ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Alkan ALKAYA

NOVEL DATA DRIVEN-BASED FAULT DETECTION FOR ELECTROMECHANICAL AND PROCESS CONTROL SYSTEMS

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

ADANA, 2012

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

NOVEL DATA DRIVEN-BASED FAULT DETECTION FOR ELECTROMECHANICAL AND PROCESS CONTROL SYSTEMS

Alkan ALKAYA

PhD THESIS

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

We certify that the thesis titled above was reviewed and approved for the award of degree of the Doctor of Philosophy by the board of jury on 16/01/2012.

…………………............	…………………………	………....................................
Assoc. Prof. Dr. İlyas EKER	Prof. Dr. Mehmet TÜMAY	Asst. Prof. Dr. Ramazan ÇOBAN
(Supervisor)	(Member)	(Member)

………………….........................	……….…………………….………………...
Assoc. Prof. Dr. Hüseyin CANBOLAT	Assoc. Prof. Dr. Zekeriya TÜFEKÇİ
(Member)	(Member)

This PhD Thesis is written at the Department of Electrical and Electronics Engineering of Institute of Natural and Applied Sciences of Çukurova University.
Registration Number:

Prof. Dr. İlhami YEĞİNGİL
Director
Institute of Natural and Applied Sciences

This thesis was supported by the Scientific Research Project Unit of Çukurova University for my thesis (Project Number: MMF2010D2).

ABSTRACT

PhD THESIS

NOVEL DATA DRIVEN-BASED FAULT DETECTION FOR ELECTROMECHANICAL AND PROCESS CONTROL SYSTEMS

Alkan ALKAYA

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

Supervisor : Assoc. Prof. Dr. İlyas EKER
Year: 2012, Pages: 165
Jury : Assoc. Prof. Dr. İlyas EKER
: Prof. Dr. Mehmet TÜMAY
: Assist. Prof. Dr. Ramazan ÇOBAN
: Assoc. Prof. Dr. Hüseyin CANBOLAT
: Assoc. Prof. Dr. Zekeriya TÜFEKÇİ

Fault Detection and Diagnosis (FDD) has become an attractive topic with increasing attention to improve efficiency, reliability and safety of modern engineering. The methodology used in FDD is clearly dependent on process and sort of available information, divided in two categories: model-based methods and data-driven-based methods. In the present research, observer-based and statistical information based Principal Component Analysis (PCA) method have been performed.

PCA is a statistical process monitoring technique that has been widely used in industrial applications. PCA methods for Fault Detection (FD) use data collected from a steady-state process to monitor T^2- and Q- statistics with a fixed threshold. For the systems where transient values of the processes must be taken into account, the usage of fixed threshold in PCA method causes false alarms and missing data that significantly compromise the reliability of the monitoring systems. In this thesis, two new methods based on PCA are proposed to overcome false alarms which occur in the transient states according to changing process conditions and the missing data problem. The proposed methods are implemented and validated experimentally on an electromechanical and process control system.

Keywords: Observer, PCA, Fault detection, Wavelet, Experimental application.

ÖZ

DOKTORA TEZİ

ELEKTROMEKANİK VE PROSES KONTROL SİSTEMLERİ İÇİN VERİ TABANLI YENİ BİR HATA ALGILAMA YÖNTEMİ

Alkan ALKAYA

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Doç. Dr. İlyas EKER
Year: 2012, Sayfa: 165
Jüri : Doç. Dr. İlyas EKER
: Prof. Dr. Mehmet TÜMAY
: Yrd. Doç. Dr. Ramazan ÇOBAN
: Doç. Dr. Hüseyin CANBOLAT
: Doç. Dr. Zekeriya TÜFEKÇİ

Hata Algılama ve Teşhisi (HAT), modern mühendislik alanındaki sistemlerin verimi, güvenliği ve güvenilirliğinin geliştirilmesine gösterilen ilgilin artması ile birlikte çok önemli ve dikkat çekici bir konu haline gelmiştir. Hata algılamada kullanılan yöntem, süreç modeline ve mevcut bilgilerin elde edilmesine bağlıdır ve bu iki kategoriye ayrılır: model tabanlı teknikler ve veri tabanlı teknikler. Bu çalışmada, gözleyici tabanlı ve istatistiksel bilgiye dayalı Temel Bileşenler Analizi (TBA) metodu gerçekleştirilmiştir.

TBA, endüstriyel uygulamalarda yaygın olarak kullanılan istatistiksel bir süreç takip tekniğidir. TBA metodu süreçlerin kararlı durumlarından toplanan verilerin T^2 ve Q görüntülerinde sabit bir eşik değeri ile takip ederek hataları algılar. Sistemlerin geçiş süreçleri de hesaba katıldığı anda bu klasik TBA yöntemi hata algılama sistemlerinin güvenliğini riske atacak yanlış alarmlar ve eksik hata verileri ortaya çıkarmaktadır. Bu tez çalışmasında, geçiş sürecinde meydana gelen yanlış alarmları ve eksik hata verilerini giderecek TBA yöntemine dayalı iki yeni metot önerilmiştir. Önerilen metotlar deneysel olarak elektromekanik ve süreç kontrol sistemleri üzerinde test edilmiştir.

Anahtar Kelimeler: Gözleyici, TBA, Hata algılama, Dalgacık, Deneysel uygulama.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude towards my supervisor, Assoc. Prof. Dr. İlyas EKER, for his unlimited support, guidance, patience and motivation for the successful completion of this research work.

I also thank my committee members, Prof. Dr. Mehmet TÜMAY and Prof. Dr. Nusret TAN, for their valuable help, support and suggestions.

Special thank is also for my colleagues and friends, Research Assistant Murat FURAT and Research Assistant Mustafa Berkan BİÇER, who always gave me constant support and help during my research studies.

I would like also to thank and acknowledge the financial supported by Scientific Research Project Unit of Çukurova University for my thesis (Project Number: MMF2010D2).

I also wish to express my deepest gratitude to my parents and extended family for their endless support, encouragement and patience.

Finally, I would like to special thanks to my wife Günseli who with love and patience encouraged me during crucial moments of my thesis.

CONTENTS **PAGE**

LIST OF TABLES **PAGE**

LIST OF FIGURES **PAGE**

LIST OF SYMBOLS

T_{α} : Fixed threshold

T_{adp} : Adaptive threshold

T_{comb} : Combination of fixed and adaptive threshold

T_{vsa} : Variance sensitive adaptive threshold

t : Time (independent variable)

$u(t)$: System input

$x(\mathrm{t})$: System state

$y(t)$: System output

A, *B*, *C*: System matrices

R : Real number vector

$\hat{x}$: Estimated system state

$\hat{y}$: Estimated system output

K : Observer gain

R^{+} : Positive real number

$f(t)$: Fault

$d(t)$: Disturbance

$r(t)$: Residual

$\tilde{x}$: Estimation error

$w_i(t)$: *i*th PC

X : Data matrix

n : Number of measurements

m : Number of variable

$\hat{T}$: Score matrix

$\hat{P}$: Loading matrix

Σ : Covariance matrix

Λ : Diagonal matrix

V : Eigenvectors

λ : Eigenvalues

a : Number of principal component

α : Confidence level

c_α : Normal distribution

μ : Mean

σ^2 : Variance

$\varnothing(z)$: Distribution function

$\Psi(t)$: Wavelet function

$\Psi^*(t)$: Conjugated wavelet

η : Scaling parameter

τ : Time localization parameter

j : Scale parameters

k : Translation parameters

$g(t)$: Energy signal

a_i : ith approximations

d_i : ith details

H_0 : High pass filter

L_0 : Low pass filter

Υ : Wavelet threshold

Γ : Wavelet coefficient

ζ : Wavelet parameters

Ω : Projection matrix

$\ddot{u}$: Number of inputs

$\breve{g}$: Number of states

ş : Number of outputs

LIST OF ABBREVIATIONS

2D	: Two-Dimensional
AHU	: Air Handling Unit
AIC	: Akaike Information Criteria
CPV	: Cumulative Percentage Variance
CUSUM	: Cumulative Sum
DAQ	: Data Acquisition Card
Db	: Daubechies
DC	: Direct Current
DPCA	: Dynamic PCA
DWT	: Discrete Wavelet Transform
EWMA	: Exponentially Weighted Moving Average
FD	: Fault Detection
FDD	: Fault Detection and Diagnosis
FDI	: Fault Detection and Isolation
FFT	: Fast Fourier Transform
FT	: Fourier Transform
ICA	: Independent Component Analysis
IFAC	: International Federation of Automatic Control
ISAS	: Integral Squared Alarm Signal
KPCA	: Kernel Principal Component Analysis
MA	: Moving Average
MBPCA	: Multi-block PCA
MPCA	: Moving PCA
MRA	: Multi Resolution Analysis
MSPC	: Multivariate Statistical Process Control
MSPCA	: Multiscale Principal Component Analysis
NLPCA	: Nonlinear Principal Component Analysis
ODEs	: Ordinary Differential Equations
PC	: Principle Components

