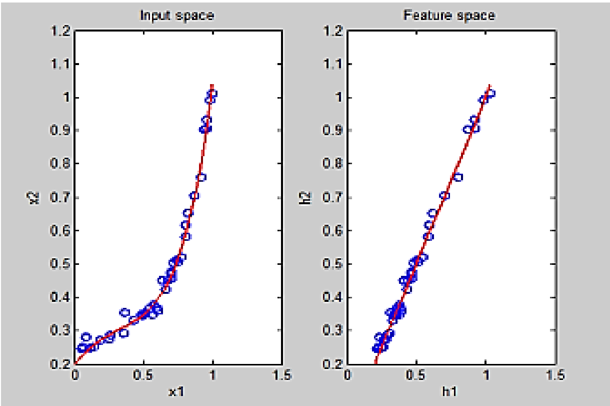

Figure 8. The principle of Kernal PCA [26].

II.3.1. Statistical process modeling using KPCA

Let $X = [X_1 \quad X_2 \quad \cdots \quad X_N]^T$; the training set scaled to zero mean and unit variance, $X_i \epsilon R^m$, $i = 1, \dots, N$ with N observations and m process variable. By a nonlinear mapping ϕ, a measured input is projected into a hyper-dimensional feature space F as:

$$\boldsymbol{\phi}: \boldsymbol{X_i} \; \boldsymbol{\epsilon} \; \boldsymbol{R^m} \; \rightarrow \boldsymbol{\phi}(\boldsymbol{X_i}) = \boldsymbol{\phi_i} \; \boldsymbol{\epsilon} \; \boldsymbol{F} \qquad (\mathbf{2.19})$$

Note that the feature space F have an arbitrarily large, possibly infinite dimensionality equal to h [18].

KPCA seeks to decouple the nonlinear correlations among the data set through diagonalizing its covariance matrix [26], which can be expressed in the linear feature space F as:

$$\boldsymbol{C_\phi} = \frac{\mathbf{1}}{\boldsymbol{N}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{\phi}(\boldsymbol{X_i}) \boldsymbol{\phi}(\boldsymbol{X_i})^{\boldsymbol{T}} = \frac{\mathbf{1}}{\boldsymbol{N}} \boldsymbol{\chi} \boldsymbol{\chi}^{\boldsymbol{T}} \; \boldsymbol{\epsilon} \; \boldsymbol{R^{h \times h}} \qquad (\mathbf{2.20})$$

With $\chi = [\phi_1 \quad \phi_2 \quad \cdots \quad \phi_N]^T$ ϕ_i is the training data arranged in the feature space F.

To diagonalize the covariance matrix, the following eigenvalue problem in the feature space is solved,

$$\boldsymbol{\lambda v = C_{\phi} v} \qquad (\mathbf{2.21})$$

Where $\lambda \geq 0$ are eigenvalues and $v \, \epsilon \, F \backslash \{0\}$ their corresponding eigenvectors. The importance of the eigenvectors is indicated by the magnitude of their corresponding eigenvalues. By combining equation **(2.20)** and **(2.21),** we get:

$$\boldsymbol{C_{\phi} v = \frac{1}{N} \sum_{i=1}^{N} \phi(X_i) \phi(X_i)^T v}$$

$$\boldsymbol{= \frac{1}{N} \sum_{i=1}^{N} \langle \phi(X_i), v \rangle \ \phi(X_i)} \qquad (\mathbf{2.22})$$

Where $\langle X, Y \rangle$ represents the dot product between X and Y. This implies that all solutions v with $\lambda \neq 0$ lie in the span of $\phi(x_1), \dots, \phi(x_N)$, so that,

$$\boldsymbol{\lambda \langle \phi(X_k), v \rangle = \langle \phi(X_k), C_{\phi} v \rangle \qquad k = 1, \dots, N} \qquad (\mathbf{2.23})$$

Therefore, there must exist coefficients α_i , $i = 1, \dots, N$ such that every eigenvector v of C_{ϕ} can be linearly expanded by:

$$\boldsymbol{v = \sum_{i=1}^{N} \alpha_i \, \phi(X_i)} \qquad (\mathbf{2.24})$$

Combining equation **(2.23)** and **(2.24)**, we get **(2.25)**:

$$\boldsymbol{\lambda \sum_{i=1}^{N} \alpha_i \langle \phi(X_k), \phi(X_i) \rangle = \frac{1}{N} \sum_{i=1}^{N} \alpha_i \, \langle \phi(X_i), \sum_{j=1}^{N} \phi(X_j) \rangle \times \langle \phi(X_{ij}), \phi(X_i) \rangle} \, ,$$

$$\boldsymbol{k = 1, \dots, N} \qquad (\mathbf{2.25})$$

Note that the eigenvalue problem shown in equation **(2.25)** only involves dot products of mapped vectors in the feature space. The mapping need not be explicitly computed and only the dot products of two vectors in the feature space are needed [26, 27].

Now defining a matrix $K \epsilon \, R^{N \times N}$ by:

$$\boldsymbol{[K]_{ij} = K_{ij} = k(x_i, x_j) = \langle \phi(X_i), \phi(X_j) \rangle} \qquad (\mathbf{2.26})$$

The left hand side of equation **(2.25)** can be expressed as:

$$\boldsymbol{\lambda \sum_{i=1}^{N} \alpha_i \langle \phi(X_k), \phi(X_i) \rangle = \lambda \sum_{i=1}^{N} \alpha_i K_{ki}} \qquad (\mathbf{2.27})$$

And the right hand side of equation **(2.25)** can be given by:

$$\frac{1}{N}\sum_{i=1}^{N}\alpha_i\,\langle\phi(X_i),\sum_{j=1}^{N}\phi(X_j)\rangle\times\langle\phi(X_{ij}),\phi(X_i)\rangle=\frac{1}{N}\sum_{i=1}^{N}\alpha_i\sum_{j=1}^{N}K_{ij}K_{ji}\quad(2.28)$$

Combining equations **(2.26)** and **(2.27)**, we get:

$$\lambda NK\alpha=K^2\alpha \qquad (2.29)$$

Where $\alpha=[\alpha_1 \quad \cdots \quad \alpha_N]^T$.To get solutions of equation **(2.29)**, we solve the following eigenvalue problem,

$$N\lambda\alpha=K\alpha \qquad (2.30)$$

A justification of this procedure is given by Schölkopf in [18].

Now, performing PCA in the feature space F is equivalent to solving the eigenvalue problem shown in equation **(2. 30)**. This yields eigenvectors $\alpha_1 \ldots \alpha_N$ with eigenvalues $\lambda_1 \ldots \lambda_N$ [26].The dimensionality of the problem can be reduced by retaining only the first ℓ eigenvectors using the cumulative percent variance method given in equation **(2.11)**. The eigenvector $\alpha_1 \ldots \alpha_N$ can be normalized by normalizing the corresponding eigenvectors in the feature space F, i.e,[27]

$$\langle v_k,v_k\rangle=1 \quad k=1,\ldots,\ell \qquad (2.31)$$

Using $v_k=\sum_{i=1}^{N}\alpha_i^k\phi(X_i)$ equation **(2.31)** leads to,

$$1=\langle\sum_{i=1}^{N}\alpha_i^k\phi(X_i)\;,\sum_{i=j}^{N}\alpha_j^k\phi(X_j)\;\rangle$$

$$=\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i^k\alpha_j^k\langle\phi(X_i),\phi(X_j)\rangle$$

$$=\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i^k\alpha_j^kK_{ij}$$

$$=\langle\alpha_k,K\alpha_k\rangle$$

$$=\lambda_k\langle\alpha_k,\alpha_k\rangle \qquad (2.32)$$

Thus, the associated eigenvectors $\alpha_1,\ldots,\alpha_N$ can be expressed

$$\langle\alpha_k,\alpha_k\rangle=\frac{1}{\lambda_k} \qquad k=1,\ldots,N \qquad (2.33)$$

Which shows that $v_1,\ldots,v_N$ are given by

$$v_i=\sum_{j=1}^{N}\frac{\alpha_j^i}{\sqrt{\lambda_i}}\phi_j=\frac{1}{\sqrt{\lambda_i}}\chi^T\alpha^i \qquad (2.34)$$

We denote $\widehat{P_f} = [v_1, \dots, v_\ell]$ the principal subspace that spans the maximal variance between data. And $\widetilde{P_f} = [v_{\ell+1}, \dots, v_N]$ the residual subspace that contains the noises [55]. Using equation **(2.34)**, $\widehat{P_f}$ and $\widetilde{P_f}$ can be expressed as:

$$\widehat{\boldsymbol{P_f}} = \left[\frac{\mathbf{1}}{\sqrt{\boldsymbol{\lambda_1}}}\boldsymbol{\chi^T\alpha^1} \quad \dots \quad \frac{\mathbf{1}}{\sqrt{\boldsymbol{\lambda_\ell}}}\boldsymbol{\chi^T\alpha^\ell}\right] = \boldsymbol{\chi^T\widehat{P}\ \widehat{\Lambda}^{-1/2}} \qquad (\mathbf{2.35})$$

$$\widetilde{\boldsymbol{P_f}} = \left[\frac{\mathbf{1}}{\sqrt{\boldsymbol{\lambda_{\ell+1}}}}\boldsymbol{\chi^T\alpha^{\ell+1}} \quad \dots \quad \frac{\mathbf{1}}{\sqrt{\boldsymbol{\lambda_N}}}\boldsymbol{\chi^T\alpha^N}\right] \qquad (\mathbf{2.36})$$

Where $\hat{P} = [\alpha^1 \quad \dots \quad \alpha^\ell]$ and $\widehat{\Lambda} = diag(\lambda^1 \quad \dots \quad \lambda^\ell)$ are the ℓ principal eigenvectors and eigenvalues of K respectively, corresponding to the largest eigenvalues in descending order.

By projecting $\phi(X)$ onto eigenvectors in the feature space F, we get,

$$\boldsymbol{t_k} = \langle \boldsymbol{v_k}, \boldsymbol{\phi}(\boldsymbol{X})\rangle = \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{N}} \boldsymbol{\alpha_i^k}\langle \boldsymbol{\phi}(\boldsymbol{X_i}), \boldsymbol{\phi}(\boldsymbol{X})\rangle, \quad \boldsymbol{k} = \mathbf{1}, \dots, \boldsymbol{N} \qquad (\mathbf{2.37})$$

The projection on the principal and residual spaces respectively are given by:

$$\hat{\boldsymbol{t}} = \widehat{\boldsymbol{P_f}}^{\boldsymbol{T}}\boldsymbol{\phi} \quad \boldsymbol{\epsilon}\ \boldsymbol{R^\ell} \qquad (\mathbf{2.38})$$

$$\tilde{\boldsymbol{t}} = \widetilde{\boldsymbol{P_f}}^{\boldsymbol{T}}\boldsymbol{\phi} \quad \boldsymbol{\epsilon}\ \boldsymbol{R^{h-\ell}} \qquad (\mathbf{2.39})$$

When solving the eigenvalue problem given in equation **(2. 30)** or projecting from the input space into the KPCA space using equation **(2.37)**, we can avoid computing the nonlinear mappings or the dot products in the feature space by introducing a *Kernel function* of the form $K(X, Y) = \langle\phi(X), \phi(Y)\rangle$ [18,20]. There exist a number of kernel functions:

- Polynomial Kernel: $K(X, Y) = \langle X, Y\rangle^d$
- Sigmoid Kernel: $K(X, Y) = \tanh(\beta_0 \langle X, Y\rangle + \beta_1)$
- Radial basis Kernel: $K(X, Y) = \exp(\frac{-\|X-Y\|^2}{C})$

Where β_0, β_1 and C are specified a priori by the user [54, 56].

These kernel functions provide a low-dimensional KPCA subspace that represents the distributions of the mapping of the training vectors in the feature space [55]. A specific selection of the kernel function implicitly determines the mapping ϕ and the feature space$F$.

Before applying KPCA, mean centering in the high-dimensional space must be performed [27]. This can be done by substituting the kernel matrix K with,

$$\widetilde{\boldsymbol{K}} = \boldsymbol{K} - \mathbf{1}_{\boldsymbol{N}}\boldsymbol{K} - \boldsymbol{K}\mathbf{1}_{\boldsymbol{N}} + \mathbf{1}_{\boldsymbol{N}}\boldsymbol{K}\mathbf{1}_{\boldsymbol{N}} \qquad (\mathbf{2.40})$$

Where, $1_N = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \epsilon R^{N \times N}$

II.3.2. Fault detection using KPCA

The KPCA-based monitoring method is similar to that using PCA in that the three indices: Q, The Hotelling's T^2 and the combined and can be interpreted in the same way in the feature space [21].

II.3.2.1. Q-statistic

Q-statistic or the SPE index is defined as the norm of the residual vector in the feature space. It is calculated as the squared norm of the residual components [27]

$$\boldsymbol{SPE = \tilde{t}^T \tilde{t} = \phi^T \widetilde{P_f} \widetilde{P_f}^T \phi = \phi^T \widehat{C}_f \, \phi} \quad (2.41)$$

Since we do not know the dimension of the feature space, it is not possible to know the number of residual components there. Thus, we cannot calculate explicitly the loading matrix $\widetilde{P_f}$. However, we can calculate the product $\widetilde{P_f}\widetilde{P_f}^T$ as the orthogonal projection to the principal component space [27], which is

$$\boldsymbol{\widehat{C}_f = \widetilde{P_f} \widetilde{P_f}^T = I_N - \widehat{P_f} \widehat{P}_f^{\,T}} \quad (2.42)$$

And leads to

$$\boldsymbol{SPE = \phi^T (I_N - \widehat{P_f} \widehat{P}_f^{\,T}) \phi = \phi^T \phi - \phi^T \widehat{P_f} \widehat{P}_f^{\,T} \phi} \quad (2.43)$$

Using equations **(2.26)** and **(2.35)** The SPE index can be calculated as a function of input vector x as

$$\begin{aligned} \boldsymbol{SPE} &= \boldsymbol{k(x,x) - \phi^T \chi^T \widehat{P} \widehat{\Lambda}^{-1} \widehat{P}^T \chi \phi} \\ &= \boldsymbol{k(x,x) - k(x)^T \widehat{P} \widehat{\Lambda}^{-1} \widehat{P}^T k(x)} \\ &= \boldsymbol{k(x,x) - k(x)^T \widehat{C} k(x)} \end{aligned} \quad (2.44)$$

Where $\hat{C} = \hat{P} \widehat{\Lambda}^{-1} \hat{P}^T$.

The threshold developed for linear PCA in equation **(2. 13)** can be directly used for KPCA.

II.3.2.2. T^2-statistic

The Hotelling's T^2 index is calculated in the feature space as [27]

$$\boldsymbol{T^2 = \hat{t}^T \widehat{\Lambda}^{-1} \hat{t}} \quad (2.45)$$

The T^2 is calculated using kernel functions as

$$\boldsymbol{T^2 = k(x)^T \widehat{P} \, \widehat{\Lambda}^{-2} \widehat{P}^T k(x)}$$

$$= \boldsymbol{k}(\boldsymbol{x})^{\boldsymbol{T}} \boldsymbol{D} \boldsymbol{k}(\boldsymbol{x}) \qquad (\mathbf{2.46})$$

Where $D = \hat{P}\,\hat{\Lambda}^{-2}\hat{P}^{T}$.The $(1-\alpha)\times 100\%$ control limit for T² is calculated using F distribution as in conventional PCA given by equation **(2.15)**.

III.3.2.3. $\boldsymbol{\varphi}$-Statistic

Similarly to PCA, φ-Statistic incorporates the SPE and T² in a balance way which is given by equation **(2.16)**. And its control limits can be approximated by $g\chi^2(h)$ with g and h are parameters given by equation **(2.17)** and **(2.18)** [25].

Table 03. KPCA-based fault detection algorithm [57].

Algorithm 2: KPCA-based fault detection
1. Developing the normal operating condition (NOC) model - Acquire normal operating data and normalize the data using the mean and standard deviation of each variable. - Decide on the type of kernel function to use and determine the kernel parameter. - Compute the kernel matrix of the NOC using equation **(2.26)** and normalize it using equation **(2.40)** - Solve the eigenvalue problem given in equation **(2.30)** and normalize the eigenvectors using equation **(2.33)** - Calculate the monitoring statistics (T² and SPE) of the normal operating data. - Determine the control limits of the T² and SPE charts. **2. Online monitoring** - Obtain new data for each sample and scale it with the mean and variance obtained at step 1 of the modeling procedure. - Given the m-dimensional scaled test data $x_t \epsilon R^m$, compute the kernel vector $k_t \epsilon R^{1\times N}$by $[k_t]_j = k_t(x_t, x_j)$where $x_j \epsilon R^m, j = 1, \ldots, N$ is the normal operating data. - Mean center the test kernel vector k_t as follows: $$\widetilde{\boldsymbol{k_t}} = \boldsymbol{k_t} - \boldsymbol{1_t K} - \boldsymbol{k_t 1_N} - \boldsymbol{1_t K 1_N} \qquad (\mathbf{2.47})$$ - Where 1_N and K are obtained from step 3 from the modelling procedure and $1_t = {}^{1}/_{N}\,[1 \quad \ldots \quad 1]\,\epsilon\, R^{1\times N}$ - Calculate the monitoring statistics (T² and SPE) of the test data. - Monitor whether T² or SPE exceeds its control limit calculated in the modeling procedure.

II.3.3. KPCA main drawback

Although capable of capturing nonlinear relationships between variables with a high degree of accuracy, kernel principal component analysis (KPCA) suffers from a high computational cost and requires the storage of the symmetric kernel matrix (computation time increases with the number of samples), Given a set of n training examples, kernel PCA needs to perform eigendecomposition of the nxn kernel matrix K. As it takes $O(n^2)$ space to store K and $O(n^3)$ time to eigendecompose it, kernel PCA is prohibitively expensive for big data, where n is very large [4, 22].

In order to surmount the problem of high computational cost and storage, we have investigated the use of a reduced version of KPCA called reduced kernel principal component (RKPCA).

II.4. The proposed approach Reduced Kernel principal component (RKPCA)

When a large number of highly dependent variables are recorded, process data often contain redundant information as well as noise due to measurement errors. It is therefore possible to model variable relationships via a data-reduction method. (Okba et al. (2015)) proposed a RKPCA method that consists on approximating the retained principal components given by the KPCA method by a set of observation vectors which point to the directions of the largest variances with the retained principal components [27].

In this work, a new RKPCA is developed. The key idea of the proposed RKPCA method is to reduce the number of observations (samples) in the data matrix using the distance between samples such that only one observation is preserved in case of redundancy or similarity.

The proposed reduced KPCA (RKPCA) method selects a reduced number of observations $X_b \epsilon \{X_i\}\ i = 1, \ldots, N$ among the N measurement variables of the information matrix such that the retained observations can be used as a new data matrix.

It consists on calculating the distance between the observations in order to eliminate the similar ones. Once the reduced matrix is formed, conventional KPCA can be applied.

Assume a data matrix $X = [X_1 \quad X_2 \quad \cdots \quad X_N]^T$ with N observations and m process variables. Data consisting of measures of dissimilarity between all pairs of two observations can be represented using a dissimilarity matrix D as

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{D_{11}} & \dots & \boldsymbol{D_{1N}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{D_{N1}} & \dots & \boldsymbol{D_{NN}} \end{bmatrix} \qquad (\mathbf{2.48})$$

The elements $\boldsymbol{D_{ij}}\ \ \boldsymbol{i}, \boldsymbol{j} = \boldsymbol{1}, \dots, \boldsymbol{N}$ are called dissimilarities [58].

D can be constructed by means of a distance measure.

For $\boldsymbol{X_i} = [\boldsymbol{X_{i1}} \quad \boldsymbol{X_{i2}} \quad \cdots \quad \boldsymbol{X_{im}}]$,$\boldsymbol{X_j} = [\boldsymbol{X_{j1}} \quad \boldsymbol{X_{j2}} \quad \cdots \quad \boldsymbol{X_{jm}}]$ the most commonly used distance between observations is the Euclidean distance.

- Euclidean distance: is the ordinary straight-line distance and it is the sum of the squared differences between two observations [59].

$$\boldsymbol{D_{ij}} = \sqrt{\sum_{\boldsymbol{k=1}}^{\boldsymbol{m}} \left(\boldsymbol{X_{ik}} - \boldsymbol{X_{jk}}\right)^{\boldsymbol{2}}} \qquad (\boldsymbol{2.49})$$

The Euclidean distance may not be appropriate if the measurements are from different units. Thus, before calculating the dissimilarity matrix D, data matrix X should be normalized.

Table 04. RKPCA-based fault detection algorithm.

Algorithm 3: RKPCA-based fault detection
- Obtain data under normal operating conditions (NOC) and scale the data using the mean and standard deviation of the columns of the data set which represent the different variables - Calculate the elements of the dissimilarity matrix D using equation **(2.49)** - Pick the smallest distance $D_{ij(s)}$ and eliminate all the j^th^ observation from the non-normalized matrix. - Normalize the new data matrix and construct the reduced kernel matrix - Estimate the reduced KPCA model using algorithm 1 (the eigenvalues and vectors of the reduced kernel matrix). - Calculate T^2, Q and φ - Determine the control limits T^2_α, Q_α, φ_α - Obtain test data X_{tt} and normalize it using the mean and standard deviation of the new data matrix obtained in step 4 - Calculate T^2, Q and φ and compare them to their respective thresholds T^2_α, Q_α, φ_α . if the control limits are violated then an abnormal case holds. - Choose another distance D_{ij} such that $D_{ij} > D_{ij(s)}$,eliminate the corresponding j^{th} observation from the non-normalized matrix and repeat from step 3 - Keep increasing the distance D_{ij} as long as the T^2, Q and φ of the testing set are not deteriorated.

II.5. Conclusion

During this chapter, the mathematical basis of PCA models was established as well as the selection of the model parameters, such as the number of PCs and the loadings. Furthermore, KPCA approach and its theorical background has been proposed to cope with the problem of linearity. However, Due to its high computational cost, the RKPCA technique was introduced with its theory to deal with this issue.

The foundations of fault detection using this three methods was presented using the well-known Hotelling's T^2 and Q-statistics, in addition to a new proposed index called φ.

Chapter III.

Application, Results & Discussion

III.1. Introduction

In this chapter, the aforementioned PCA, KPCA and the proposed RKPCA methods are applied to the data of the cement plant collected under its NOC state and several common faults that can occur and disturb the normal operation of the system. The obtained results are shown in different figures and in term of False Alarm Rate (FAR), Missed Alarm Rate (MDR), Detection Time Delay (DTD), the Cost function (J) and the execution time (ET). This is followed by a discussion about these results and a comparison between PCA, KPCA and the proposed RKPCA approach. Before exhibiting the results and the discussion, the system and the data used in this experiment are described so that to make the experiment clear to the reader and make the results obtained more credible.

III.2. Process description

Cement is a substance which is made of grinded gypsum and cement clinker which itself is produced from a burned mixture of limestone and clay in certain percentages.
Cement manufacturing is a complex process that begins with mining and then grinding raw materials that include limestone and clay, to a fine powder, called raw meal, which is then heated to a sintering temperature as high as 1450 °C in a cement rotary kiln. In this process, the chemical bonds of the raw materials are broken down and then they are recombined into new compounds. The result is called clinker, which are rounded nodules between 1mm and 25mm across. The clinker is ground to a fine powder in a cement mill and mixed with gypsum to create cement [60].

Cement rotary kiln is the most vital part of a cement factory whose outcome is cement clinker. In Ain El Kebira cement plant in the Algerian east, where the work is conducted, a short rotary kiln of 5.4m shell diameter (Without brick and coating) and 80 m length, with 3˚ incline is used. Two 560 (kws) asynchronous motors spin the kiln at a low variable speed around 2.14 (rpm) producing clinker of density varying from 1300 to 1450 (kg/m3) under normal conditions. Two natural gas burners are used, the main one in the discharge end and the secondary in the first level of preheater tower [45]. The relevant signals that monitor the kiln system were collected from the historian of the aforementioned plant; they are listed and described in Table 05 and Figure 09. It gives simplified schematic of the installation.

Table 05. Process variables of the cement rotary kiln [45].

Signal	Description	Unit
1, 3, 5, 7	Depression of gases in outlets of cyclones: one, two, three, and four respectively, tower I	mbar
2, 4, 6, 8	Temperature of gases in outlets of cyclones: one, two, three, and four respectively, tower I	°C
10	Depression of gas in inlet of cyclone four, tower I	mbar
17, 19, 21, 23	Depression of gases in outlets of cyclones: one, two, three, and four respectively, tower II	mbar
18, 20, 22, 24	Temperature of gases in outlets of cyclones: one, two, three, and four respectively, tower II	°C
12, 25	Temperature of the material entering the kiln from tower I, and tower II respectively	°C
9, 15	Power of the motor driving the exhauster fans of tower I, and tower II respectively	kW
11, 16	Speed of the exhauster fans of tower I, and tower II respectively	r.p.m
13	Depression of gas in the outlet of the smoke filter of tower I	mbar
14, 26	Temperature of gas in the outlet of the smoke filter: tower I, and tower II respectively	°C
27	The sum of the powers of the two motors spinning the kiln	kW
28	Temperature of excess air from the cooler	°C
31	Temperature of the secondary air	°C
29, 32, 33	pressure of air under the static grille, repression of: fan I, fan II, and fan III respectively	mbar
30, 34	Speed of the cooling fan I, and fan III respectively	r.p.m
35, 37, 39	pressure of air under the chamber I, II, and III of the dynamic grille, repression of fan IV, fan V, and fan VI respectively	mbar
36, 38, 40	Speed of cooling fan IV, fan V, and fan VI respectively	r.p.m
41	Speed of the dynamic grille	strokes/min
42	Command issue of the pressure regulator for the speeds of the draft fans of cooler filter	r.p.m
43	Flow of fuel (natural gas) to the main burner	m^3/h
44	Flow of fuel (natural gas) to the secondary burner (pre-calcination level)	m^3/h

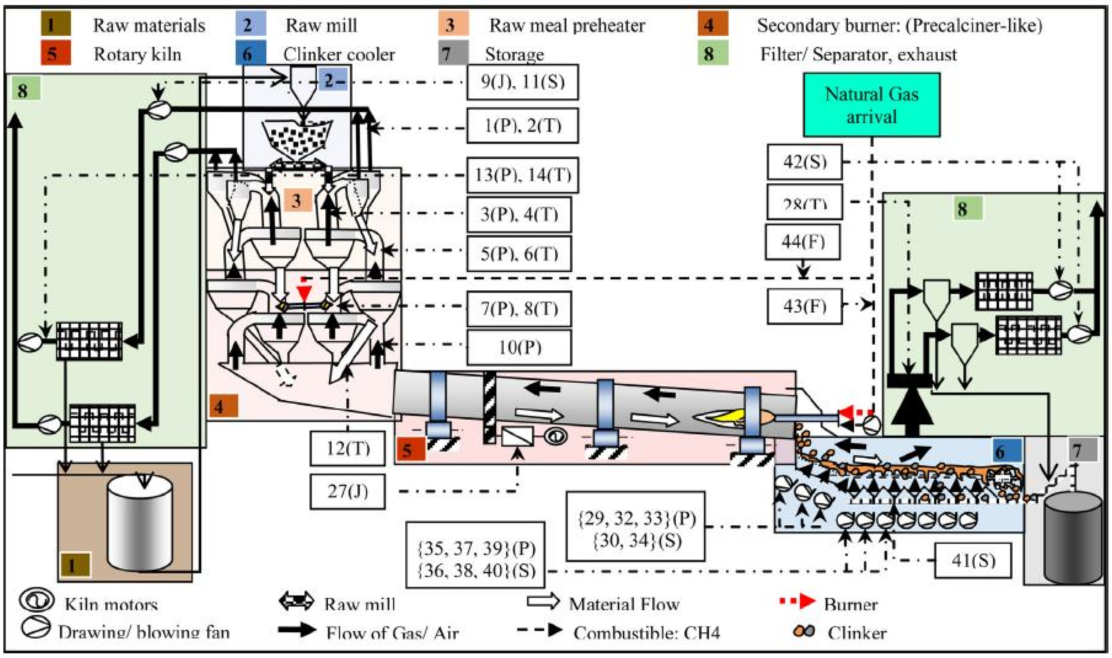

Figure 09. An overview of the manufacturing process in cement plant including signals used in the application (1-44) [45].

III.3. Computation of monitoring performance metrics

Monitoring performance was based on five metrics: False Alarms rate (FAR), Missed Alarms rate (MDR), Detection Time Delay (DTD), the Cost function J, Execution time (ET).

III.3.1. False Alarms rate

Calculated as the percentage of normal samples identified as faults (or abnormal) during the normal operation of the plant.

$$\boldsymbol{FAR = \frac{N_{N,F}}{N_N} \times 100\%} \qquad (3.1)$$

Where, $N_{N,F}$ is the number of normal samples detected as faults and N_N is the number of normal samples [45].

III.3.2. Missed Alarms rate

Calculated as the percentage of faulty samples identified as healthy samples during abnormal operation of the plant [45].

$$\boldsymbol{MDR = \frac{N_{F,N}}{N_F} \times 100\%} \qquad (3.2)$$

Where, $N_{F,N}$is the number of faulty samples identified as normal, and N_F is the number of faulty samples [45].

III.3.3. Detection Time Delay

Defined as the time required for indicating the fault after its occurrence.

$$\boldsymbol{DTD = t_d - t_o} \qquad (3.3)$$

Where, t_d and $t_o$is the detection and occurrence time respectively [45].

III.3.4. The Cost function

Another approach to evaluate a given method is to use the cost function. Using the three evaluation criteria introduced previously, a cost function of the form

$$\boldsymbol{J = q_1 \frac{FAR}{FAR_d} + q_2 \frac{MDR}{MDR_d} + q_3 \frac{DTD}{DTD_d}} \qquad (3.4)$$

is used.

Where, FAR_d, MDR_d and DTD_d are the desired values of FAR, MDR and DTD respectively.Whereas q_1, q_2 and $q_3$are their respective weights. By changing the values of $q's$ one can tailor the cost function of the problem at hand. For example, if one wanted to emphasize the FAR and ignore the amount of MDR, then one would increase q_1 and decrease q_2. The optimal solution would be the configuration that minimizes the cost function J [61].

In this study herein, FAR, MDR and DTD are of the same importance i.e. $q_1 = q_2 = q_3 = 1$.And FAR_d, MDR_d and DTD_d values are 1%, 1% and 1s respectively.

III.3.5. Execution time

The execution time or CPU time of a given task is defined as the time spent by the system executing that task in other way you can say the time during which a program is running [62]. In this work, a computer with Intel ® Core™ i5-2450M CPU @ 2.50 GHz is used.

III.4. Application procedure

This work is based on the real-time data collected by process computers from the cement plant. Table 6 lists the different data sets used in order to construct and test the proposed RKPCA then evaluate and compare its performances to that of PCA and KPCA.

Table 06. Data sets used in the application [45].

Data sets	Size	Sampling interval (s)	description
Training set	$X_0 \in R^{768\times44}$	20	Normal operation data, used to construct FD scheme
Testing set	$X_0 \in R^{11000\times44}$	1	Normal operation data, used to test the FD scheme
Process fault	$X_0 \in R^{2084\times44}$	1	Normal/Faulty operation, process fault
Sensor faults (6 sets)	$X_0 \in R^{1500\times44}$	1	Sensor fault simulations

III.4.1. Data generation and description

Set 1, Training data set: consists of 44 variables listed in Table 5 and indicated in Figure 09. About 768 observation samples (more than 4 h) are collected during normal condition operation NOC, with a sampling rate of 1 sample each 20s.This set is used to construct the fault detection model: (i) building the PCA, KPCA and RKPCA model, (ii) computing the upper limits of T^2, Q and φ for each model [45].

Set 2, Testing data set: consists of 11 000 sample, collected from the plant during healthy operation with a sampling interval of 1s. This data set is used to test the accuracy of the model compared to the upper limits calculated using the training set in terms of the false alarms contributed by the T^2, Q and φ statistics[45].