PCA	: Principal Component Analysis
CEG	: Cause and Effect Graph
PI	: Proportional and Integral
PID	: Proportional and Integral and Derivative
PLS	: Partial Least Squares
PRESS	: Prediction Sum of Squares
QPT	: Qualitative Process Theory
QSIM	: Qualitative Simulation
QTA	: Qualitative Trend Analysis
RPCA	: Recursive PCA
SDG	: Signed Directed Graph or Signed Digraph
SPC	: Statistical Process Control
SPE	: Square Prediction Error
STFT	: Short-Time Fourier Transform
SVD	: Singular Value Decomposition
WT	: Wavelet Transform

1. INTRODUCTION

1.1. Background and Motivation

The growing demand for high performance, efficiency, safety and reliability and increasing complexity of the technical processes has been of great interest in the development of fault detection methods (Angeli, 2004). The advent of the computer in 1970s and its increasing application in decentralized process automation systems since 1975 was the beginning of computationally more involved and soft-based fault detection algorithms (Isermann, 2006). First publications about the fault detection methods appeared in connection with the aerospace systems (Beard, 1971) and chemical plants (Himmelblau, 1978). The early detection of faults may help to avoid system breakdowns and product deterioration. Fault Detection (FD) algorithms and their applications to a wide range of industrial processes have been the subject of intensive research over the past two decades (Isermann, 2005; Odgaard et al., 2008; Treetrong et al., 2009; Karami et al., 2010). The methods on Fault Detection and Diagnosis (FDD) can be divided into two main groups, the model-based and data driven-based methods (Venkatasubramanian et al., 2003a; 2003b; 2003c).

Model-based FD methods are based on comparing the behaviors of the actual plant and a mathematical model of the system (Hammouri et al., 2010). The method uses signal residuals, which indicate changes between the real process and the process model. Residual generation can be performed in different ways: parity equations (Zhong et al., 2009), observer-based generation (Peng et al., 2010), and the methods based on parameter estimation (Fischer et al., 2007). Neural networks and fuzzy systems have also been used in model-based FDI algorithms (Yüksel et al., 2010). On the other hand, reported applications or real-time implementation of the schemes are still very few. Observer-based and parameter estimation methods are the most frequently applied methods for the fault detection (Isermann, 1997). However, obtaining a complete and robust mathematical model is difficult due to process complexity and dimension. Therefore the method is generally suitable only for additive faults and limited to processes with a small number of variables (Yoon et al., 2000).

The data-based FD methods can be used to solve these problems (Venkatasubramanian et al., 2003a). The advantage of these methods is that the model of the system is not necessary to know in order to make a conclusion on a fault appearance. This means that the method is appropriate for the systems that cannot be easily or ever modeled, or for which the model is nonlinear, hybrid, or structurally ill-posed. Another advantage of the method is that, in addition to the additive faults, it is possible to detect multiplicative faults too. For the data-based methods, only the availability of large amount of historical process data is needed (Venkatasubramanian et al., 2003a). With the development of computer and data storage facilities, nearly every industrial process now routinely collects and stores massive amounts of data on many process variables. Efficient utilization of this large pool of data can lead to significant improvement in two areas. First, the data can be used to monitor the performance of the process over time for the fault detection. Second, frequently measured process variables can be used to infer the quality variables and an inferential control scheme can be developed. In situations, multivariate statistical methods such as Principal Component Analysis (PCA) and Partial Least Squares (PLS) can play a major role. These methods have been successfully applied to solve a wide range of multivariate problems in the chemical processes (Elshenawy et al., 2010), semiconductor processing (Zhiqiang et al., 2010), machining processes (Tsung, 2000), waste water treatment (Lennox and Rosen, 2002), nuclear power systems (Baraldi et al., 2010), air-conditioning processes (Wang et al., 2010), building central chilling systems (Youming and Lili, 2009; 2010). Most processes are well equipped with the sensors to realize automatic monitoring and control. However, most of the research studies are based on simulation results (Tsung, 2000; Baraldi et al., 2010).

Although the conventional PCA method has been highly successfully used for monitoring purposes, its best applications are restricted to analyze steady-state data containing linear relationships between the variables (Zvokelj et al., 2010). Because these processes are fairly rare, various modifications of the PCA method have been developed, including the dynamic PCA (Pöllänen et al., 2006), non-linear PCA (Jia et al., 2010), multi-block PCA (Westerhuis et al., 1998), multi-way PCA (Wold et al., 1987), moving PCA (Kano et al., 2001) and recursive PCA (Li et al., 2000).

Existing methods for the fault detection have largely focused on the steady-state operations and are not directly applicable during the transitions (Pöllänen et al., 2006; Jia et al., 2010; Westerhuis et al., 1998; Kano et al., 2001; Li et al., 2000). Applying a PCA method to such a transient process (like servo systems) can produce excessive number of false alarms or missed detection of process faults which significantly compromises the reliability of the monitoring system. Therefore, a novel PCA fault detection method is required that explicitly caters for the non-steady states and wide operating condition changes during transitions.

The data collected from industrial process often contain measurement noise that causes missing fault (interrupted) signal components even if the PCA method is used**.** Therefore, the noise has to be removed before PCA analysis for robust fault detection.

Extraction of the weak signals and de-noising are very important for fault diagnostics, especially for early fault detection, in which cases features are very weak and masked by the noise (Peng et al., 2004). The noises are often stochastic signals with broadband, whose frequency band will overlap with the interested signals. Therefore it is difficult to eliminate the noise from the signals effectively with general filter-based methods such as exponential, polynomial, median and Kalman filters (Shao et al., 1999). In addition, to implement some filtering algorithms it is necessary to have future values, e.g. in the median filter. In this respect they are unsuitable for on-line application. The Wavelet Transform (WT) addresses some of these limitations (Shao et al., 1999)**.**

The wavelet transform has attracted recent interest in applied mathematics for signal processing and fault detection (Hui et al., 2011; Wu et al., 2011; Rafiee et al., 2011)**.** This new mathematical technique has been demonstrated to be fast in computation with localization and quick decay properties in contrast to existing popular methods, especially, the Fast Fourier Transform (FFT). One of the main features of WT is that it may decompose a signal directly according to the frequency and represent it in the frequency domain distribution state in the time domain. In the transformation, both time and frequency information of the signal are retained.

1.2. Research Objectives

The present research focuses on model and data driven-based fault detection and diagnosis methods and experimental application to electromechanical and process control systems.