Set 3, Faulty process data: The faulty data is collected from the Cement plant for more than half of an hour (2084 s) with a sampling interval of 1s. The set consists of two regions; first region corresponds to healthy operation, whereas the second region corresponds to the process faulty operation. The fault occurs after 7 min and evolves slowly in the process. We use this data to check the ability of the proposed detection method to detect the fault [45].

Set 4, seven simulated sensor/actuator faults: This set contains 7 simulated sensor faults occurring in the rotary kiln process. It includes abrupt, random, intermittent, and slow drift additive faults that might be single as well as multiple faults. Each simulation covers 1500s of process operation. The original data is taken during healthy operation of the rotary kiln, and then faults are introduced from 500^{th} sample to 1000^{th} sample. Thus each simulation has three regions (healthy; faulty; healthy). Each region lasts for 500s. The simulation is done by adding a fault of magnitude 2% to the data collected during normal (healthy) operation. In the intermittent case, the fault is introduced in the following intervals: 500-580, 610-660, 700-740, 800-830, 870-900 and 975-1000 with amplitude 5.5%, 4.5%, 5%, 5.5%, 5% and 4.5% respectively[45].

Table 7 lists the 7 sensor faults, with type and magnitude of each fault.

Table 07. Simulated sensor faults introduced at 500-1000 s [45].

Fault	Faulty variables	Fault magnitude	Description of the fault
SFault(1)	16	(0, 5%)	Additive random fault, with mean 0, and variance 0.05
SFault(2)	44	−2%	Abrupt additive fault, bias $b = -0.02$
SFault(3)	30	+2%	Additive fault: Linear drift from 0% to 2%; slope $Ks = 4 \times 10^{-5}$
SFault(4)	34	−2%	Additive fault: Linear drift from 0% to −2%; slope $= -4 \times 10^{-5}$
SFault(5)	12, 18, 43	[+, −, +]2%	Abrupt additive fault ±2% (multiple)
SFault(6)	4, 6, 8, 14, 24	[+,+,+,−,−]2%	Additive fault: Linear drift from 0% to ±2% (multiple); slope $= \pm 4 \times 10^{-5}$
SFault(7)	11	+4.5%–5.5%	Additive fault: Intermittent fault, changing intervals and amplitudes

III.4.2. Application of PCA, KPCA and the proposed approach RKPCA to cement rotary kiln process

In this part, PCA, KPCA and the proposed approach RKPCA are applied to the previously described cement plant and the followed steps to build and evaluate the models are explained.

III.4.2.1. PCA monitoring model

PCA has been applied to cement rotary kiln system. The training data in Table 6 under normal operation condition is first scaled (to have zero mean and unit variance) then used to build the PCA model and to determine the statistical control limits $T_\alpha^2, Q_\alpha, \varphi_\alpha$ at the confidence levels 95% and 99%. In this study herein, the CPV method is utilized to find out the optimum number of retained PCs. The CPV method is set to capture 90% of the total variance in the training data, which results in retaining 20 PC's$(\ell = 20)$. The performance of the PCA model is evaluated in the training part using the monitoring performance metric *FAR.*

The testing set in Table 6 are used to validate the model by comparing the time evolution of the fault detection indices T^2, Q and φ to their respective thresholds $T_\alpha^2, Q_\alpha, \varphi_\alpha$ (at the confidence levels 95% and 99%) obtained in the training part in terms of FAR.

The PCA model is then applied to 8 fault sets (Tab. 07) in order to assess the model's efficiency based on the monitoring performances *FAR* (for the healthy region), *MDR* and *DTD* (for the faulty region).

The cost function has been also calculated to simplify the comparison between different methods where an average value of J of all the fault sets is deduced and compared to other J's obtained from other methods. Concerning the execution time criteria the Tic/Toc matlab command is used to assess the model in terms of time consuming.

III.4.2.2. KPCA monitoring model

As described in Algorithm 2, the fault-free training data (Table 6) is used to construct a KPCA reference model to be used in fault detection. This data is first scaled (to have zero mean and unit variance), and then used to construct the KPCA model. In this study, the same criteria as in PCA model has been used for comparison purposes. Hence, the CPV method is also utilized to find out the optimum number of retained PCs that captures 90% of the total variance, which results in retaining 20 PCs $(\ell = 20)$. Another, important parameter for kernel-based methods in model development for process monitoring is the choice of the kernel function and its width. The radial basis kernel $K(X,Y) = \exp(\frac{-\|X-Y\|^2}{c})$ is utilized in this work and normalized using equation **(2.40)**. The value of the kernel parameter c depends on the process being monitored and has been set to c=13 000 in this application. The fault detection techniques through Q, T^2 and φ statistics are first carried out using the training fault-free data and their respective thresholds

$T_{\alpha}^2, Q_{\alpha}, \varphi_{\alpha}$ at the confidence levels 95% and 99% are calculated. The KPCA model is then assessed using FAR.

The testing data in table 6 is used to validate the KPCA model. First, the data is scaled by using the mean and standard deviation of the training set, and then the Kernel matrix of the testing part is calculated and normalized as described in algorithm 2. KPCA model is also applied to the process fault and the seven simulated sensor faults (Table 7) and has been assessed in terms of FAR (For healthy region), MDR and DTD (for faulty region).

The cost function has been calculated using the monitoring indices where an average value of J of all the fault sets is deduced.

For KPCA modeling, an important criteria is taken in consideration: the execution time. As stated previously, kernel PCA is known by its high computational and storage problem and our purpose is to decrease this time by the proposed RKPCA approach. The Tic/Toc matlab command is used to calculate the execution time.

III.4.2.3. The proposed RKPCA monitoring model

As mentioned in the previous chapters, the proposed approach is a reduced version of KPCA, which means, once the number of observations is reduced in the training set, a conventional KPCA can be applied.

The reduction part consists on calculating the dissimilarity matrix from the normalized 768×768 training data matrix. The dissimilarities or the distances are sorted in ascending order to simplify. As a first step, one observation from each pair with the smallest dissimilarity is eliminated from the non-normalized training set 768×768. Hence, the number of observations has been reduced. After normalizing the new reduced training data, a conventional KPCA model is applied. The new model is then evaluated in terms of FAR and execution time ET, and the number of samples obtained in each model is recorded.

Due to a deterioration of the FAR values noticed in the testing set, this procedure was repeated 53 times by selecting 53 distance from the total distances.

To select the optimum distance that provides a considerable reduction in samples, the reduced models or the distances that provides the minimum values of FAR when using $T_{95}^2, T_{99}^2, Q_{95}, Q_{99}, \varphi_{95}, \varphi_{99}$ are selected. In our case, 7 models or distances are retained. These models are then applied to the process fault and the seven simulated sensor faults.

For each model and with each fault, FAR, MDR, DTD and J of the three indices using the thresholds $T_{95}^2, T_{99}^2, Q_{95}, Q_{99}, \varphi_{95}, \varphi_{99}$ are calculated. At the end, an average value of J is obtained for each model or distance, the optimal distance is then the distance with the minimum average J.

III.5. Results and discussion

III.5.1. PCA monitoring model

III.5.1.1. NOC results

The thresholds of T^2, Q and φ deduced from the training set are found to be:

$$\boldsymbol{T_{95}^2 = 35.74\,, T_{99}^2 = 51.13\,, Q_{95} = 7.57\,, Q_{99} = 10.71, \varphi_{95} = 1.21, \varphi_{99} = 1.65}$$

The performance of the fault detection model based PCA in terms of FAR set is summarized in table 8. From this table, it can be seen that globally the three monitoring indices (T^2, Q, φ) showed good results of FAR. The FAR contributed by the PCA model in the training set are good. This clearly indicates the accuracy of the model. In the testing set, a low FAR is obtained for the rotary kiln monitoring using PCA. Besides, it is even negligible when using a confidence level of 99%.

The execution of the model lasted only 3.57 min and that because PCA has performed an eigendecomposition of a 44×44 covariance matrix.

Table 08. FAR contributed by T^2, Q, φ under NOC using PCA.

Method	**NOC Data**	**index**	**Confidence level**	FAR (%)	Execution time (min)
PCA	**training data**	**T^2**	**95%**	4.95	3min 34 s
			99%	1.04	
		Q	**95%**	4.95	
			99%	1.04	
		φ	**95%**	4.95	
			99%	1.04	
	testing data	**T^2**	**95%**	4.45	
			99%	0.37	
		Q	**95%**	5.84	
			99%	0.82	
		φ	**95%**	5.53	
			99%	0.61	

Figure 10 shows the time evolution of the fault detection index φ based PCA for training and testing set. The horizontal line in red represent the threshold with a confidence level of 95%, whereas the one in green represent the threshold with a confidence level of 99%.

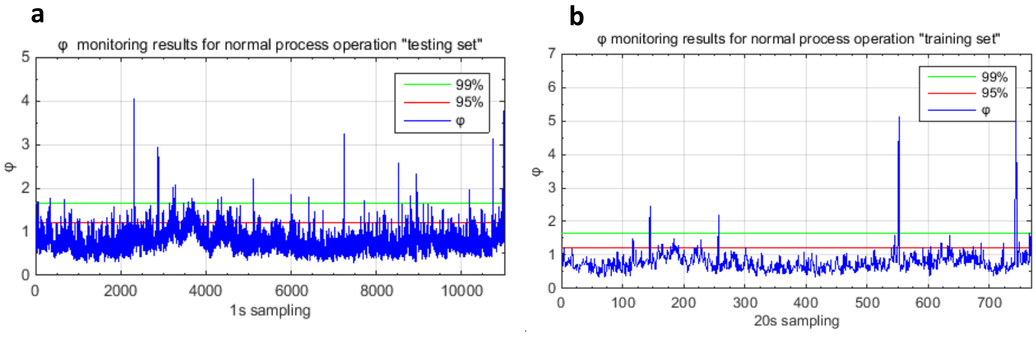

Figure 10. φ monitoring results of healthy process operation using PCA. (a) testing data set; (b) training data set.

III.5.1.2. the involuntary real process fault results

Figure 11 shows the occurrence of an involuntary process fault using φ statistics with its corresponding thresholds.

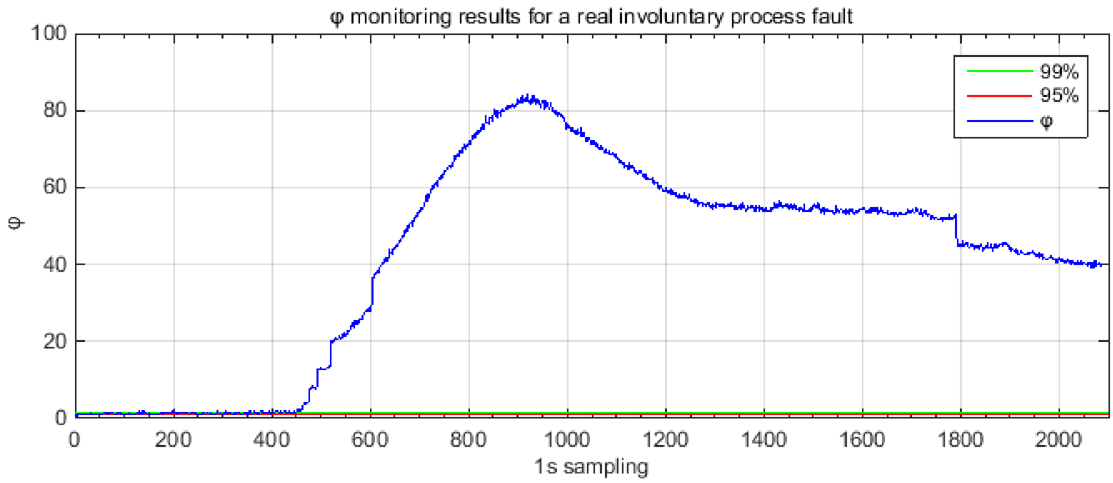

Figure 11. φ monitoring of real involuntary process fault in the cement rotary kiln using PCA.

The performances of the fault detection approach based on PCA model in terms of MDR, FAR, DTD and the Cost function J are summarized in Table 9.

The involuntary process fault was promptly detected by the Q statistic at its two thresholds Q_{95}, Q_{99}; and T^2 and φ at their thresholds T^2_{99}, φ_{99} respectively. However, the detection time is delayed when detected by T^2_{95}, φ_{95} to 30s and 2s respectively. In terms of MDR, the PCA based φ index has provided the best performance at confidence level of 95% (MDR=0.00%), whereas negligible values using T^2 and Q monitoring indices.

In terms of FAR, T², Q and φ showed high values except using T^2_{99} where FAR equals 1.67%.As a summarized index, the process fault is efficiently detected when using Q and φ at their thresholds Q_{99} and φ_{99} respectively.

Table 09. Missed detection rate (MDR), False Alarm Rate (FAR), detection time delay (DTD) and the cost function J values for the eight faults of cement rotary kiln using PCA.

Faults	Performances	T²		Q		φ	
		95%	99%	95%		95%	99%
Process fault	MDR	1.38	1.98	0.24	1.68	0.00	1.14
	FAR	15.48	1.67	59.05	14.05	71.67	14.52
	DTD	0.00	30.00	0.00	0.00	0.00	2.00
	J	**16.86**	**33.65**	**59.29**	**15.73**	**71.67**	**17.66**
Random and single	MDR	2.00	2.40	1.20	1.20	1.20	1.40
	FAR	1.10	0.30	5.60	0.80	3.90	0.60
	DTD	1.00	1.00	1.00	1.00	1.00	1.00
	J	**4.10**	**3.70**	**7.80**	**3.00**	**6.10**	**3.00**
Abrupt and single	MDR	0.00	0.00	0.40	7.60	0.00	0.00
	FAR	1.80	0.90	5.50	1.70	3.80	1.30
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.80**	**0.90**	**5.90**	**9.30**	**3.80**	**1.30**
Drift and single 1	MDR	4.60	6.80	2.40	10.80	2.60	5.80
	FAR	4.20	0.70	4.60	0.40	5.60	1.10
	DTD	18.80	27.00	0.00	5.00	5.00	27.00
	J	**27.60**	**34.50**	**7.00**	**16.20**	**13.20**	**33.90**
Drift and single 2	MDR	13.40	21.20	1.40	2.20	1.40	2.40
	FAR	0.70	0.10	4.00	0.10	1.70	0.10
	DTD	1.50	90.00	1.40	1.90	2.50	10.50
	J	**15.60**	**111.30**	**6.80**	**4.20**	**5.60**	**13.00**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.10	0.20	5.80	1.50	3.40	0.60
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.10**	**0.20**	**5.80**	**1.50**	**3.40**	**0.60**
Drift and multiple	MDR	17.00	22.60	12.60	15.60	12.20	15.80
	FAR	2.40	1.00	9.60	1.30	5.60	1.50
	DTD	55.00	111.40	54.90	71.40	54.40	54.90
	J	**74.40**	**135.00**	**77.10**	**88.30**	**72.20**	**72.20**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	5.14	1.04	9.80	1.85	8.67	1.37
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**5.14**	**1.04**	**9.80**	**1.85**	**8.67**	**1.37**
J average		**18.32**	**40.03**	**22.43**	**17.51**	**23.08**	**17.87**

III.5.1.3. The simulated sensor faults results

In this section, the performance of the PCA based fault detection is evaluated in detecting single as well as multiple sensor faults of abrupt, random, intermittent, and drift types (Fig. 12).

The performances of all the seven simulated sensor faults are tabulated in Table 9.