The objectives are to

- Describe the terminology used in the field of fault detection and diagnosis, fault types and classification.
- Provide an overview of various detection methods from different perspectives.
- Develop and validate model-based (observer) fault detection scheme for abrupt, incipient and intermittent faults.
- Develop effective data-driven methodologies for fault detection and diagnosis.
- Develop conventional PCA method for process fault detection (Conventional PCA based Q and T^2 – statistics for steady-state conditions).
- Propose a new threshold based PCA method that is sensitive to transient states and changes in operation (servo tracking system).
- Improve the conventional monitoring charts for PCA fault detection purpose. These monitoring charts are:

  a) Combined fixed and adaptive threshold, T_{comb}.

  b) Variance sensitive adaptive threshold, T_{vsa}.
- Propose a wavelet based PCA fault detection method.

- Evaluate performance of the proposed method with experimental applications (Electromechanical and process control systems).
- Results of the proposed methods and conventional methods are compared in terms of sensitivity to
  a) Detect fault accurately.
  b) Eliminate false alarm occurring in transient states.
  c) Eliminate missing fault problem.

1.3. Contributions of Research

The main differences and important contributions of this research can be summarized as follows:

- Application of the observer-based fault detection method to an electromechanical system.
- A combined threshold PCA method is proposed and implemented to prevent false alarm when transient states taken into account.
- A variance sensitive adaptive threshold PCA method is proposed and implemented to overcome missing fault alarm.
- Wavelet-based combined PCA method is implemented to prevent the false alarms and to produce uninterrupted fault alarm signal.
- Conventional and proposed PCA methods are implemented experimentally on a DC motor system and process control systems.

The framework of the present research is presented in Figure 1.1. There are two main branches, one is the general theory development, and the other is the application to dynamical systems. The circles in the chart denote the existing theories that have been considered in the research and the patterned square are the original contributions of the current study.

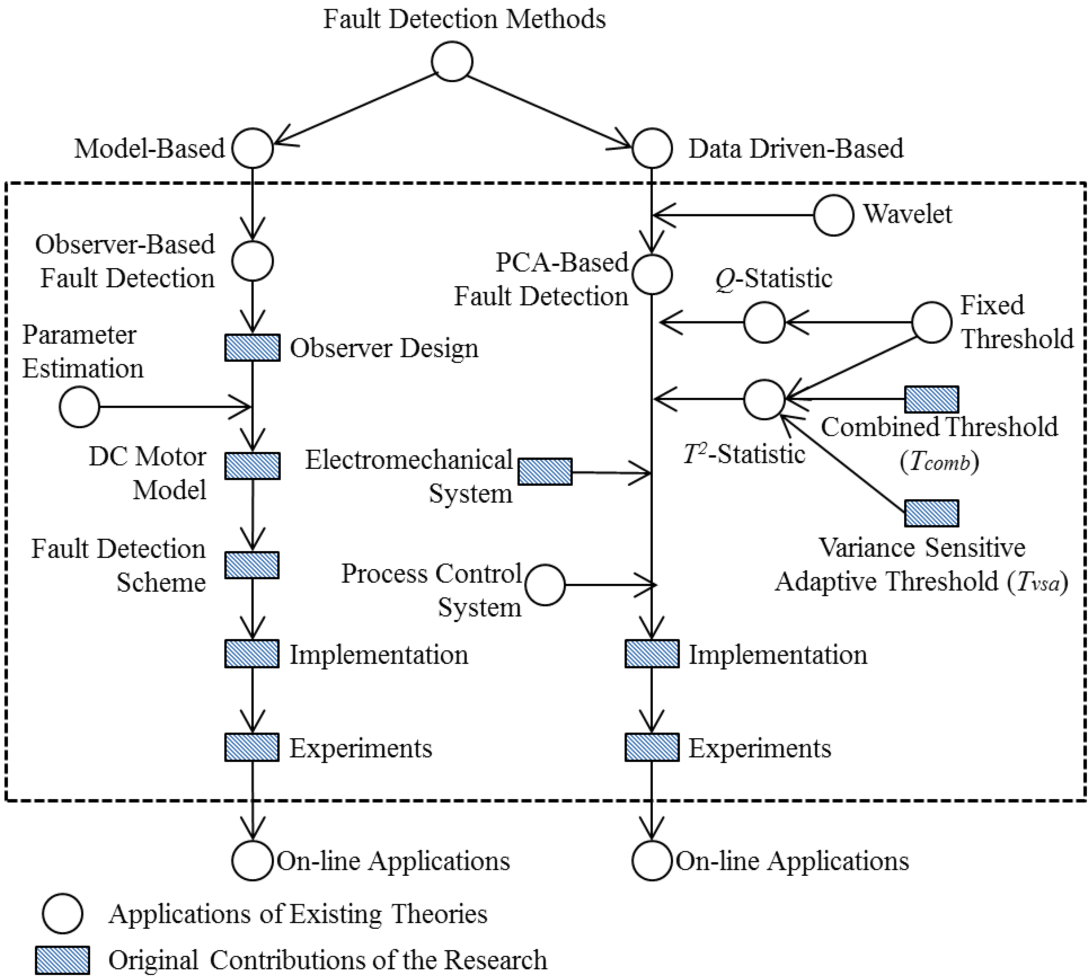


Figure 1.1. Framework of the Present Research

1.4. Thesis Outline

The thesis is organized as follows:

Chapter 1: Introduction

The first chapter comprises of the introduction of the research, research background, motivation, objectives and contributions of the current research.

Chapter 2: Fault Detection and Diagnosis (FDD)

The second chapter of the thesis first starts to give requirements of FDD methods in engineering systems. A common terminology and fault types in the fault detection and diagnosis framework in order to comment on some developments in the field of FDD are presented. Literature survey for the various fault detection methods from different perspectives is also provided. Strengths and weaknesses of each fault detection methods are outlined. Basic principles of the model-based fault detection methods are discussed, followed by a detailed description of observer-based fault detection method studied in the thesis. In addition, the robustness is introduced.

Chapter 3: Data-Driven-Based FDD: Concept and Theory

The third chapter discusses a multivariate statistical data-driven fault detection method, PCA. Different PCA methods such as conventional PCA, Dynamic PCA, Nonlinear PCA, Multi block PCA, Recursive PCA, Moving PCA, and Multi-scale PCA, are presented. The methodology and applications area are given for each method. This chapter also presents calculation of PCA model and monitoring statistics (T^2 and Q) for fault detection. Combined Threshold (T_{comb}) and Variance Sensitive Adaptive Threshold (T_{vsa}) are described and presented step by step. Wavelet theory including de-nosing, Multi Resolution Analysis (MRA) and applications in fault detection is emphasized.

Chapter 4: Experimental Setup

Experimental setup and preliminaries are introduced in chapter 4. Mathematical model of the electromechanical plant and properties of the process control systems are described. Technical information of Data Acquisition Card (DAQ) and computer used in the experiments are presented.

Chapter 5: Experiments and Results

Observer-based fault detection, variance sensitive adaptive threshold and proposed wavelet-based combined PCA methods are implemented. Several experimental tests are performed to demonstrate the performance of proposed methods. The capabilities of the proposed methods are deeply analyzed during steady-state and transient-state operation. The results are presented and compared with the results obtained by means of conventional approach.

Chapter 6: Conclusions

The last chapter gives concluding remarks, summarizes the main contributions of the thesis and discusses directions for future work.

Finally, all the references used in the thesis, biographical information of the author and sections of the Appendix are presented.

1.5. Conclusions

The first chapter of the thesis is to provide basic features of the research consisting the research motivation, problem statement, research objectives, contributions and scopes. The contents of the chapters composing the thesis and the main contributions were presented.

2. FAULT DETECTION AND DIAGNOSIS (FDD)

2.1. Importance of FDD

Fault detection and diagnosis has been becoming more and more important for process monitoring because of the increasing demand for higher performance as well as for increased safety and reliability of dynamic systems (Isermann, 2011). Fault detection and diagnosis deals with the timely detection, diagnosis and correction of abnormal conditions of faults in a process. The early detection of the occurrence of faults is critical in avoiding product deterioration, performance degradation, major damage to the machinery itself and damage to human health or even loss of lives. The quick and correct diagnosis of the faulty component then facilitates the making of appropriate and optimal decisions on emergency and corrective actions, and on repairs. These aspects can minimize downtime, increase the safety of plant operations and reduce manufacturing costs (Isermann, 2011). Hence, fault diagnosis is a major research topic attracting considerable interest from industrial practitioners as well as academic researchers.

A typical operation and maintenance process using automated FDD can be viewed as having four distinct functional processes (Katipamula and Brambley, 2005), as shown in Figure 2.1 and similar process descriptions have been provided in (Isermann, 1984; Rossi and Braun, 1997). The first step is to monitor the physical system or device and detect any abnormal conditions (problems). This step is generally referred to as fault detection. When an abnormal condition is detected, fault diagnosis is used to evaluate the fault and determine its causes. These two steps constitute the FDD process. Following diagnosis, fault evaluation assesses the size and significance of the impact on system performance (in terms of energy use, cost, availability, or effects on other performance indicators). Based on the fault evaluation, a decision is then made on how to respond to the fault (e.g., by taking a corrective action or possibly even no action).

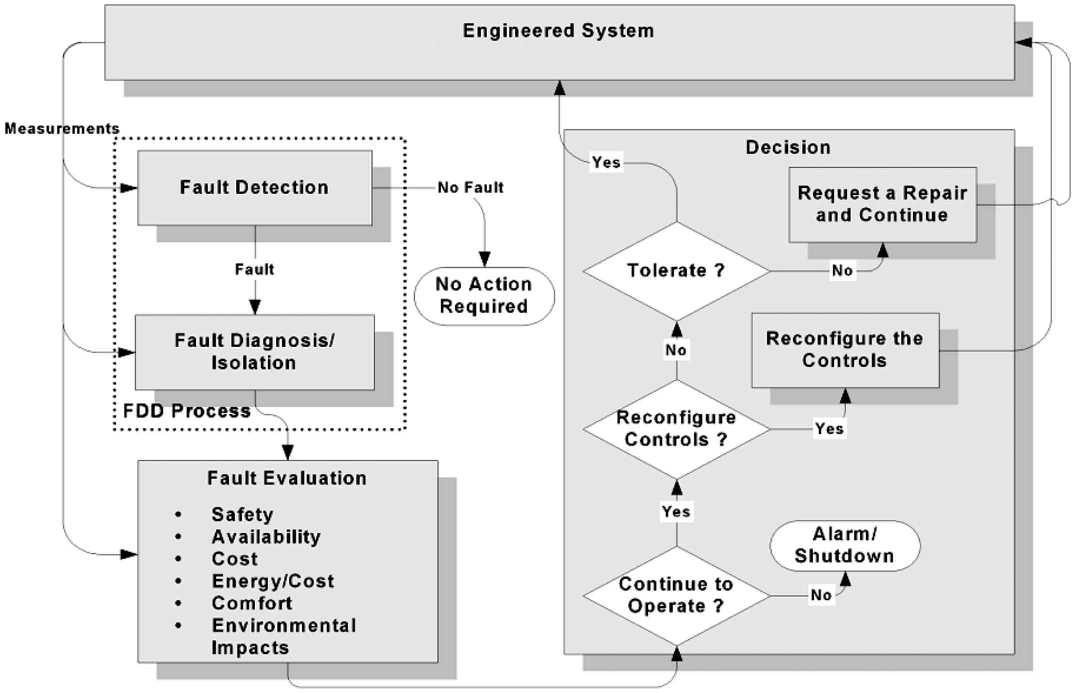


Figure 2.1. Generic Application of Fault Detection and Diagnostics to Operation and Maintenance of Engineered Systems.

FDD itself is frequently described as consisting of three key processes: fault detection, fault isolation, and fault identification. The first, fault detection, is the process of determining that some fault has occurred in the system. The second involves isolating the specific fault that occurred, including determining the kind of fault, the location of the fault, and the time of detection. The third process, fault identification, includes determining the size and time-variant behavior of a fault. Together, fault isolation and fault identification are commonly termed fault diagnosis.

There is a wide literature documentation on process fault diagnosis ranging from analytical methods to artificial intelligence and statistical approaches (Katipamula and Brambley, 2005). From a modeling perspective, there are methods that require accurate process models, semi-quantitative models or qualitative models. However, some methods that do not assume any form of model information rely only on historic process data. Therefore, this chapter is mainly devoted to introduce the terminology used in the field of fault detection and diagnosis, fault types and classification and to provide an overview of various diagnostic methods from different perspectives.

2.2. Fault Classification

The types of faults depend basically on their location within the system, the number of components that can be affected and their temporal evolution (Isermann, 2011). Taking into account the effects of the faults, these are classified as *additive faults* (those which correspond to sensor and actuator faults) and *multiplicative faults* (or parametric).

Additive process faults: These are unknown inputs acting on the plant, which are normally zero and, when present, can cause a change in the plant outputs independent of the known inputs.

Multiplicative process faults: These are changes (abrupt or gradual) in some plant parameters. They may cause changes in the plant outputs which also depend on the magnitude of the known inputs. Such faults describe the deterioration of the plant equipment, such as contamination, clogging, or the partial or total loss of the power.

In literature faults can take place in different parts of a system, and are classified as actuator faults, sensor faults and component faults (Kanev, 2004)**.**

Actuator faults represent partial or complete loss of control action. Total actuator fault can occur, for instance, as a result of a breakage, cut or burned wiring, shortcuts, or the presence of outer body in the actuator. Despite of the input applied to an actuator, it produces no actuation. This is an example of a completely lost actuator (stuck actuator). Partially failed actuator produces only a part of the normal (i.e., under nominal operating condition) actuation. It can result from, e.g., hydraulic or pneumatic leakage, increased resistance or fall in the supply voltage.

Sensor faults represent incorrect reading from the sensors. They also are subdivided into partial and total. Produced information is not related to value of the measured physical parameter in case of the total actuator fault. They can be due to broken wires, lost contact with the surface, etc. The output containing useful information could still be retrieved. This can, for instance, be a gain reduction, a biased measurement or increased noise.

Component faults are faults in the components of a complex system, i.e., all faults that cannot be categorized as sensor or actuator faults. These faults represent changing in the damping constant, etc., that are often due to structural damages. They often result in a change in the dynamical behavior of a nonlinear complex system. They are the most frequently encountered types in fault family to deal with.

Regarding the time dependency of faults, they can be distinguished as illustrated in Figure 2.2.

Abrupt faults occur instantaneously often as a result of hardware damage. Usually they are very severe as they affect the performance or the stability of the system.

Incipient faults represent slow in time parametric changes, often as a result of aging. They are more difficult to detect.

Intermittent faults are faults that appear and disappear repeatedly; for instance, due to a partially damaged wiring.

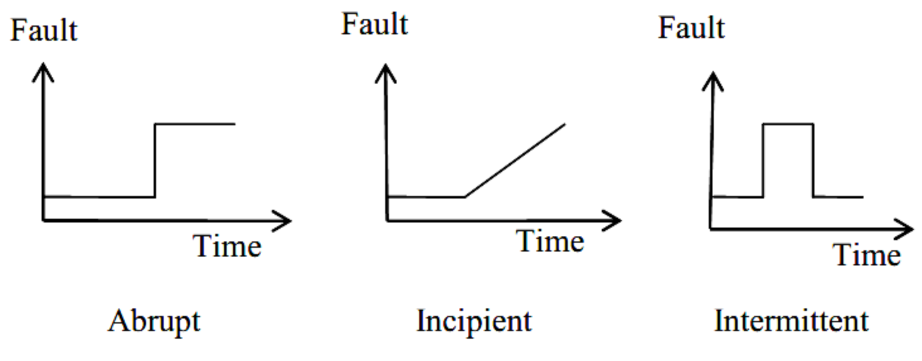


Figure 2.2. Time Characteristics of Faults.