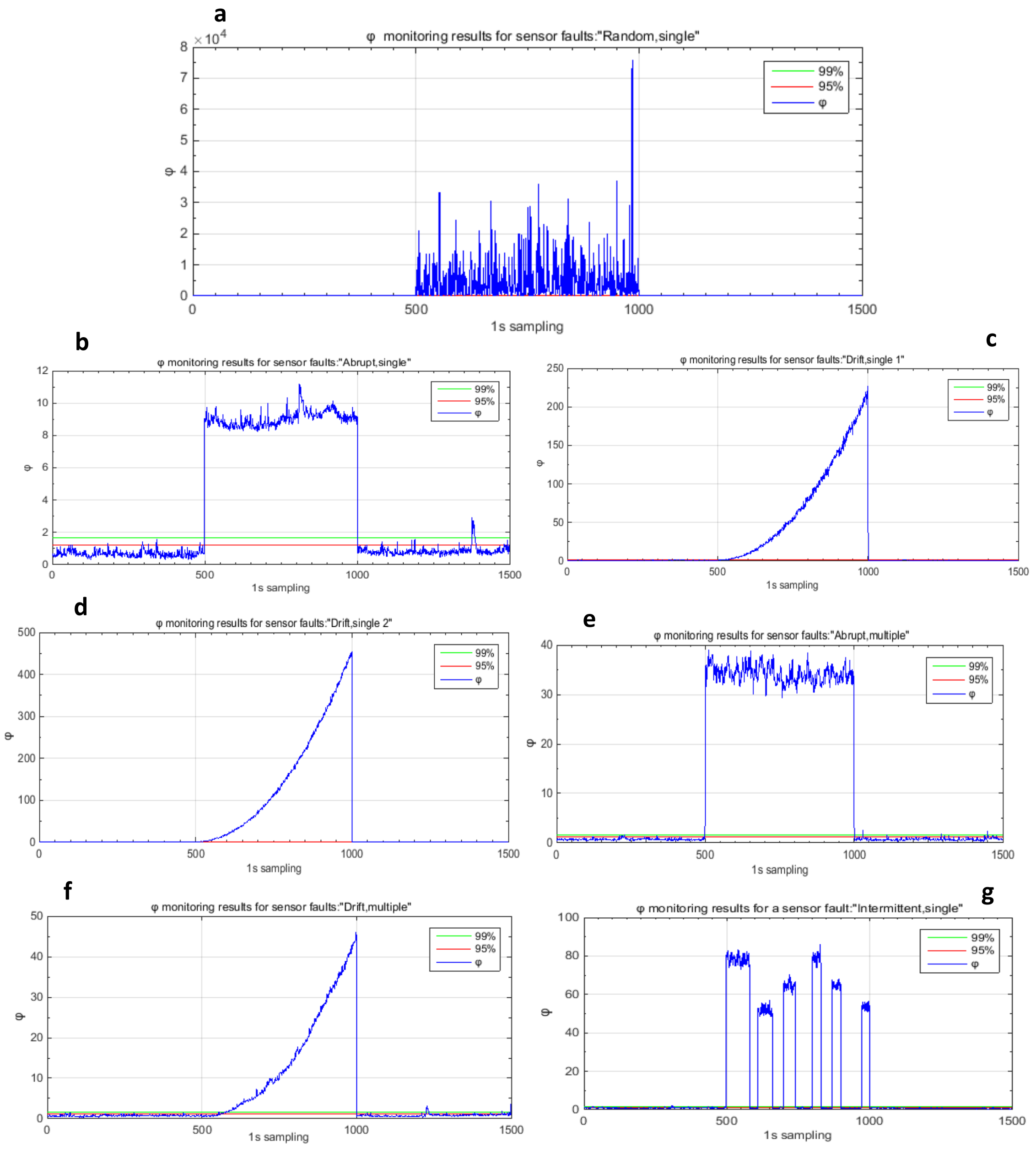

Figure 12. φ monitoring results of sensor faults using PCA. **(a)** Sfault1; **(b)** Sfault2; **(c)** Sfault3; **(d)** Sfault4; **(e)** Sfault5; **(f)** Sfault6; **(g)** Sfault7.

For abrupt fault in single sensor (SFault2) or multiple sensors (SFault5) and intermittent fault in single sensor (SFault7), the faults are instantly detected.

The detection time of the random fault in single sensor (SFault1) and drift faults in single sensor (SFault3, SFault4) are slightly delayed. However, the PCA model have detected the drift fault in multiple sensors (SFault6) after approximately 2min using T_{99}^2 and approximately 1min using the remaining thresholds.

In terms of MDR and FAR, Abrupt fault in single sensor (Sfault2) or multiple sensors (Sfault5) and intermittent fault in single sensor (Sfault7) showed the best performance with zero found missed alarms (MDR= 0.00%) and negligible values of false alarms.

Acceptable values of FAR and MDR are seen when detecting the Random fault in single sensor (Sfault1), especially when using T_{99}^2 where FAR equals 0.30%. However, a large amount of missed alarms is noticed when detecting drift fault in multiple sensors (Sfault6) using all indices, drift fault in single sensor (Sfaul3) when using the confidence level of 99% and drift fault in single sensor (Sfault4) when using T^2 monitoring index.

As a summarized index, the cost function is calculated in each fault and for each monitoring index (with its two confidence levels 95% and 99%). J shows the overall performance of the PCA based fault detection. From table 9, J has shown efficient results when detecting Sfault1 and Sfault7 especially with confidence level 99% and Sfault2 and Sfault5 when using thresholds T_{99}^2 and φ_{99}.

Sfault3 and Sfault4 have provided acceptable values with thresholds Q_{99} and T_{99}^2 respectively whereas high values when the remaining indices are used.

PCA has presented a big J value using the three monitoring indices for the detection of Sfault6.

III.5.2. KPCA monitoring model

III.5.2.1. NOC results

The thresholds of T^2, Q and φ deduced from the training data are found to be:

$$\boldsymbol{T_{95}^2 = 35.68\,, T_{99}^2 = 50.84\,, Q_{95} = 0.12 \times 10^{-4}\,, Q_{99} = 0.17 \times 10^{-4}},$$

$$\boldsymbol{\varphi_{95} = 1.23, \varphi_{99} = 1.69}$$

The performances of the fault detection model based KPCA in terms of FAR are tabulated in table 10.

From table 10, the FAR values provided by the three monitoring indices T^2, Q and φ shows good results which confirm the accuracy of the KPCA model. In the validation phase, the model has provided the best performances in the three monitoring indices.

1h 37min represents the time consumed in processing and analyzing the NOC data sets, which is a considerable time. This is because KPCA performed an eigendecomposition of a 768×768 kernel matrix.

Table 10. FAR contributed by T^2, Q, φ under NOC using KPCA.

Method	**NOC Data**	**index**	**Confidence level**	FAR (%)	Execution time (min)
KPCA	**training data**	**T^2**	**95%**	4.95	1h37min
			99%	1.04	
		Q	**95%**	4.95	
			99%	1.04	
		φ	**95%**	4.95	
			99%	1.04	
	testing data	**T^2**	**95%**	0.37	
			99%	0.08	
		Q	**95%**	0.83	
			99%	0.65	
		φ	**95%**	0.96	
			99%	0.57	

Figure 13 shows the monitoring results of KPCA model in normal operation (training and testing set) using the index φ. The horizontal line in red represent the index threshold with a confidence level of 95%, whereas the one in green represent the index threshold with a confidence level of 99%.

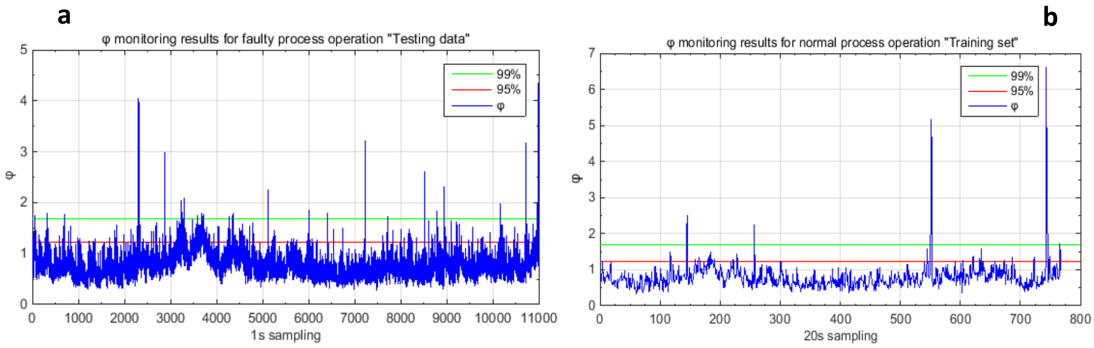

Figure 13. φ monitoring results of healthy process operation using KPCA. (a) testing data set; (b) training data set.

III.5.2.2. The involuntary real process fault results

Table 11 summarizes the performances of KPCA-based model in terms of MDR, FAR, DTD and Cost function J.

The table shows that the detection time of this fault is lightly delayed when the three monitoring indices are used with their two confidence level (95%, 99%).

In terms of FAR and MDR, high amount of false alarms is shown by the three indices except for T^2_{99} where FAR equals to 1.67.Whereas, a slight number of missed detected samples is noticed especially when using φ_{95} for the detection where zero amount of missed samples is provided.

In order to give a global view of the performances in detecting this fault, the cost function is used. Minimum values of J are seen when the thresholds φ_{99} and Q_{99} are utilized.

Table 11. Missed detection rate (MDR), False Alarm Rate (FAR), detection time delay (DTD) and the cost function J values for the eight faults of cement rotary kiln using KPCA.

Faults	Performances	T²		Q		φ	
		95%	99%	95%	99%	95%	99%
Process fault	MDR	1.44	1.98	0.18	1.26	0.00	1.14
	FAR	14.29	1.67	62.62	14.76	70.71	11.90
	DTD	0.00	30.00	1.10	2.00	0.00	2.00
	J	**15.73**	**33.65**	**63.90**	**18.02**	**70.71**	**15.04**
Random and single	MDR	2.00	48.60	1.20	1.20	1.20	1.40
	FAR	1.10	0.20	5.70	0.80	3.60	0.50
	DTD	0.35	2.00	1.00	1.00	1.00	1.00
	J	**3.45**	**50.80**	**7.90**	**3.00**	**5.80**	**2.90**
Abrupt and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.80	0.90	5.30	1.70	3.70	1.30
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.80**	**0.90**	**5.30**	**1.70**	**3.70**	**1.30**
Drift and single 1	MDR	4.60	6.80	1.80	7.00	2.60	6.00
	FAR	4.10	0.70	4.70	0.40	5.40	1.10
	DTD	19.00	27.00	0.50	5.50	5.00	27.00
	J	**27.70**	**34.50**	**7.00**	**12.90**	**13.00**	**34.10**
Drift and single 2	MDR	13.60	22.00	1.40	2.20	1.20	2.60
	FAR	0.70	0.10	4.20	0.10	1.70	0.10
	DTD	1.50	90.00	1.50	1.90	2.50	10.40
	J	**15.80**	**112.10**	**7.10**	**4.20**	**5.40**	**13.10**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.10	0.20	6.00	1.50	3.40	0.60
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.10**	**0.20**	**6.00**	**1.50**	**3.40**	**0.60**
Drift and multiple	MDR	17.00	22.80	12.40	15.60	12.20	16.00
	FAR	2.40	1.00	10.00	1.30	5.40	1.40
	DTD	55.00	111.40	55.00	71.50	55.00	55.00
	J	**74.40**	**135.20**	**77.40**	**88.40**	**72.60**	**72.40**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	5.14	1.04	9.96	1.85	8.59	1.29
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**5.14**	**1.04**	**9.96**	**1.85**	**8.59**	**1.29**
J average		**18.14**	**46.04**	**23.07**	**16.44**	**22.90**	**17.59**

The time evolution of the fault detection index φ based KPCA of the real process fault are depicted in Figure 14.

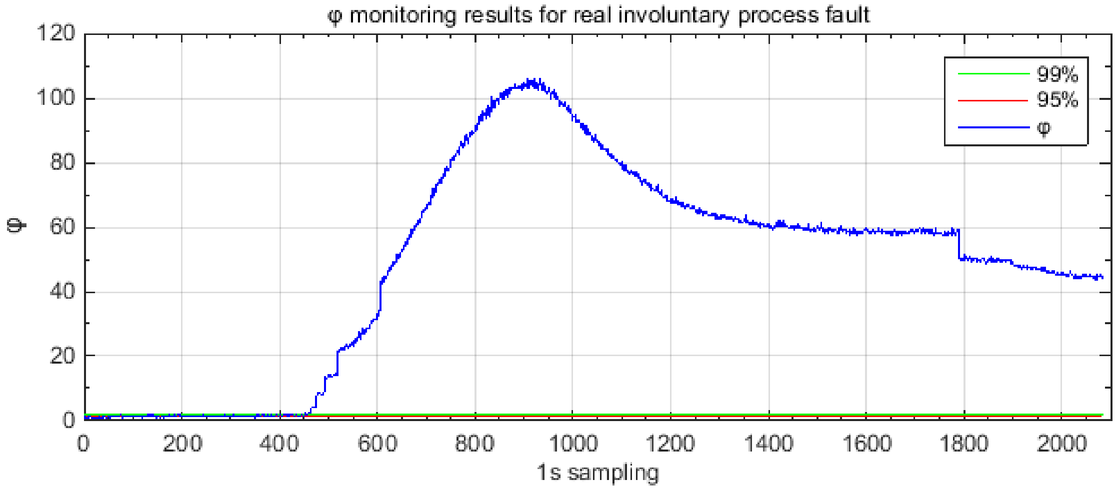

Figure 14. φ monitoring of real involuntary process fault in the cement rotary kiln using KPCA.

III.5.2.3. Simulated sensor faults results

The performances of the KPCA based model in detecting all these faults are summarized in table 11.

From the table, abrupt fault in single sensor (Sfault2) or multiple sensors (Sfault5) and intermittent fault in single sensor (Sfault7) are detected immediately and efficiently with no delay and zero amount of missed detected samples. As well, the quantity of false alarms is negligible except when the intermittent fault in single sensor (Sfault7) is detected using the three indices with confidence limit of 95%.

The KPCA based model has also promptly detected the random fault (Sfault1) and the drift fault in single sensor (Sfault3) with the three monitoring indices. Whereas, the detection is delayed to 1min30s when using T^2_{99} to detect the drift fault in single sensor (Sfault4).

In terms of FAR, negligible values in detecting random fault in single sensor (Sfault1) and drift fault in single sensor (Sfault3, Sfault4) are noticed. Similarly, the amount of missed alarms are acceptable except the detection of: Sfault1 using T^2_{99}, Sfault3 using the three indices with confidence level of 99% and Sfault4 using the monitoring index T^2.

Globally, the performance of KPCA model in terms of average J has shown that abrupt fault in single sensor (Sfault2) or multiple sensors (Sfault5) and intermittent fault in single sensor (Sfault7) are efficiently detected particularly when using the confidence limit 99%. Whereas, random fault (Sfault1) and drift fault in single sensor (Sfault4) are effectively detected using the two statistics Q andφ as well as (Sfault3) using the Q. Drift fault in multiple sensors (Sfault6) has shown high values of J in the three monitoring indices.

Figure 15 show the time evolution of the φ monitoring index using KPCA. The detection of single as well as multiple sensor faults of abrupt, random, intermittent, and drift types can be noticed from the graphs.