2.3. Historical Background and Literature Survey

FDD methods can be classified into two broad categories, model-based methods and data-driven-based methods (Venkatasubramanian et al., 2003a; 2003b; 2003c). The two categories differ by the knowledge used to diagnose the cause of faults, although both may use simulation models and measurement data. Model-based methods use “priori knowledge” (knowledge available in advance) to identify the differences between model simulation results and actual operation measurements. Simulation models are commonly based on first principles and do provide process insight. However, they may not fit the process data that well and are not able to explain systematic variation. Data-based methods may not use any physical knowledge; instead, they can be driven completely by recorded measurement data. These data driven models fit the data properly, but cannot be generalized to different situations and do not always generate good process insight.

Model-based methods are further divided into quantitative and qualitative modeling methods. Quantitative models are based on mathematical relationship derived from the underlying physical knowledge. Quantitative methods rely on explicit mathematical models of a system to detect and diagnose faults. By understanding the physical relationships and characteristics of a system, mathematical equations to represent each component of the system can be developed and solve to simulate the steady and transient behavior of the systems. Another broad method is qualitative modeling, which uses rule-based methods developed based on priori knowledge. Qualitative models use the qualitative rule relationships to detect and diagnose faults instead of quantitative mathematical equations. The rules are derived from expert knowledge, process history data and quantitative models simulation data. Expert knowledge is normally summarized to a database in the form of if-then statements.

Data-based models are derived from process history data, and are subdivided into black box model and gray box model. Their difference is whether model parameters have physical meaning. Black box models use non-physical based relationship to represent the characteristics of a system. Model parameters do not represent actual physical properties. Black box models use techniques such as linear or multiple linear regression, artificial neural networks, and fuzzy logic. In a gray box model, the model parameters are determined based on physical principles. Parameter estimation techniques are often used to obtain those parameters from measurement data. Comparing with the black box modeling, the gray box modeling needs higher-level user expertise to form the model parameters and estimate parameter values.

An overview on general FDD concepts and a chart for classification of diagnostic algorithms is given (Venkatasubramanian et al., 2003a) as shown in Figure 2.3. FDD methods are broadly classified into two categories: model-based methods and data-driven-based methods.

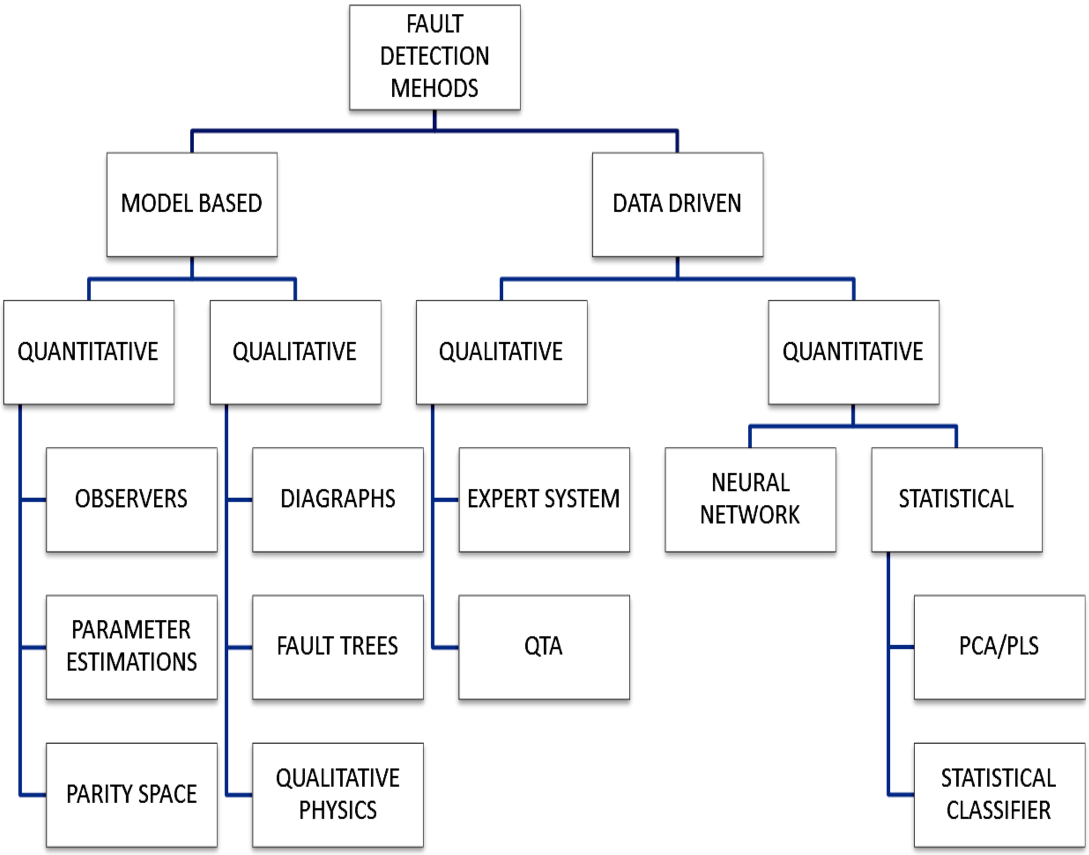


Figure 2.3. Classification of Fault Detection Methods.

2.3.1. Model-based Fault Detection Methods

2.3.1.1. Quantitative Model-Based Methods

Most of the work on quantitative model-based approaches has been based on using general input-output and state space models to generate residuals. Different approaches for fault detection using mathematical models have been developed in the last few decades (Isermann, 2005). These approaches can be classified into observer, parity space and parameter estimation methods (Isermann, 2011).

Observer or filter-based: The basic idea of the observer or filter-based approaches is to estimate the states or outputs of the system from the measurements by using either Luenberger's observers in a deterministic setting (Chen et al., 2010) or Kalman's filters in a stochastic case (Villez et al., 2011). The flexibility in selecting observer gains has been studied (Frank and Ding, 1997). Integrated design of observer based Fault Detection (FD) for a class of uncertain nonlinear systems with Lipschitz non-linearities is studied (Chen et al., 2011). The freedom in the design of the observer can be used to enhance the residuals for isolation. The dynamics of the response can be controlled, within certain limits, by placing the poles of the observer.

In recent years, several model-based methods have been developed (Isermann, 2006) and especially observer-based methods have been given more attention (Edwards, et al., 2007; Shields, 2005). For example sliding mode approaches (Edwards et al., 2000), geometric approach (Persis and Isidori, 2001) and adaptive control (Zhang et al., 2002) are combined successfully with observer-based fault detection and diagnosis techniques.

Parity space: Parity equations are rearranged and usually transformed variants of the input-output or state-space models of the plant (Zhong et al., 2009). The essence is to check the parity (consistency) of the plant models with sensor outputs (measurements) and known process inputs. The idea of this approach is to rearrange the model structure to get the best fault isolation. Concept of the parity relations was introduced by (Chow and Willsky, 1984). Further developments have been made by

(Gertler et al., 1995; Zhang et al., 2006) among others. There is a fundamental equivalence between parity relations and the observer-based methods. Both methods produce identical residuals if the generators are designed for the same specification (Ding and Jeinsch, 1999).

Parameter Estimation: The model-based FDI can also be achieved by means of using the system identification techniques if the basic structure of the model is known (Isermann, 2011). This approach is based on the assumption that faults are reflected in the physical system parameters such as friction, mass, resistance, etc. The basic idea is that the parameters of the actual process are estimated on-line using well known parameter estimation methods and the results are compared with the parameters obtained initially under the fault-free case. Any discrepancy indicates a fault. The parameter estimation may be more reliable than the analytical redundancy methods, but it is also more demanding in terms of on-line computation and input excitation requirements. A relationship has been found between parity relations and parameter estimation as well (Gertler, 2000).

It can be seen that one of the major advantages of using the quantitative model-based approach is that we will have some control over the behavior of the residuals. However, several factors such as system complexity, high dimensionality, process nonlinearity and/or lack of good data often render it very difficult even impractical, to develop an accurate mathematical model for the system. This, of course, limits the usefulness of this approach in real industrial processes (Venkatasubramanian et al., 2003a).

Advantages and Disadvantages of Quantitative Models

Advantages of FDD based on quantitative models include (Katipamula and Brambley, 2005):

- 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.

Disadvantages of FDD based on quantitative models include (Katipamula and Brambley, 2005):

- 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.
- Extensive user input creates opportunities for poor judgment or input errors that can have significant impacts on results.

2.3.1.2. Qualitative Model-based Methods

Based on various forms of qualitative knowledge used in fault diagnosis, qualitative model-based approaches can be classified into digraphs, fault trees and qualitative physics methods (Venkatasubramanian, 2003b).

Causal model approaches using digraphs: A Signed Directed Graph or Signed Digraph (SDG), as a qualitative model, effectively and graphically represents a process system. Cause-effect relations or models can be represented in the form of signed digraphs. A digraph is a graph with directed arcs between the nodes and SDG is a graph in which the directed arcs have a positive or negative sign attached to them. The directed arcs lead from the ‘cause’ nodes to the ‘effect’ nodes. SDGs provide a very efficient way of representing qualitative models graphically and have been the most widely used form of causal knowledge for process fault diagnosis (Venkatasubramanian et al., 2003b). Using signed digraphs (SDG) for fault diagnosis was first proposed and what is called a cause-effect graph (CE graph) was derived from SDG in 1979 (Iri et al., 1979). SDG can be obtained from differential algebraic equations for the process (Umeda et al., 1980). The issue of conditional arcs in SDG is addressed and also extended the idea of SDG to include five-range patterns instead of the usual three-range pattern used in the standard SDG (Shiozaki et al., 1985). Partial system dynamics, statistical information about equipment failure, and digraphs to represent the failure propagation network for identifying fault location are used (Kokawa et al., 1983). Rule-based methods using SDG have been used for fault diagnosis (Kramer and Palowitch, 1987). An important work in the field of steady-state Qualitative Simulation (QSIM) using SDG has been presented in (Oyeleye and Kramer, 1988).

In recent years, the problem of fault diagnosis using what is called Possible Cause and Effect Graph (PCEG) models have been approached (Wilcox et al., 1994a; Wilcox et al., 1994b). Digraph-based models for automated HAZOP analysis have been used (Vaidhyanathan et al., 1995). Use of SDGs for multiple fault detection is demonstrated (Vedam et al., 1997a). Improvement of fault resolution in SDG models through the use of fuzzy set theory are discussed (Han et al., 1994). A framework for process supervision using fuzzy logic-based fault diagnosis has been presented (Genovesi et al., 1999). How fuzzy digraphs can be used for qualitative and quantitative simulation of the temporal behavior of process systems has been presented (Li and Wang, 2001).

Fault trees approaches: Fault trees are used in analyzing system reliability and safety. Fault tree analysis was originally developed at Bell Telephone Laboratories in 1961 (Yang, 2004). Fault tree is a logic tree that propagates primary events or faults to the top level event or a hazard. The tree usually has layers of nodes. At each node different logic operations like AND and OR are performed for propagation. Fault-trees have been used in a variety of risk assessment and reliability analysis studies (Kelly and Lees, 1986; Ulerich and Powers, 1988).

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). Quantitative models have an advantage in modeling the transient behavior of the systems more precisely than any other modeling technique.

The first approach is to derive qualitative equations from the differential equations termed as confluence equations (Yang, 2004). Considerable work has been done in this area of qualitative modeling of systems and representation of causal knowledge (Iwasaki, 1986). 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. Sacks examines piece-wise linear approximations of nonlinear differential equations through the use of a qualitative mathematical reasoned to deduce the qualitative properties of the system (Sacks, 1988). Kuipers predicts qualitative behavior by using Qualitative Differential Equations (QDEs) that are an abstraction of the ODEs that represent the state of the system (Kuipers, 1986). Bendapudi and Braun provide a detailed list of available dynamic models for vapor compression equipment. They also developed a dynamic centrifugal chiller model from first principles for FDD (Bendapudi et al., 2002). In terms of applications of qualitative models in fault diagnosis, Qualitative Simulation (QSIM) and Qualitative Process Theory (QPT) have been the popular approaches. Examples of research work in QSIM include in (Kay and Kuipers, 1993). Examples of using the QPT framework in process fault diagnosis include in (Falkenhainer and Forbus, 1991).

***Advantages and Disadvantages* of Qualitative Models**

Advantages of qualitative models are (Katipamula and Brambley, 2005):

- 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.

Disadvantages of FDD based on qualitative models include (Katipamula and Brambley, 2005):

- 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.

2.3.2. Data-Driven Based Fault Detection Methods

2.3.2.1. Quantitative Data-Driven Methods

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 (Venkatasubramanian et al., 2003c).

Multivariate statistical approaches: In multivariate situations, the probability that a process is completely under normal operating control region is less than that in the univariate case (Montgomery, 1996). Similarly the probability that a multivariate process is completely out-of-control is less than that of a univariate case. Using multivariate control charts the desired confidence level can be maintained by taking advantage of the cross correlation information between variables. Hence, the process can be analyzed for its stability without the added complication of maintaining many control charts at the same time.

Classical multivariate statistical process control methods, for example latent variable 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 (Kresta et al., 1991; MacGregor and Kourti, 1995). The use of PCA assumes data are approximately normally distributed and time independent (Jollife, 1986). Generally, industrial processes are dynamic in nature, and exhibit highly auto-correlated process variables. Moreover, correlations between variables tend to be nonlinear. These characteristics can lead to an excess of false alarms or a significant loss of information when using linear PCA for process monitoring.

To address these limitations, several modifications to basic PCA have been proposed. Nonlinear Principal Component Analysis (NLPCA) is used to capture nonlinear relationships among variables. Compared to linear PCA, NLPCA can explain more variance in smaller dimensions (Dong and McAvoy, 1996; Kramer, 1991; Tan and Mavrovouniotis, 1995). Similarly, dynamic PCA has been proposed to eliminate the effect of autocorrelation in process data by augmenting the data matrix with time-lagged variables (Ku et al., 1995; Luo et al., 1999; Lin et al., 2000). Adaptive PCA updates the model parameters continuously by exponential smoothing so as to get the model adjusted to suit new operating conditions (Wold, 1994). Multiway and multiblock PCA are suitable for batch process operations (Nomikosi and MacGregor, 1994; MacGregor et al., 1994; Wold et al., 1996). Moreover, multiblock PCA allows for efficient computation of very large datasets.

Conventional multivariate process monitoring methods detect fault conditions at a single scale since they represent the data in terms of basis functions at a fixed resolution or scale in time and frequency. An early development of a multiscale framework for statistical process monitoring can be attributed to Bakshi (1998) who proposed use of wavelets to decompose data into several views or scales prior to the application of PCA. This has a two-fold effect, namely decorrelation across variables and elimination or reduction of autocorrelation individual variables. Wavelets are appropriate in this regard due to their time-frequency localization property. Several combinations of PCA with wavelets have been developed to monitor the process because of the ability of wavelets to compress multiscale features of the signal and approximately remove serial or auto correlations in time signals (Bakshi, 1998; Misra et al., 2002; Maulud et al., 2006; Rosen and Lennox, 2001). Multiscale Principal Component Analysis (MSPCA) approach adapts to the nature of the signal features and this approach has been extended to a nonlinear MSPCA by using neural networks to extract the latent nonlinear structure from the PCA transformed data (Fourie and Devaal, 2000; Shao et al., 1999; Zhinqiang and Qunxiong, 2005).

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 (Ocak, 2003; Rengaswamy et al., 2000).

Neural network approaches: Considerable interest has been shown in the literature to the application of neural networks for fault diagnosis. 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, radial basis and so on; and (ii) the learning strategy such as supervised and unsupervised learning (Yang, 2004).