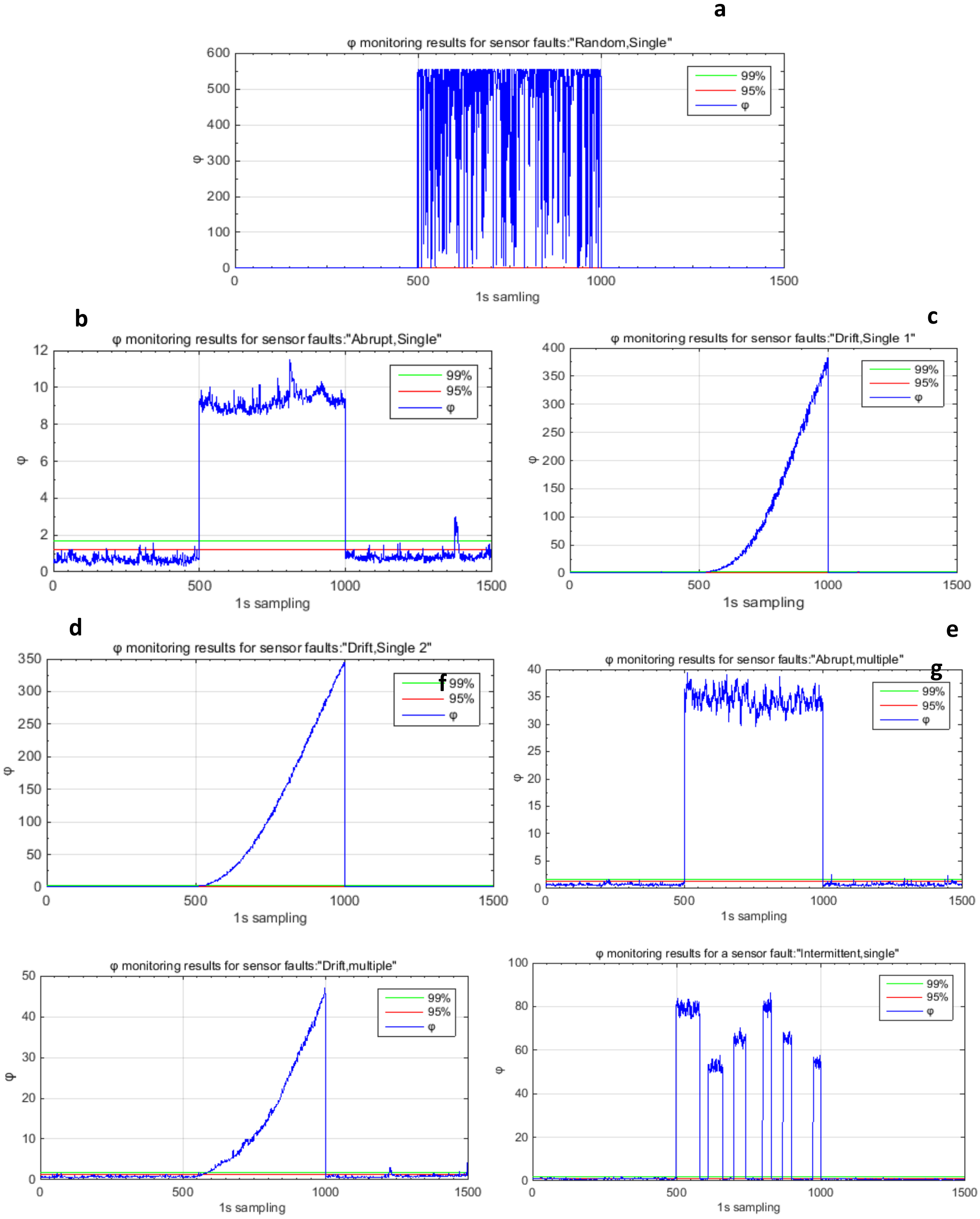

Figure 15. φ monitoring results of sensor faults using KPCA. **(a)** Sfault1; **(b)** Sfault2; **(c)** Sfault3; **(d)** Sfault4; **(e)** Sfault5; **(f)** Sfault6; **(g)** Sfault7.

III.5.3. The proposed RKPCA monitoring model

III.5.3.1. NOC results

The monitoring results of the 53 RKPCA model in normal operation (training and testing set) using the three indices T^2, Q and φ (with confidence levels 95% and 99%) are obtained. The models are assessed using the FAR values, as well as the number of samples and the execution time for each model (see table 01. Appendix).

The FAR values shown in Table 01 in Appendix are depicted as a graph in Figure 16.

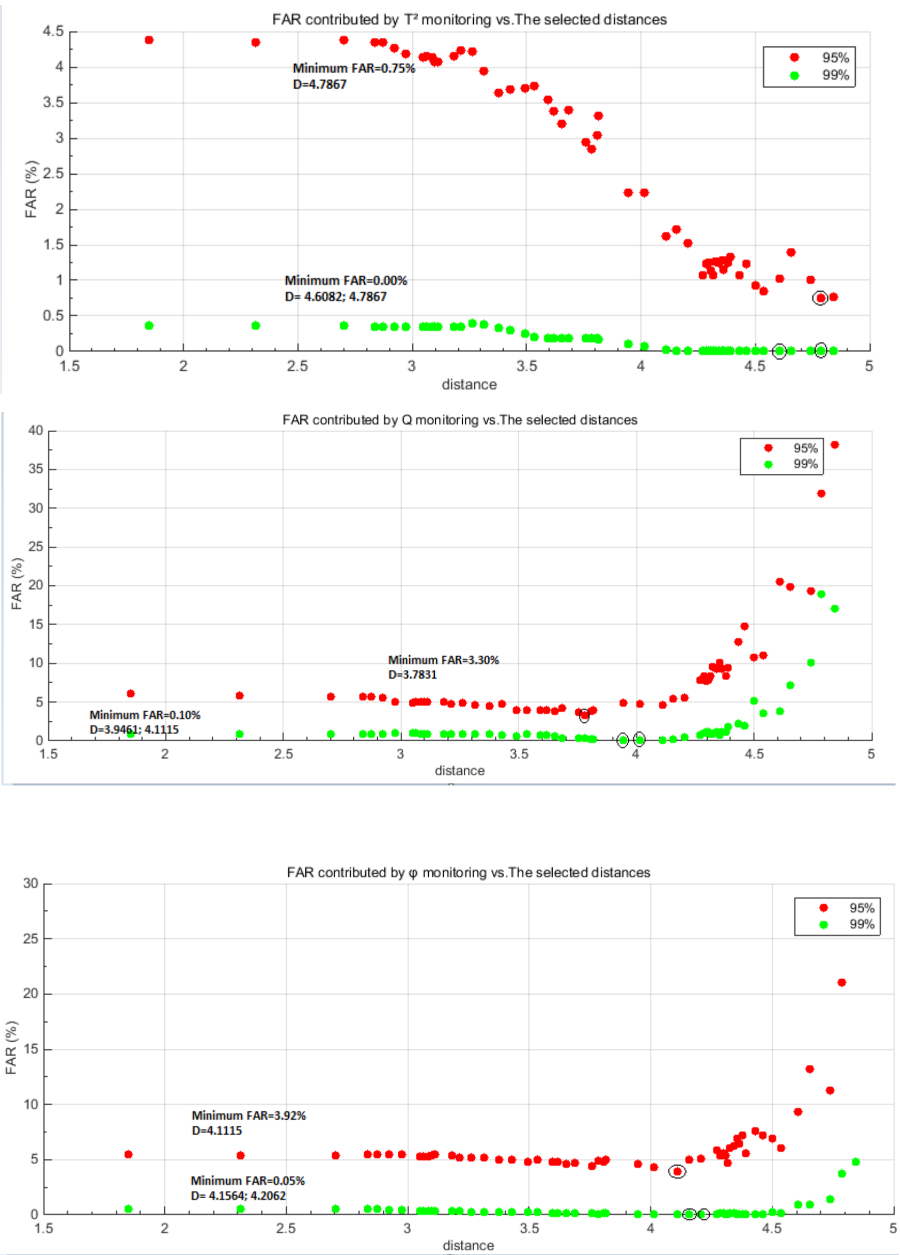

Figure 16. FAR contributed by T^2, Q and φ under NOC (Testing set) versus the 53 selected distances.

From the figure 16, an increasing FAR of Q (with confidence limit 95% and 99%) is noticed starting from distance 4.7867 (108 sample). The dropping of the number of samples from 768 to 108 has caused a loss of information and a big variability in the residual space due to the lack of fit of the new data to the RKPCA. Since φ (with confidence levels 95% and 99%) is expressed in terms of Q, an increasing in its FAR is obvious. At the same time, T^2 has shown a decreasing FAR values which shows that the informations explained by the PC's are not affected by this reduction.

The encircled points in Figure 16 represents the seven retained distances and their relative minimum FAR. The models that match these distances are applied to the various faults, and evaluated in terms of average J (see tables 2-8 in appendix).

Distance 4.6082 with 144 sample and 9min30s execution time has provided the best performances (minimum average J) among the 7 distances in each of Q with confidence level 99% and φ with the two confidence levels 95% and 99% (Tab. 12). Thereby, the model that matches this distance is selected as the RKPCA model.
The bold values highlight the best performance of the proposed method.

Table 12. The average value of the cost function J of the seven retained models.

Distances	N° of samples	Execution time (min)	T^2 95%	T^2 99%	Q 95%	Q 99%	φ 95%	φ 99%
			Average J					
3.7831	447	30.08	27.98	51.74	20.31	21.89	23.76	27.54
3.9461	359	23.88	12.39	47.96	20.23	22.98	23.30	28.27
4.1115	298	22.66	18.48	48.31	19.10	23.41	23.35	28.18
4.1564	281	22.09	18.29	55.19	19.88	20.20	22.99	27.02
4.2062	267	20.28	24.53	56.36	18.43	16.11	21.31	24.80
4.6082	**144**	**9.51**	30.87	69.35	23.89	**12.17**	**17.09**	**12.19**
4.7867	108	6.82	31.30	65.78	37.34	26.27	30.78	15.64

The thresholds of T^2, Q and φ deduced from the training set are found to be:

$$\boldsymbol{T^2_{95} = 27.64\,, T^2_{99} = 78.28\,, Q_{95} = 9.54 \times 10^{-4}\,, Q_{99} = 0.13 \times 10^{-2}, \varphi_{95} = 1.04, \varphi_{99} = 1.42}$$

The performance of the fault detection model based RKPCA in terms of FAR in the NOC is summarized in table 13. The FAR contributed by the RKPCA model in the training set are good especially for confidence level 99%. This clearly indicates the model's precision. In the testing set, T^2 monitoring has shown the best performance in terms of FAR. Also, small values of FAR are seen in Q and φ using the confidence level 99% whereas large values are noticed when using confidence level 95%.

Table 13. FAR contributed by T², Q, φ under NOC using the proposed RKPCA.

Method	NOC Data	index	Confidence level	FAR (%)	Execution time (min)
RKPCA	**training data**	**T²**	**95%**	4.86	9 min 30 s
			99%	0.69	
		Q	**95%**	4.86	
			99%	0.69	
		φ	**95%**	4.86	
			99%	0.69	
	testing data	**T²**	**95%**	1.02	
			99%	**0.00**	
		Q	**95%**	20.55	
			99%	3.75	
		φ	**95%**	9.31	
			99%	0.88	

Figure 17 shows the time evolution of the fault detection index φ based RKPCA for training and testing set. The horizontal line in red represent the threshold with a confidence level of 95%, whereas the one in green represent the threshold with a confidence level of 99%.

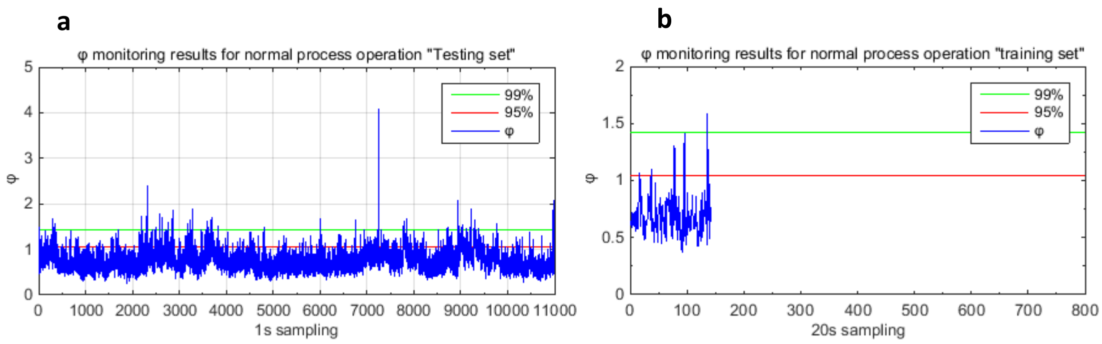

Figure 17. φ monitoring results of healthy process operation using RKPCA. (a) testing data set; (b) training data set.

III.5.3.2. The involuntary real process fault

Table 14 summarizes the performances of RKPCA-based model in terms of MDR, FAR, DTD and Cost function J.

From the table, φ index with its two confidence levels 95% and 99% has detected immediately the process fault (DTD=0.00s).Whereas a slight delay is noticed when using T² and Q indices. In terms of MDR, the proposed RKPCA has shown small amount of missed detected samples when using all the indices. However, FAR provided considerable values when using all indices with confidence level 95%.

Table 14. Missed detection rate (MDR), False Alarm Rate (FAR), detection time delay (DTD) and the cost function J values for the eight faults of cement rotary kiln using the proposed RKPCA.

Faults	Performances	T^2		Q		φ	
		95%	99%	95%	99%	95%	99%
Process fault	MDR	0.60	3.31	1.02	1.92	0.60	2.04
	FAR	54.29	0.00	13.10	0.95	20.71	1.19
	DTD	4.80	55.00	3.00	3.90	0.00	0.00
	J	**59.69**	**58.31**	**17.12**	**6.77**	**21.31**	**3.23**
Random and single	MDR	30.80	44.40	1.00	1.60	1.00	1.80
	FAR	1.30	0.20	18.00	1.80	9.10	0.80
	DTD	0.00	0.00	1.00	1.00	1.00	1.00
	J	**32.10**	**44.60**	**20.00**	**4.40**	**11.10**	**3.60**
Abrupt and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.00	0.20	25.10	4.10	11.40	1.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.00**	**0.20**	**25.10**	**4.10**	**11.40**	**1.20**
Drift and single 1	MDR	4.80	8.80	2.40	7.00	2.80	5.80
	FAR	3.30	0.10	7.30	0.40	2.60	0.10
	DTD	22.00	42.00	0.50	8.30	5.00	8.70
	J	**30.10**	**50.90**	**10.20**	**15.70**	**10.40**	**14.60**
Drift and single 2	MDR	10.00	27.00	1.20	1.80	1.20	1.80
	FAR	0.50	0.10	10.40	0.90	3.30	0.30
	DTD	0.00	118.9	1.30	1.74	1.30	1.95
	J	**10.50**	**146.00**	**12.90**	**4.44**	**5.80**	**4.05**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.30	0.20	34.70	7.60	15.60	2.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.30**	**0.20**	**34.70**	**7.60**	**15.60**	**2.20**
Drift and multiple	MDR	20.60	47.00	8.00	10.80	10.00	12.20
	FAR	1.30	0.10	11.90	1.60	4.70	0.50
	DTD	90.00	206.60	32.00	38.80	38.50	55.00
	J	**111.90**	**253.70**	**51.90**	**51.20**	**53.20**	**67.70**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.37	0.96	19.20	3.21	7.95	0.96
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.37**	**0.96**	**19.20**	**3.21**	**7.95**	**0.96**
J average		**30.87**	**69.35**	**23.89**	**12.17**	**17.09**	**12.19**

To summarize, J has shown that the real process fault is detected effectively using Q and φ, particularly with confidence level 99%. However, large values are provided by the T^2 index.

The time evolution of the fault detection index φ based RKPCA of the real process fault is depicted in Figure 18.

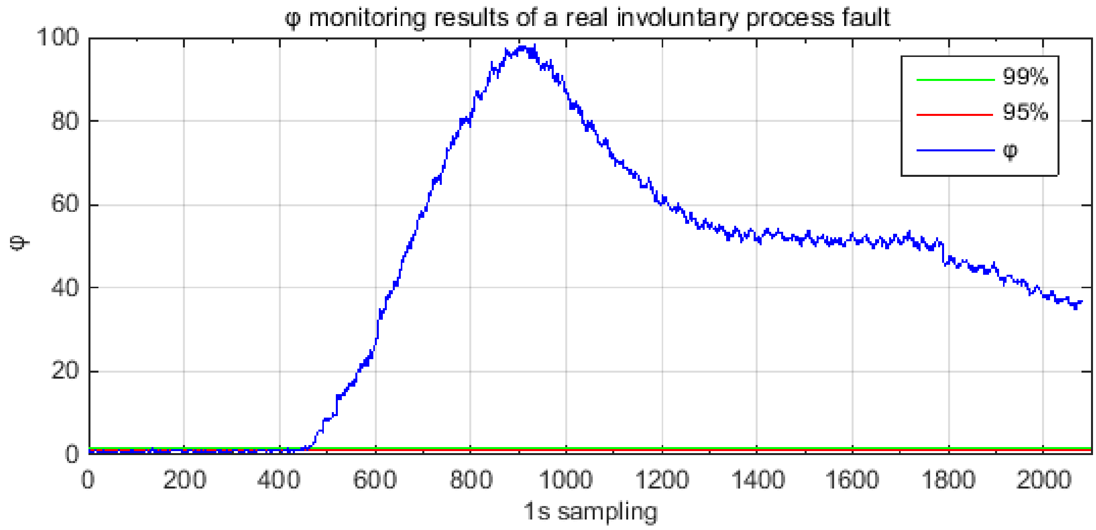

Figure 18. φ monitoring of real involuntary process fault in the cement rotary kiln using the proposed RKPCA.

III.5.3.3. Simulated sensor faults

The performances of the RKPCA based model in detecting all these various faults are summarized in table 14.

From the table, abrupt fault in single sensor (Sfault2) or multiple sensors (Sfault5) and intermittent fault in single sensor (Sfault7) are detected immediately and efficiently with zero missed detected samples and with no delay. In addition, the evaluation using FAR showed effective results.

The T^2 index provided zero detection time with its two confidence limits when detecting the random fault (Sfault1) and with 95% confidence limit when detecting the drift fault in single sensor (Sfault3, Sfault4).As well as, small values when using the remaining indices. In terms of MDR, all of the three monitoring indices (T^2, Q and φ) have detected the Drift fault in single sensor (Sfault3) efficiently (negligible values of MDR), whereas the random fault (Sfault1) and the drift fault in single sensor (Sfault4) has shown large values when T^2 index is used. In terms of FAR, random fault (Sfault1) and drift fault in single sensor (Sfault4) have been detected effectively using the three monitoring indices except Q index with confidence level 95% where a considerable value of FAR is noticed (FAR=18.00% and 10.40% respectively).

The detection of the drift fault in multiple sensors (Sfault6) is delayed to 3min26s and 1min30s when detected using T^2 with confidence level 99% and 95% respectively, whereas slightly delayed when using the remaining indices. In terms of MDR and FAR, large amount of missed detected samples are noticed in the three indices, however, negligible amount of false alarms are shown expect when detected using Q with confidence level 95%.

To recapitulate, the average values of J of the φ index (with its two confidence levels 99% and 95%) and Q (with confidence level 99%) have shown the best performances in detecting all the faults (process fault and 7 simulated sensor faults).

Figure 19 shows the time evolution of the RKPCA based monitoring indices, where it is possible to detect single as well as multiple sensor faults of abrupt, random, intermittent, and drift types.

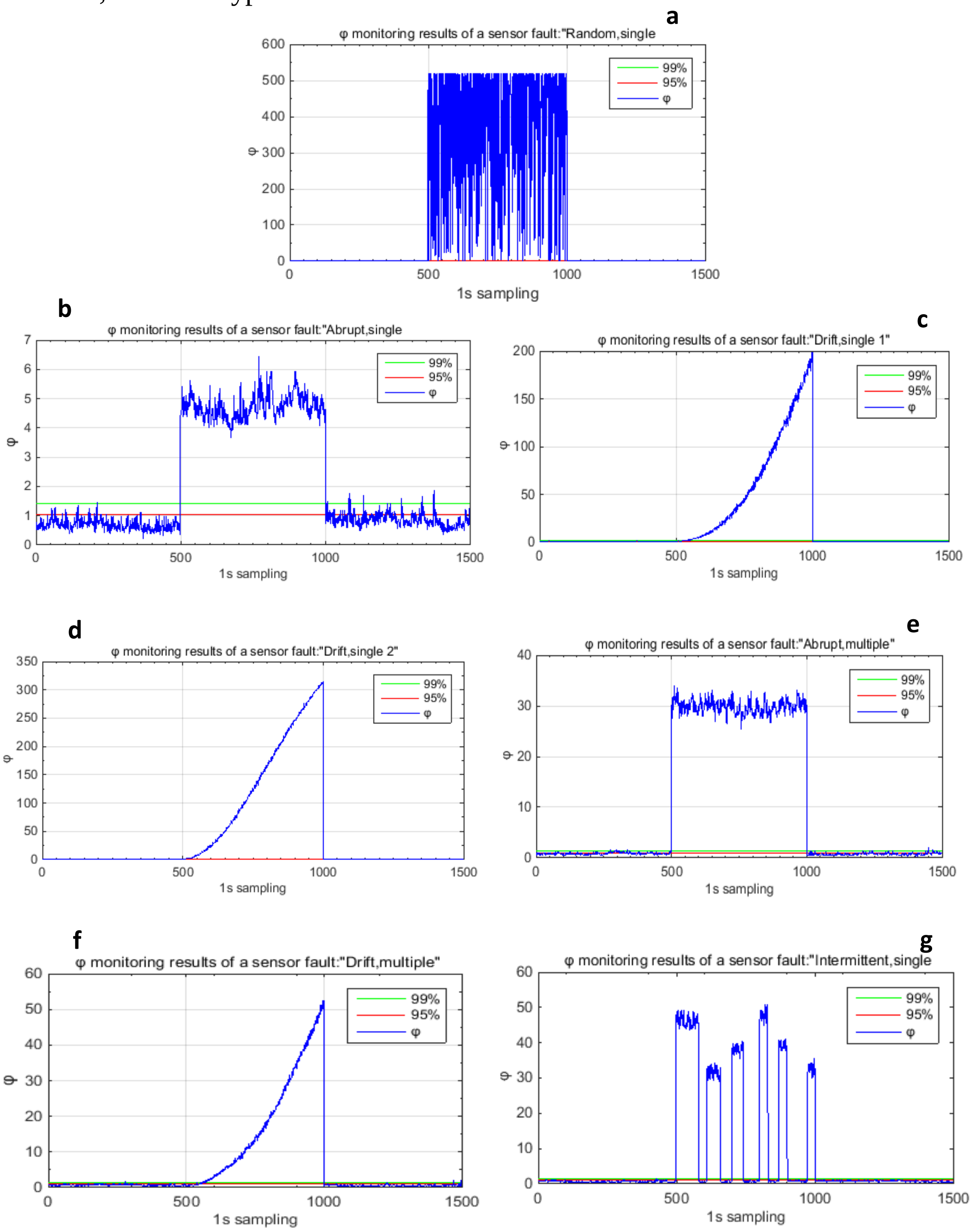

Figure 19. φ monitoring results of sensor faults using RKPCA. **(a)** Sfault1; **(b)** Sfault2; **(c)** Sfault3; **(d)** Sfault4; **(e)** Sfault5; **(f)** Sfault6; **(g)** Sfault7.

III.5.4. Comparison between PCA, KPCA and the proposed approach RKPCA

Based on tables 08, 10 and 13 the proposed RKPCA approach has achieved the best reduction in the execution time compared to KPCA technique. The computational time has dropped from 1h37min to only 9min30s.This is due to the reduction in the number of samples from 768 to 144 sample ,which means the number of observations has been reduced by approximately 5 times, and this huge. This reduction in time is very important in nowadays processes.

The performances of the proposed method in terms of FAR using the three indices (in testing set) has also provided good results when the confidence level 99% is used, as well as the PCA technique. However, KPCA technique has provided the best FAR reduction compared to PCA and RKPCA. This implies that KPCA technique has dealt with the non-captured nonlinearities when PCA is used, and the loss of information resulted from the reduction in RKPCA.

Based on table 15 and compared to PCA and KPCA, RKPCA based φ index has provided the best average J reduction. The bold values highlight the best performance of the proposed method.

The proposed approach can perfectly detect all the faults with an average J of 12.17, 17.09 and 12.19 when Q with confidence level 99% and φ index (with confidence levels 95% and 99%) are used respectively. Whereas, it has some difficulties when the detection is performed using T^2 index.

Table 15. A comparative table between PCA, KPCA and the proposed RKPCA using the average J value contributed by T^2, Q and φ.

Method	index	Confidence level	Average J
PCA	T^2	95%	18.32
		99%	40.03
	Q	95%	22.43
		99%	17.51
	φ	95%	23.08
		99%	17.87
KPCA	T^2	95%	18.14
		99%	46.04
	Q	95%	23.07
		99%	16.44
	φ	95%	22.90
		99%	17.59
RKPCA	T^2	95%	30.87
		99%	69.35
	Q	95%	23.89
		99%	**12.17**
	φ	95%	**17.09**
		99%	**12.19**

III.6. conclusion

In this part of work, PCA, KPCA and the proposed approach RKPCA were applied to a set of healthy and faulty data obtained from a cement plant. The three methods were assessed in terms of FAR, MDR, DTD, the cost function J and the execution time. The results provided by the proposed methodology has proven its capability in reducing the computational time problem introduced by KPCA.As well as its efficiency to detect different types of faults (Abrupt , Drift, Intermittent).

GENERAL CONCLUSION

Fault detection and diagnosis is an important problem in process engineering nowadays. Therefore, early detection and diagnosis of process faults while the plant is still operating can help avoid abnormal event progression and reduce productivity loss. Several methods were developed and have been suggested to deal with the fault detection problem. Method based on multivariate statistical process control (MSPC) are among the most efficient techniques that have seen growth in the last decade.

In this work, PCA, KPCA and the suggested approach RKPCA were used as multivariate statistical methods to monitor the cement rotary kiln system. The monitoring was based on the common used indices: Hotelling's T^2 and Q, in addition to a new proposed index called the combined index φ.

KPCA technique was first introduced to cope with the linear assumption of the PCA. It has provided negligible values in terms of FAR during the healthy operation and good results in detecting the several faults compared to PCA method .However, it has shown a big problem in terms of time consuming.

The new RKPCA approach was proposed to deal with the high computational time resulted from using KPCA method. The idea behind the novel approach was to reduce the number of samples in the data matrix. The three methods PCA, KPCA and RKPCA were compared to each other in terms of the performances: FAR, MDR, DTD, the cost function J and the execution time. The proposed approach has shown the best results in reducing the execution time and its ability to efficiently detect all the faults with φ index (with its two confidence levels 95% and 99%) and Q index using confidence level 99%. Whereas, it has some difficulties when the detection is performed using T^2 index.

Further improvements to the performance of the proposed approach may be obtained by improving the choice of the bandwidth used in the kernel function. Additionally, parameters such as the number of the retained principal component may be tuned based on the process considered to obtain a better performances in terms of false alarm and missed detection rates.

REFERENCES

[1] S.J.Qin, *"Statistical process monitoring: basics and beyond",* J.Chemom, VOL.17, PP.480-502, 2003.

[2] T.J.McAvoy*, "Intelligent "Control" applications in the process industries"*, Annu Rev Control, VOL.26, PP.75-86, 2002.

[3]. V.Venkatasubramanian, R. Rengaswamy, K. Yin Surya N. Kavuri, "*A review of process fault detection and diagnosis Part I: Quantitative model-based methods",* Computers and Chemical Engineering, vol. 27, pp.293-311, 2003.

[4] C.Aldrich, L.Auret, "Chapter 2: Overview of process fault diagnosis", in *Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods*, Ed.London: Springer, PP.17-70, 2013.

[5] M.Kano, Sh.Hasebe, I.Hashimoto, H.Ohna, *"A new multivariate statistical process monitoring method using principal component analysis"*, Computers and Chemical Engineering, Vol.25, PP.1103-1113, 2001.

[6] E.Russell, L. Chiang, R. Braatz, "*Data-driven methods for fault detection and diagnosis in chemical processes"*, Springer-Verlag London, 2000.

[7] L.Chiang, E.Russell, R.Braatz, "Fault *detection and diagnosis in industrial systems"*, Springer-Verlag, London, 2001.

[8] S.Yin, X. Li, H.Gao, O. Kaynak, *"Data-based techniques focused on modern industry: An overview",* IEEE Transactions on Industrial Electronics, VOL.62 (1), PP. 657–667.

[9] I.Jolliffe, "*Principal component analysis* (2nd ed)", New York, NY: Springer-Verlag, (2002).

[10] S.Wold, K. Esbensen, P.Geladi, *"Principal component analysis",* Chemometrics and Intelligent Laboratory Systems, VOL.*2* (1–3), PP.37–52, 1987.

[11] B.M.Wise and N.B Gallagher, "*The process Chemometrics approach to process monitoring and fault detection"*, Journal of Process Control, VOL.*6*, PP. 329–348, (1996).

[12]A.M Kramer, "*Auto-associative neural networks",* Computers and Chemical Engineering, VOL.16 (4), PP.313–328, 1992.

[13] D.Dong, T.J.McAvoy, *"Nonlinear principal component analysis – Based on nonlinear principal curves and neural networks",* In Proceedings of the American Control Conference. American Control Conference, Baltimore, MD, USA, PP. 1284–1288, 1996.

[14] H.G.Hiden, M.J.Willis, T.M.Tham, A.G.Montague, *"Nonlinear principal component analysis using genetic programming",* Computer and Chemical Engineering, VOL.23, PP.413-425, 1999.

[15] C.Cheng, M.S.Chiu, *"Nonlinear process monitoring using JITL-PCA",* Chemometrics Intelligent Laboratory Systems, VOL.76, PP.1-13, 2005.

[16] U.Kruger, D.Antory, J.Hahn, G.W.Irwin, G.McCullough, *"Introduction of a nonlinearity measure for principal component models",* Computers and Chemical Engineering, VOL.29, PP.2355-2362, 2005.

[17] A.Maulud, D.Wang, J.A.Romagnoli, *"A multiscale orthogonal nonlinear strategy for multivariate statistical process monitoring",* Journal of Process Control, VOL.16, PP.671-683, 2006.

[18] B.Schölkopf, A.Smala, K.R.Müller, *"Nonlinear component analysis as kernel eigenvalue problem",* Neural Comput, VOL.10, PP.1299-1319, 1998.

[19] S.Mika, B.Schölkopf, A.Smola, K.R Müller, M.Scholz, G.Rätsch, *"Kernel PCA and de-noising in feature spaces",* Advances in Neural Information Processing Systems, VOL.11, PP.536-542, 1999.

[20] S.Romdhani, S.Gong, A. Psarrou, *"A multi-view nonlinear active shape model using kernel PCA",* Proceedings of BMVC, Nottingham,UK, PP. 483 492, 1999.

[21] K.Kazor, R.W.Holloway, T.Y.Coth, A.S.Hering, *"Comparison of linear and nonlinear dimension reduction techniques for automated process monitoring of a decentralized Wastewater treatment facility"*, Stoch Environ Res Risk Assess, VOL.33, PP.1544-1544, 2016.

[22] L.Zhong, T.Yang, J.Yi, R.Jin, Z.H.Zhou, *"stochastic optimization for Kernel PCA",* AAAI Conference on Artificial Intelligence, 2016.

[23] D.Lopez-Paz, S.Sra,A. Smola,Z.Ghahramani, B.Schölkopf, *"Randomized nonlinear component analysis"*, In Proceedings of the 31st International Conference on Machine Learning, 2014.

[24] K.I.Kim, M.O.Franz, B.Schölkopf, *"Iterative kernel principal component analysis for image modeling"*, IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL.27, PP.1351–1366, 2005.

[25] H. Yue and S. Joe Qin, *"Reconstruction based fault identification using a combined index",* Ind. Eng. Chem. Res, VOL.40, PP.4403–4414, 2001.

[26] J.M.Lee, C.Yoo, SK.Choi, PA.Vanrolleghemb, I. Lee, *"Nonlinear process monitoring using Kernel principal component analysis*", Chem Eng Sci, VOL.59, PP.223–234, 2004.

[27] O.Taouali, I.Jaffel, H.Lahdhiri, M.Faouzi Harkat, H.Messaoud, "*New fault detection method based on reduced kernel principal component analysis (RKPCA)",* Int J Adv Manuf Technol ,2015.

[28]. D.M.Himmelblau, *"Fault Detection and Diagnosis in Chemical and Petrochemical Processes",* Elvisar Press, Amsterdam, Himmelblau, 1978

[29]. R.Iserman, *"Model-based Fault Detection and Diagnosis - Status and Applications",* Annual Reviews in Control, Vol.29, PP. 71–85, 2005

[30]. S. Katipamula, M.R. Brambley, "*Methods for fault detection, diagnostics and prognostics for building systems—a review, Part I"*, HVAC&R Research, Vol.12 (2), PP. 169–187, 2005.