The most popular supervised learning strategy in neural networks has been the back-propagation algorithm. There are a number of papers that address the problem of fault diagnosis using back-propagation neural networks. In chemical engineering, Watanabe et al. (1989), Ungar et al. (1990) and Hoskins et al. (1991) were among the first researchers to demonstrate the usefulness of neural networks for fault diagnosis.

Most of the work on improvement of performance of standard back-propagation neural networks for fault diagnosis is based on the idea of explicit feature presentation to the neural networks by Fan et al. (1993), Tsai and Chang (1995), and Maki and Loparo (1997). Modifications to the selection of basis functions have also been suggested to the standard back-propagation network. Different network architectures have been used for the problem of fault diagnosis. For example, Bakshi and Stephanopoulos (1993) proposed Wave-net: a multiresolution hierarchical neural network. Self-organizing neural network structures such as the ART2 network (Carpenter and Grossberg, 1988) have also been extensively used in fault diagnosis. Whiteley and Davis (1994) demonstrate the use of the ART2 network for the interpretation of sensor data. Chen et al. (1999) and Wang et al. (1999) discuss the integration of wavelets with ART networks for the development of diagnostic systems.

2.3.2.2. Qualitative Data-Driven Methods

Two of the major methods that extract qualitative history information are expert systems and Qualitative Trend Analysis (QTA) (Venkatasubramanian, 2003c).

Expert system approaches: Rule-based feature extraction has been widely used in expert systems for many applications. 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 (Venkatasubramanian et al., 2003c).

Initial attempts at the application of expert systems for fault diagnosis can be found in Henley (1984), Chester et al., (1984) and Niida (1985). Rich et al. (1989) discuss a diagnostic expert system for a whipped topping process.

There are a number of other researchers who have proposed on application of expert systems for diagnostic problems. Basila et al. (1990) have developed a supervisory expert system that uses object-based knowledge representation to represent heuristic and model-based knowledge. Becraft and Lee (1993) have proposed an integrated framework comprising of a neural network and an expert system. Tarifa and Scenna (1997) have proposed a hybrid system that uses signed directed graphs (SDG) and fuzzy logic. Wo et al. (2000) have presented an expert fault diagnostic system that uses rules with certainty factors. Leung and Romagnoli (2000) have presented a probabilistic model-based expert system for fault diagnosis.

Qualitative trend analysis approaches: 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 (Venkatasubramanian et al., 2003c). Cheung and Stephanopoulos (1990) have built a formal framework for the representation of process trends. Janusz and Venkatasubramanian (1991) identify a comprehensive set of primitives by which any trend can be represented. Rengaswamy and Venkatasubramanian (1995) have shown how primitives can be extracted from raw noisy sensor data by treating the problem of primitive identification as a classification problem using neural networks. Vedam and Venkatasubramanian (1997b) proposed a wavelet theory based adaptive trend analysis framework and later proposed a dyadic B-Splines based trend analysis algorithm (Venkatasubramanian et al., 2003a). Recently, Rengaswamy et al. (2001) have discussed the utility of trend modeling in control loop performance assessment.

Advantages and Disadvantages of Data Driven-Based Models

The advantages of FDD methods based on process history are (Katipamula and Brambley, 2005):

- These methods are well suited to problems for which theoretical models of behavior are poorly developed or inadequate to explain observed performance.
- They are suited where training data are plentiful or inexpensive to create or collect.
- Black-box models are easy to develop and do not require an understanding of the physics of the system being modeled.
- Computational requirements vary, but they are generally manageable.
- There is a wealth of documented information available on the underlying mathematical methods.

Disadvantages of process history-based methods of FDD include (Katipamula and Brambley, 2005):

- Gray-box models based on first principles require a thorough understanding of the system and expertise in statistics.
- Most models cannot be used to extrapolate beyond the range of the training data.
- 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.

2.4. Model-Based Fault Detection Method

Model-based fault detection methods in dynamic systems have been received much attention over the last decades, both in research context and in the domain of application studies on real plants (Renganathan and Bhaskar, 2010). Model-based FD methods are based on comparing the behaviors of the actual plant and a mathematical model of the system (Hammouri et al., 2010). The method uses signal residuals, which indicates changes between the real process and the process model. A general principle of the model-based fault detection is shown in Figure 2.4. In the figure, $u(t)$ is system input variables, $y(t)$ is system output, $r(t)$ is the residual signal, which is used to compare the difference between real process and system model. The model-based approach (analytical redundancy) is a widely accepted modern approach for fault detection and isolation (FDI). It is based on the idea that the measurements from dissimilar sensors are functionally related. Any violation of these relationships indicates the occurrence of faults. It also indicates the essential problem in model-based fault detection and isolation (FDI) is to generate a good residual model describing the behavior of the monitored system.

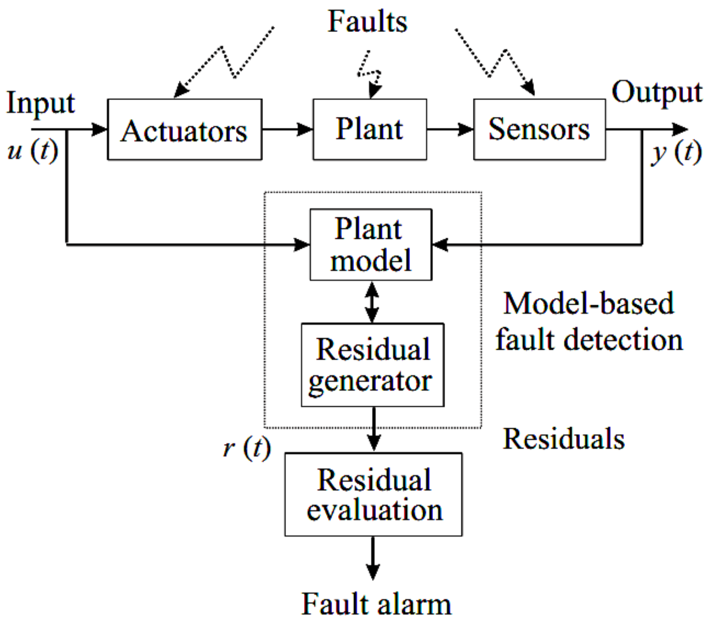


Figure 2.4. Scheme for the Model-Based Fault Detection.

Residual generation can be performed in different ways: parity equations (Zhong et al., 2009), observer-based generation (Peng et al., 2010), and the methods based on parameter estimation (Fischer et al., 2007). Key references of model-based FDI can be found in Chen and Patton (1999); Gertler (1998); Patton and Chen (1997); Isermann and Ballé (1997); Frank (1996); Willsky (1976); Garcia and Frank (1997); Blanke et al. (2003).

Observer-based fault detection method is one of the most effective methods and has obtained much more attention (Depersis and Isidori, 2000; Odgaard et al., 2008; Chen and Chowdhury, 2010; Fanglai and Feng, 2010). Over the years, different linear time-domain designed observers have been investigated. A state space observer which can be used for fault detection is the Kalman filter (Karami et al., 2010). A drawback of the Kalman filter is that the feedback gain matrix is determined in a way that considers the sensitivity with respect to disturbances and not faults. Therefore, Luenberger observer (Chen and Chowdhury, 2010) is commonly used in the residual generation part. The process model in an observer can be extended with integral state variables that represent functions of the faults which will then be estimated.

2.4.1. Observer-Based Fault Detection Method

The basic idea of an observer based FDI scheme is to reconstruct the outputs of the system from the measurements or subsets of the measurements with the help of an observer and using the estimation error as a residual for the detection and isolation of the faults. For a given linear state-space system:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \tag{2.1}$$

where the input $u(t) \in \mathfrak{R}^{ü}$, the state $x(\mathrm{t}) \in \mathfrak{R}^{ğ}$, the output $y(t) \in \mathfrak{R}^{ş}$. $\mathfrak{R}$ denotes real number vector. Assuming that $A \in \mathfrak{R}^{ğxğ}$, $B \in \mathfrak{R}^{ğxü}$, $C \in \mathfrak{R}^{şxğ}$ are known plant matrices. *ü*, *ğ*, *ş* denotes number of inputs, number of states and number of outputs, respectively.