[31]. H.Cheng, M.Nikus, Sirkka-Liisa Jämsä-Jounela, *"Application of a causal digraph based fault diagnosis method with discrete state space model on a paper of machine simulator",* Helsinki University of Technology Laboratory of Process Control and Automation Kemistintie 1, FI-02150 HUT, Finland.

[32]. E.E.HURDLE, L.M.BARTLETT and J.D.ANDREWS, *"System fault diagnostics using fault tree analysis. IN: Proceedings of the 16th Advances in Reliability Technology Symposium (ARTS)"*, Loughborough, UK, PP. 203-222, 2005.

[33] Q. YANG, *"Model-Based and Data Driven Fault Diagnosis Methods With Applications to Process Monitoring",* PhD Thesis, Electrical Engineering and Computer Sciences, Case Western Reserve University, Ohio, USA, 2004.

[34] Y.IWASAKI and H.A.SIMON, *"Causality in device behavior. Artificial Intelligence",* 29(1):3 – 32, 1986.

[35]. Alkan Alkaya, "*Novel Data Driven-Based Fault Detection for Electromechanical and Process Control Systems*", Ph.D. thesis, Çukurova University, Institute of natural and applied sciences Department of electrical and electronics engineering, 2012

[36]. V.Venkatasubramanian, R. Rengaswamy, S.Kavuri, *"A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies",* Computers & Chemical Engineering, 2003

[37]. V. Venkatasubramanian, R.Rengaswamy, S Kavuri, K.Yin, *"A review of process fault detection and diagnosis: Part III: Process history based methods".* Computers & Chemical Engineering, Vol.46, PP. 27-327, 2003

[38]. J.V.KRESTA, J.F.MACGREGOR and T.E.MARLIN, *"Multivariate Statistical Monitoring of Process Operating Performance".* The Canadian, 1991.

[39]. J.F.MACGREGOR, and T.KOURTI, *"Statistical Process Control of Multivariate Processes. Control Engineering Practice",* 3(3):403 – 414, 1995.

[40]. S.Wold, K.Esbensen, & P.Geladi, *"Principal component analysis»*, Chemometrics and Intelligent Laboratory Systems, *2*(1–3), PP. 37–52, (1987).

[41]. H.OCAK, *"Fault Detection, Diagnosis and Prognosis of Rolling Element Bearings: Frequency Domain Methods and Hidden Markov Modeling".*
PhD thesis, Case Western Reserve University, 2003.

[42]. R.RENGASWAMY and V.VENKATASUBRAMANIAN, *"A fast training neural network and its updation for incipient fault detection and diagnosis".* Computers and Chem. Eng., 24(2-7):431–437, 2000.

[43]. J.C.Benneyan, R. C. Lloyd, P.E. Plsek, "Statistical process control as a tool for research and healthcare improvement", Qual Saf Health Care, 2003.

[44] D.M.Himes, R.H.Storer, C.Georgakis, "*Determination of the number of principal components for disturbance detection and isolation*." In: Proceedings of American control conference, IEEE Press, Piscataway, New Jersey, PP. 1279–1283, 1994

[45] A.Bakdi, A.Kouadri, A.Bensmail, "*Fault detection and diagnosis in a cement rotary kiln using PCA with EWMA-based adaptive threshold monitoring scheme*", Control Engineering Practice, VOL.66, PP.64-75, 2017.

[46] G.Diana, C.Tommasi, *"Cross-validation methods in principal component analysis: a comparison",* Statistical Methods & Applications, VOL.11, PP.71–82, 2002.

[47] J.L.Horn, *"A rationale and test for the number of factors in factor analysis",* Psychometrika, VOL.30, PP.179–185, 1965.

[48] S.B.Franklin, D.J. Gibson, P.A. Robertson, J.T.Pohlmann, & J.S. Fralish, *"Parallel analysis: a method for determining significant principal components", Journal of Vegetation Science*, VOL. *6*, PP.99–106, 1995.

[49] L.Guttman, *"Some necessary conditions for common factor analysis"*, Psychometrika, VOL.*19*, PP. 149–161.

[50] M.Mansouri, M.F.Destain, H. Nounou, M.Nounou, *"Enhanced monitoring of environmental processes".* Int. J. Environ. Sci. Dev, 7 (7), 525, 2016.

[51] J.E.Jackson, *"Quality control methods for several related variables*", Technometrics, PP.359–377, 1959

[52] F.Carlos, Alcala, S. Joe Qin, "*Reconstruction-based Contribution for Process Monitoring with Kernel Principal Component Analysis",* American Control Conference Marriott Waterfront, Baltimore, MD, USA ,June 30-July 02, 2010.

[53] B.Schölkopf, A.Smala, Müller.K.R, *"Kernel principal component",* PP.583-588, 1997.

[54] S.Haykin, "*Neural Networks",* Prentice-Hall, Englewood Cliffs, 1999

[55] M.Kallas, G.Mourat, K.Anani, J.Ragot, D.Maquin, *"Fault detection and estimation using kernel principal component"*, IFAC papersOnline, VOL.50, PP.1025-1030, 2017.

[56] N.Christianini, J.Shawe-Taylor, *"An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods",* Cambridge university press, UK, 2000.

[57] R.T.Samuel, Y.Cao, *"Nonlinear process fault detection and identification using kernel PCA and kernel density estimation",* Systems science and control engineering, VOL.4, PP.165-174, 2016.

[58] http://www.imm.dtu.dk/~perbb/MAS/ST116/module02/index.html

[59] https://en.wikipedia.org/wiki/Euclidean_distance

[60] S.Masoud, A.Fatehi, *"Identification, Prediction and Detection of process Fault in a Cement Rotary Kiln by Locally Linear Neuro-Fuzzy Technique",*Wolrd Academy of Science, Engineering and Technology, International Journal of mechanical and mechatronics Engineering, VOL.3, 2009.

[61] MM.Y.Chow, R.N.Sharpe, *"On the Application and Design of Artificial Neural Networks for Motor Fault Detection-Part II",* IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL.40, Aplril.1993.

[62] https://www.geeksforgeeks.org/measure-execution-time-with-high-precision-in-c-c/

APPENDIX

TABLE 01

The table below shows the FAR values contributed by T², Q and φ with confidence levels 95% and 99% using the 53 RKPCA models. In addition, the number of samples and the time execution of each model are given.

These FAR values are depicted in Figure 16 in order to obtain the minimum FAR values and their relative distances provided by the testing set.

The values in bold highlight the seven selected distances with their corresponding execution time and the number of samples obtained for each model.

			FAR for training data						FAR for testing data					
			T²		Q		φ		T²		Q		Φ	
Distances	TC (min)	N° of samples	95%	99%	95%	99%	95%	99%	95%	99%	95%	99%	95%	99%
1.8484	56.60	767	5.10	0.75	4.20	0.60	4.50	0.60	4.38	0.37	6.09	0.82	5.46	0.56
2.3127	54.26	764	4.97	1.05	4.97	1.05	4.97	1.05	4.35	0.37	5.81	0.84	5.41	0.55
2.7010	52.79	758	5.01	1.06	5.01	1.06	5.01	1.06	4.38	0.36	5.70	0.88	5.41	0.55
2.8357	51.97	754	5.04	1.06	5.04	1.06	5.04	1.06	4.35	0.35	5.69	0.87	5.46	0.55
2.8732	51.60	750	4.93	0.93	4.93	0.93	4.93	0.93	4.35	0.35	5.74	0.87	5.42	0.55
2.9219	49.74	743	4.98	0.94	4.98	0.94	4.98	0.94	4.27	0.35	5.55	0.88	5.44	0.44
2.9732	49.20	737	5.02	0.95	5.02	0.95	5.02	0.95	4.18	0.34	4.98	0.95	5.45	0.46
3.0485	49.07	721	4.99	0.97	4.99	0.97	4.99	0.97	4.14	0.35	4.93	0.94	5.31	0.38
3.0637	48.88	719	5.01	0.97	5.01	0.97	5.01	0.97	4.15	0.35	5.02	0.94	5.32	0.38
3.0851	48.61	714	5.04	0.98	5.04	0.98	5.04	0.98	4.14	0.35	5.05	0.93	5.31	0.38
3.0978	48.20	708	4.94	0.99	4.94	0.99	4.94	0.99	4.07	0.35	5.07	0.89	5.35	0.35
3.1096	47.76	705	4.96	0.99	4.96	0.99	4.96	0.99	4.07	0.35	5.04	0.86	5.46	0.35
3.1812	47.66	693	5.05	1.01	5.05	1.01	5.05	1.01	4.15	0.35	5.05	0.85	5.35	0.33
3.2135	47.42	686	4.96	1.02	4.96	1.02	4.96	1.02	4.24	0.34	4.78	0.86	5.17	0.30
3.2630	45.25	672	5.06	1.04	5.06	1.04	5.06	1.04	4.21	0.40	4.85	0.89	5.21	0.29
3.3139	44.20	663	4.98	1.06	4.98	1.06	4.98	1.06	3.95	0.38	4.55	0.86	5.20	0.27
3.3765	43.97	643	4.98	0.93	4.98	0.93	4.98	0.93	3.64	0.33	4.45	0.86	4.98	0.24
3.4289	43.46	623	4.98	0.96	4.98	0.96	4.98	0.96	3.68	0.30	4.75	0.77	4.96	0.25
3.4923	42.04	601	4.99	1.00	4.99	1.00	4.99	1.00	3.70	0.25	3.95	0.66	4.80	0.21
3.5338	41.03	574	5.05	1.05	5.05	1.05	5.05	1.05	3.74	0.20	3.96	0.82	4.95	0.21
3.5937	40.17	545	4.95	0.92	4.95	0.92	4.95	0.92	3.54	0.19	3.88	0.77	4.82	0.17
3.6185	38.24	538	5.02	0.93	5.02	0.93	5.02	0.93	3.38	0.19	3.90	0.80	4.78	0.17
3.6516	37.53	524	4.96	0.95	4.96	0.95	4.96	0.95	3.20	0.19	3.75	0.59	4.58	0.15

3.6858	31.48	502	4.98	1.00	4.98	1.00	4.98	1.00	3.40	0.18	4.25	0.31	4.67	0.14
3.7568	31.24	463	4.97	1.08	4.97	1.08	4.97	1.08	2.95	0.18	3.63	0.33	4.43	0.13
3.7831	**30.08**	**447**	**4.70**	**0.89**	**4.92**	**0.89**	**4.70**	**0.89**	**2.84**	**0.18**	**3.30**	**0.35**	**4.90**	**0.08**
3.8078	26.34	431	5.10	0.93	5.10	0.93	5.10	0.93	3.04	0.18	3.82	0.18	4.78	0.11
3.8147	25.65	427	4.92	0.94	4.92	0.94	4.92	0.94	3.31	0.17	3.93	0.20	4.94	0.11
3.9461	**23.88**	**359**	**5.01**	**1.11**	**5.01**	**1.11**	**5.01**	**1.11**	**2.24**	**0.11**	**4.89**	**0.10**	**4.55**	**0.09**
4.0147	23.17	331	5.14	0.91	5.14	0.91	5.14	0.91	2.24	0.08	4.73	0.12	4.32	0.07
4.1115	**22.66**	**298**	**5.03**	**1.01**	**5.03**	**1.01**	**5.03**	**1.01**	**1.62**	**0.01**	**4.62**	**0.10**	**3.92**	**0.06**
4.1564	**22.09**	**281**	**4.98**	**1.07**	**4.98**	**1.07**	**4.98**	**1.07**	**1.72**	**0.00**	**5.44**	**0.20**	**4.97**	**0.05**
4.2062	**20.28**	**267**	**4.87**	**1.12**	**4.87**	**1.12**	**4.87**	**1.12**	**1.53**	**0.00**	**5.60**	**0.46**	**5.07**	**0.05**
4.2707	18.53	253	5.14	1.19	5.14	1.19	5.14	1.19	1.07	0.00	7.82	0.75	5.85	0.09
4.2865	18.24	243	4.94	0.82	4.94	0.82	4.94	0.82	1.23	0.00	8.36	1.04	5.33	0.12
4.2974	18.16	238	5.04	0.84	5.04	0.84	5.04	0.84	1.25	0.00	7.64	1.11	5.55	0.10
4.3069	17.94	236	5.08	0.85	5.08	0.85	5.08	0.85	1.13	0.00	7.85	1.17	5.38	0.08
4.3153	16.42	233	5.15	0.86	5.15	0.86	5.15	0.86	1.08	0.00	8.37	0.88	4.67	0.07
4.3253	15.82	229	4.80	0.87	4.80	0.87	4.80	0.87	1.27	0.00	9.62	0.88	6.05	0.10
4.3413	15.56	226	4.87	0.88	4.87	0.88	4.87	0.88	1.25	0.00	9.29	1.10	6.26	0.10
4.3543	14.10	216	5.09	0.93	5.09	0.93	5.09	0.93	1.29	0.00	10.06	0.69	6.89	0.06
4.3626	13.97	208	4.81	0.96	4.81	0.96	4.81	0.96	1.15	0.00	9.25	1.18	6.39	0.07
4.3791	13.78	205	4.88	0.98	4.88	0.98	4.88	0.98	1.25	0.00	8.41	1.15	7.19	0.06
4.3900	13.47	199	5.03	1.01	5.03	1.01	5.03	1.01	1.33	0.00	9.41	1.84	5.54	0.09
4.4324	11.81	185	4.86	1.08	4.86	1.08	4.86	1.08	1.08	0.00	12.75	2.26	7.57	0.07
4.4592	11.38	175	5.14	1.14	5.14	1.14	5.14	1.14	1.23	0.00	14.78	1.99	7.19	0.06
4.5017	10.44	165	4.85	1.21	4.85	1.21	4.85	1.21	0.93	0.00	10.81	5.15	6.91	0.24
4.5379	10.32	160	5.00	1.25	5.00	1.25	5.00	1.25	0.85	0.00	11.08	3.52	6.03	0.18
4.6082	**9.51**	**144**	**4.86**	**0.69**	**4.86**	**0.69**	**4.86**	**0.69**	**1.02**	**0.00**	**20.55**	**3.75**	**9.31**	**0.88**
4.6555	8.54	133	5.26	0.75	5.26	0.75	5.26	0.75	1.40	0.00	19.88	7.18	13.21	0.95
4.7400	7.78	120	5.00	0.83	5.00	0.83	5.00	0.83	1.01	0.00	19.38	10.09	11.25	1.39
4.7867	**6.82**	**108**	**4.63**	**0.93**	**4.63**	**0.93**	**4.63**	**0.93**	**0.75**	**0.00**	**31.86**	**18.95**	**21.04**	**3.77**
4.8418	6.03	96	5.21	1.04	5.21	1.04	5.21	1.04	0.77	0.00	38.25	17.02	27.79	4.81

TABLE 2-8

The 7 tables below represent the performances of each of the 7 selected models in detecting the different types of faults in terms of: Missed Detection Rate (MAR), False Alarm Rate (FAR), Detection Time Delay (DTD) and The cost function J.

An average J value of the 8 faults is calculated for each model (The values in bold).

Table2. The MDR, FAR, DTD and J contributed by T^2, Q and φ of the model that matches: Distance 1=3.7831 , number of samples=447

Faults	Performances	T^2		Q		φ	
		95%	99%	95%	99%	95%	99%
Process fault	MDR	1.26	2.58	0.36	1.62	0.00	2.52
	FAR	18.57	0.00	53.81	13.10	80.48	0.00
	DTD	1.25	43.00	0.00	2.00	0.00	33.50
	J	**21.08**	**45.58**	**54.17**	**16.72**	**80.48**	**35.02**
Random and single	MDR	2.00	45.00	1.20	1.40	1.20	2.00
	FAR	0.70	0.10	4.10	0.50	3.70	0.20
	DTD	0.60	2.00	1.00	1.00	1.00	1.00
	J	**3.30**	**47.10**	**6.30**	**2.90**	**5.90**	**3.20**
Abrupt and single	MDR	0.00	0.00	0.20	5.20	0.00	0.00
	FAR	1.40	0.70	2.70	0.80	3.10	0.40
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.40**	**0.70**	**2.90**	**6.00**	**3.10**	**0.40**
Drift and single 1	MDR	4.60	7.20	4.60	9.40	3.00	7.80
	FAR	4.20	0.40	1.30	0.20	5.30	0.10
	DTD	18.70	27.50	5.00	36.00	5.50	36.00
	J	**27.50**	**35.10**	**10.90**	**45.60**	**13.80**	**43.90**
Drift and single 2	MDR	12.80	22.80	1.60	2.40	1.40	0.60
	FAR	0.60	0.10	1.20	0.10	1.00	0.10
	DTD	48.00	113.30	1.80	10.50	0.20	17.50
	J	**61.40**	**136.2**	**4.60**	**13.00**	**2.60**	**18.20**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.60	0.20	4.90	0.80	4.40	0.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.60**	**0.20**	**4.90**	**0.80**	**4.40**	**0.20**
Drift and multiple	MDR	18.00	27.40	13.20	16.40	12.20	20.00
	FAR	1.80	0.70	5.60	0.90	5.50	0.50
	DTD	85.00	120.00	55.00	71.80	55.00	98.00
	J	**104.80**	**148.10**	**73.80**	**89.10**	**72.70**	**118.50**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	3.53	0.96	4.98	1.04	7.15	0.96
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**3.53**	**0.96**	**4.98**	**1.04**	**7.15**	**0.96**
J average		**27.98**	**51.74**	**20.31**	**21.89**	**23.76**	**27.54**

Table3. The MDR, FAR, DTD and J contributed by T^2, Q and φ of the model that matches: Distance=3.9461, number of samples=359.

Faults	Performances	T^2		Q		φ	
		95%	99%	95%	99%	95%	99%
Process fault	MDR	1.20	2.58	0.24	2.76	0.00	2.52
	FAR	18.81	0.00	53.57	0.95	78.10	0.24
	DTD	3.50	43.00	0.00	0.00	0.00	33.50
	J	**24.51**	**45.58**	**53.81**	**3.71**	**78.10**	**36.26**
Random and single	MDR	1.80	43.20	1.20	1.80	1.20	2.00
	FAR	0.80	0.20	4.10	0.20	3.20	0.20
	DTD	0.75	2.00	1.00	1.00	1.00	1.00
	J	**3.35**	**45.40**	**6.30**	**3.00**	**5.40**	**3.20**
Abrupt and single	MDR	0.00	0.00	0.00	0.60	0.00	0.00
	FAR	1.40	0.40	3.70	0.40	2.80	0.40
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.40**	**0.40**	**3.70**	**1.00**	**2.80**	**0.40**
Drift and single 1	MDR	5.00	7.40	4.20	10.80	2.80	7.20
	FAR	3.00	0.10	1.90	0.10	3.80	0.10
	DTD	18.60	27.50	2.00	43.00	7.00	35.00
	J	**26.60**	**35.00**	**8.10**	**53.90**	**13.60**	**42.30**
Drift and single 2	MDR	12.60	21.60	1.60	3.00	1.40	6.00
	FAR	0.30	0.10	1.00	0.10	0.90	0.10
	DTD	1.00	71.80	1.50	10.90	0.50	17.35
	J	**13.90**	**93.50**	**4.10**	**14.00**	**2.80**	**23.45**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.80	0.20	7.90	0.20	5.00	0.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.80**	**0.20**	**7.90**	**0.20**	**5.00**	**0.20**
Drift and multiple	MDR	18.40	28.40	11.20	18.00	12.00	19.40
	FAR	1.70	0.70	10.00	0.60	5.01	5.00
	DTD	5.50	133.60	50.80	88.50	55.00	95.00
	J	**25.60**	**162.70**	**72.00**	**107.10**	**72.01**	**119.40**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	2.97	0.96	5.94	0.96	6.75	0.96
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**2.97**	**0.96**	**5.94**	**0.96**	**6.75**	**0.96**
J average		**12.39**	**47.96**	**20.23**	**22.98**	**23.30**	**28.27**

Table 4. The MDR, FAR, DTD and J contributed by T^2, Q and φ of the model that matches: Distance= 4.1115 , number of samples=298

Faults	Performances	T^2		Q		φ	
		95%	**99%**	**95%**	**99%**	**95%**	**99%**
Process fault	MDR	1.38	2.64	0.30	2.88	0.00	2.58
	FAR	24.29	0.00	50.71	0.48	77.38	0.00
	DTD	1.80	43.00	0.00	3.00	0.00	33.90
	J	**27.47**	**45.64**	**51.01**	**6.36**	**77.38**	**36.48**
Random and single	MDR	2.00	43.40	1.20	1.80	1.20	2.00
	FAR	0.80	0.20	4.20	0.20	3.90	0.20
	DTD	0.80	0.30	1.00	1.00	1.00	1.00
	J	**3.60**	**43.90**	**6.40**	**3.00**	**6.10**	**3.20**
Abrupt and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.20	0.20	4.70	0.40	3.00	0.40
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.20**	**0.20**	**4.70**	**0.40**	**3.00**	**0.40**
Drift and single 1	MDR	4.60	7.40	4.20	11.00	3.00	7.60
	FAR	2.90	0.10	5.50	0.10	3.50	0.10
	DTD	18.50	27.70	5.50	46.00	8.50	35.00
	J	**26.00**	**35.20**	**15.20**	**57.10**	J=15.00	J=42.70
Drift and single 2	MDR	11.40	21.60	1.60	2.80	1.40	6.00
	FAR	0.30	0.10	1.00	0.20	0.70	0.10
	DTD	0.00	71.80	1.50	10.50	1.3	17.40
	J	**11.70**	**93.50**	**4.10**	**13.50**	**3.40**	**23.50**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.50	0.20	6.40	0.40	4.60	0.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.50**	**0.20**	**6.40**	**0.40**	**4.60**	**0.20**
Drift and multiple	MDR	18.20	31.60	10.60	17.20	11.60	19.60
	FAR	1.60	0.30	10.70	0.40	4.60	0.40
	DTD	55.00	135.00	38.70	88.00	55.00	98.00
	J	**74.80**	**166.90**	**60.00**	**105.60**	**71.20**	**118.00**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	2.57	0.96	5.06	0.96	6.18	0.96
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**2.57**	**0.96**	**5.06**	**0.96**	**6.18**	**0.96**
J average		**18.48**	**48.31**	**19.10**	**23.41**	**23.35**	**28.18**

Table 5. The MDR, FAR, DTD and J contributed by T^2, Q and φ of the model that matches: Distance=4.1564, number of samples=281

Faults	Performances	T^2		Q		φ	
		95%	99%	95%	99%	95%	99%
Process fault	MDR	1.50	2.64	0.18	1.98	0.00	2.46
	FAR	23.57	0.00	51.19	2.14	80.00	0.00
	DTD	0.25	43.00	1.50	2.00	0.00	33.60
	J	**25.32**	**45.64**	**52.87**	**6.12**	**80.00**	**35.46**
Random and single	MDR	2.00	43.40	1.20	1.80	1.20	2.00
	FAR	0.80	0.20	5.20	0.20	4.60	0.20
	DTD	0.00	0.35	1.00	1.00	1.00	1.00
	J	**2.80**	**43.95**	**7.40**	**3.00**	**6.80**	**3.20**
Abrupt and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.10	0.20	5.80	0.50	4.00	0.40
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.10**	**0.20**	**5.80**	**0.50**	**4.00**	**0.40**
Drift and single 1	MDR	4.60	7.40	4.00	9.40	2.80	7.40
	FAR	3.70	0.10	2.30	0.10	5.00	0.10
	DTD	18.70	36.50	2.80	42.00	6.00	36.00
	J	**27.00**	**44.00**	**9.10**	**51.50**	**13.80**	**43.50**
Drift and single 2	MDR	12.20	23.60	1.60	2.40	1.60	4.80
	FAR	0.30	0.10	1.40	0.30	1.10	0.10
	DTD	0.00	113.50	7.50	10.50	1.30	12.90
	J	**12.50**	**137.20**	**10.50**	**13.20**	**4.00**	**17.80**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.50	0.20	7.40	0.50	4.90	0.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.50**	**0.20**	**7.40**	**0.50**	**4.90**	**0.20**
Drift and multiple	MDR	18.40	33.60	10.20	15.20	10.80	19.40
	FAR	1.40	0.10	11.60	0.60	5.50	0.30
	DTD	55.00	135.70	38.80	70.00	48.00	95.00
	J	**74.80**	**169.40**	**60.60**	85.80	**64.30**	**114.70**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	2.33	0.96	5.38	1.04	6.18	0.96
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**2.33**	**0.96**	**5.38**	**1.04**	**6.18**	**0.96**
J average		**18.29**	**55.19**	**19.88**	**20.20**	**22.99**	**27.02**

Table 6. The MDR, FAR, DTD and J contributed by T^2, Q and φ of the model that matches: Distance=4.2062 , number of samples=267

Faults	**Performances**	**T^2**		**Q**		**φ**	
		95%	**99%**	**95%**	**99%**	**95%**	**99%**
Process fault	MDR	1.14	2.64	0.30	1.86	0.00	2.46
	FAR	41.90	0.00	43.33	1.43	71.90	0.00
	DTD	3.50	43.00	2.00	2.50	0.00	33.60
	J	**46.54**	**45.64**	**45.63**	**5.79**	**71.90**	**36.06**
Random and single	MDR	30.40	43.20	1.20	1.60	1.20	2.00
	FAR	0.70	0.20	5.50	0.30	4.70	0.20
	DTD	0.00	0.50	1.00	1.00	1.00	1.00
	J	**31.10**	**43.90**	**7.70**	**2.90**	**6.90**	**3.20**
Abrupt and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.90	0.20	6.40	1.50	4.40	0.40
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.90**	**0.20**	**6.40**	**1.50**	**4.40**	**0.40**
Drift and single 1	MDR	4.60	7.60	3.40	7.40	2.80	7.40
	FAR	3.20	0.10	2.30	0.10	4.40	0.10
	DTD	18.80	38.50	5.50	27.30	8.50	27.50
	J	**26.60**	**46.20**	**11.20**	**34.8**	**15.70**	**35.00**
Drift and single 2	MDR	13.20	25.00	1.60	2.00	1.60	4.40
	FAR	0.30	0.10	1.60	0.30	1.20	0.10
	DTD	0.00	118.00	1.50	1.90	1.35	10.90
	J	**13.50**	**143.10**	**4.70**	**4.20**	**4.15**	**15.40**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.50	0.20	8.00	0.90	5.90	0.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.50**	**0.20**	**8.00**	**0.90**	**5.90**	**0.20**
Drift and multiple	MDR	18.40	34.60	10.20	14.00	10.80	18.20
	FAR	1.40	0.10	9.40	0.70	5.80	0.30
	DTD	55.00	136.00	38.70	63.00	39.00	88.70
	J	**74.80**	**170.70**	**58.30**	**77.70**	**55.60**	**107.20**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	2.33	0.96	5.54	1.12	5.94	0.96
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**2.33**	**0.96**	**5.54**	**1.12**	**5.94**	**0.96**
J average		**24.53**	**56.36**	**18.43**	**16.11**	**21.31**	**24.80**

Table 7. The MDR, FAR, DTD and J contributed by T^2, Q and φ of the model that matches: Distance=4.6082 , number of samples=144.

Faults	Performances	T^2		Q		φ	
		95%	99%	95%	99%	95%	99%
Process fault	MDR	0.60	3.31	1.02	1.92	0.60	2.04
	FAR	54.29	0.00	13.10	0.95	20.71	1.19
	DTD	4.80	55.00	3.00	3.90	0.00	0.00
	J	**59.69**	**58.31**	**17.12**	**6.77**	**21.31**	**3.23**
Random and single	MDR	30.80	44.40	1.00	1.60	1.00	1.80
	FAR	1.30	0.20	18.00	1.80	9.10	0.80
	DTD	0.00	0.00	1.00	1.00	1.00	1.00
	J	**32.10**	**44.60**	**20.00**	**4.40**	**11.10**	**3.60**
Abrupt and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.00	0.20	25.10	4.10	11.40	1.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.00**	**0.20**	**25.10**	**4.10**	**11.40**	**1.20**
Drift and single 1	MDR	4.80	8.80	2.40	7.00	2.80	5.80
	FAR	3.30	0.10	7.30	0.40	2.60	0.10
	DTD	22.00	42.00	0.50	8.30	5.00	8.70
	J	**30.10**	**50.90**	**10.20**	**15.70**	**10.40**	**14.60**
Drift and single 2	MDR	10.00	27.00	1.20	1.80	1.20	1.80
	FAR	0.50	0.10	10.40	0.90	3.30	0.30
	DTD	0.00	118.9	1.30	1.74	1.30	1.95
	J	**10.50**	**146.00**	**12.90**	**4.44**	**5.80**	**4.05**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.30	0.20	34.70	7.60	15.60	2.20
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.30**	**0.20**	**34.70**	**7.60**	**15.60**	**2.20**
Drift and multiple	MDR	20.60	47.00	8.00	10.80	10.00	12.20
	FAR	1.30	0.10	11.90	1.60	4.70	0.50
	DTD	90.00	206.60	32.00	38.80	38.50	55.00
	J	**111.90**	**253.70**	**51.90**	**51.20**	**53.20**	**67.70**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	1.37	0.96	19.20	3.21	7.95	0.96
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**1.37**	**0.96**	**19.20**	**3.21**	**7.95**	**0.96**
J average		**30.87**	**69.35**	**23.89**	**12.17**	**17.09**	**12.19**

Table 8. The MDR, FAR, DTD and J contributed by T^2, Q and φ of the model that matches: Distance=4.7867, number of samples=108.

Faults	Performances	T^2		Q		φ	
		95%	99%	95%	99%	95%	99%
Process fault	MDR	2.10	3.61	0.00	0.18	0.00	0.72
	FAR	33.81	0.00	41.43	23.57	54.05	10.24
	DTD	13.50	56.00	0.00	16.80	0.00	0.70
	J	**49.41**	**59.61**	**41.43**	**40.55**	**54.05**	**11.66**
Random and single	MDR	33.80	43.00	0.80	1.00	0.80	1.80
	FAR	0.50	0.20	27.00	12.10	17.80	1.80
	DTD	0.00	0.00	1.00	1.00	1.00	1.00
	J	**34.30**	**43.20**	**28.80**	**14.10**	**19.60**	**4.60**
Abrupt and single	MDR	0.00	1.20	0.00	0.00	0.00	0.00
	FAR	0.40	0.20	51.70	39.80	40.70	14.40
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.40**	**1.40**	**51.70**	**39.80**	**40.70**	**14.40**
Drift and single 1	MDR	5.60	9.20	1.20	3.20	1.80	4.40
	FAR	6.20	0.10	11.70	3.60	9.00	0.70
	DTD	23.00	42.00	4.40	4.90	0.75	8.80
	J	**34.80**	**51.30**	**17.30**	**11.70**	**11.55**	**13.90**
Drift and single 2	MDR	9.40	21.00	0.60	1.00	1.00	1.60
	FAR	0.10	0.10	18.80	7.80	9.70	0.70
	DTD	7.80	103.00	0.40	0.40	0.50	1.80
	J	**17.30**	**124.10**	**19.80**	**9.20**	**11.20**	**4.10**
Abrupt and multiple	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.20	0.20	45.30	27.00	27.80	4.10
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.20**	**0.20**	**45.30**	**27.00**	**27.80**	**4.10**
Drift and multiple	MDR	21.00	42.40	6.40	8.40	8.40	11.00
	FAR	0.60	0.10	33.60	19.80	20.50	2.70
	DTD	91.50	203.00	20.00	21.00	32.50	55.00
	J	**113.10**	**245.50**	**60.00**	**49.20**	**61.40**	**68.70**
Intermittent and single	MDR	0.00	0.00	0.00	0.00	0.00	0.00
	FAR	0.96	0.96	34.40	18.63	20.00	3.69
	DTD	0.00	0.00	0.00	0.00	0.00	0.00
	J	**0.96**	**0.96**	**34.40**	**18.63**	**20.00**	**3.69**
J average		**31.30**	**65.78**	**37.34**	**26.27**	**30.78**	**15.64**