One can design a state observer in the following form provided the system is completely observable:

$$\begin{cases} \dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(y(t) - \hat{y}(t)) \\ \hat{y}(t) = C\hat{x}(t) \end{cases} \tag{2.2}$$

where $\hat{x}$, $\hat{y}$ and K are the estimated system state, the estimated system output and the observer gain respectively, $K \in \mathfrak{R}^{+}$, $\mathfrak{R}^{+}$ denotes set of positive real numbers. The difference between the measured system output $y(t)$ and the estimated system output $\hat{y}(t)$ can be used as the residual signal for the purpose of fault detection and isolation. Figure 2.5 shows the configuration of an observer-based residual generator, where $f(t)$, $d(t)$ and $r(t)$ represent the fault, the disturbance and the residual signal respectively. Subtracting (2.2) from (2.1) gives:

$$\begin{cases} \dot{x}(t) - \dot{\hat{x}}(t) = Ax(t) - A\hat{x}(t) - K\left(Cx(t) - C\hat{x}(t)\right) \\ \qquad\qquad\quad = (A - KC)\left(x(t) - \hat{x}(t)\right) \end{cases} \tag{2.3}$$

Define the state estimation error $\tilde{x}$ as:

$$\tilde{x}(t) = x(t) - \hat{x}(t) \tag{2.4}$$

Then Equation (2.3) can be re-written as

$$\dot{\tilde{x}}(t) = (A - KC)\tilde{x}(t) \tag{2.5}$$

It can be seen from (2.5) that the dynamic behavior of the error vector $\tilde{x}$ is determined by the eigenvalues of matrix *(A-KC).* If the matrix *(A-KC)* is stable, the error vector will converge to zero for any initial error vector $\tilde{x}(0)$. Hence the residual signal $r(t)$:

$$r(t) = y(t) - \hat{y}(t) = Cx(t) - C\hat{x}(t) \tag{2.6}$$

will be very close to zero under the normal operating condition. However, if there is a fault occurred in the sensor, which can be modeled as a change in the matrix C i.e. $C = C+\Delta C$, the residual will become:

$$r(t) = C\tilde{x}(t) + \Delta C x(t) \tag{2.7}$$

where ΔC is the uncertainty caused by parameter variations, model uncertainties and external disturbances, and bounded, $|\Delta C| < \delta$, δ denotes threshold value to detect fault.

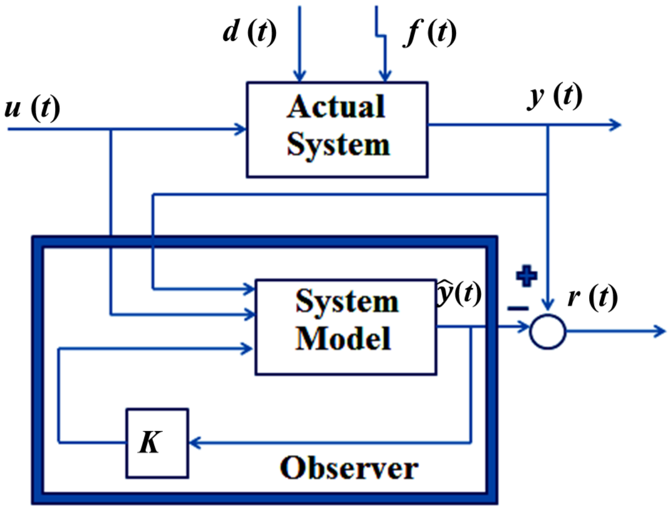


Figure 2.5. Use of Observer to Generate Residual.

It is no longer close to zero, thus testing the residual, $r(t)$, can indicate a fault. Actuator and system faults can be modeled similarly as the changes in the matrices A and B respectively. Analysis will also show that the residual would deviate from zero if a fault in the system or the actuator should occur.

2.4.2. Robustness

Model-based FDD methods are based on mathematical models; however, a precise and accurate model of a real system might not be easy to obtain. There are some obvious reasons; e.g., unknown structure of disturbances, different noise effects, and uncertain or time varying (due to aging) system parameters. FDD methods that are able to handle this kind of model uncertainty are referred to as robust. Model uncertainty can cause false and missed alarms; hence, it needs to be considered when implementing FDD systems. If it is not handled properly, it can have strong impact on FDD performance. There exist several approaches to handle the robustness issue. They are divided into two groups as active and passive robustness approaches. The active robustness approach deals with the model uncertainty in the residual generation phase. The aim is to avoid model uncertainty effects on the residuals. The passive robustness approaches are implemented in the residual evaluation phase, e.g., by using time varying thresholds, also known as adaptive thresholds. For further details about robust FDD, the research results of Chen and Patton (1999) and Frank and Ding (1997) can be seen.

2.5. 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 to fault detection from different perspectives. Towards that goal, we have classified the methods into two categories: (i) model-based methods; and (ii) data-driven-based methods. We have also compared the advantage and disadvantage of these methods. In this chapter, we interested with model-based methods and Lunberger observer-based fault detection scheme is proposed.

3. DATA DRIVEN-BASED FDD: CONCEPT AND THEORY

For large scale systems it is often difficult to use model-based methods because of the lack of accurate models (Alkaya and Eker, 2011). When large volumes of process data are available as in a modern state-of-the-art plant, data-based technologies provide an alternative approach to process monitoring that partially circumvents difficulties associated with model-based methods. This is particularly appealing route as modern industrial processes are characterized by high instrumentation and process automation and, thus it is not uncommon to have large amounts of data collected every few seconds on such plants. In principle, data-based approaches exploit structure or regularities in data to derive mathematical or statistical models that describe expected process behavior under normal operating conditions. The derived models can then be used for monitoring, control and process optimization tasks. Data-driven process monitoring statistics based on multivariate methods and their applications in fault detection in industrial processes are briefly introduced in this section.

3.1. Multivariate Statistical Process Control (MSPC)

Performance monitoring and early detection of abnormal events is critical in achieving set product quality objectives as well as general continuous process improvement. Examples of such abnormal events include among other, drifts and shifts in the mean or the variance of one or more process variables. To this end, a range of statistical process monitoring techniques has been proposed as a means for achieving stated plant objectives. These included classical charting techniques such as Shewhart, Cumulative Sum (CUSUM), and Exponentially Weighted Moving Average (EWMA) control charts used in monitoring the performance of processes to detect changes in process performance. However, these charts are not suitable for multivariate processes where observed variables tend to be significantly correlated. To effectively handle these cases, multivariate extensions of these univariate methods have been developed. These are based on the projection of measured variables onto latent structures. More specifically, methods based on the use of principal component analysis (PCA), partial least squares (PLS) and related variants have gained a lot of attention over the last couple of decades in the monitoring of multivariable processes (Kano et al., 2001; Kourti, 2005; Zvokelj et al., 2010; Zhiqiang et al., 2011). These groups of fault detection and diagnosis tools are generally referred to as multivariate statistical process control (MSPC) methods.

3.1.1. Classical Statistical Process Control

Classic univariate control charts analyze data at a fixed scale or resolution, which makes them detect changes at that single scale. More formally, the linear transformation of data in these charts has been done at fixed frequencies and extract features in the domain of time as illustrated in Figure 3.1 (Ganesan et al., 2004). Shewhart charts represent data at the sampling interval or at the finest scale which is effective for detecting large mean shifts. The Shewhart charts use only information about the process contained in the last observed point and ignore any information given by the entire sequence of points. This limitation of Shewhart charts can be overcome by the use of CUSUM, Moving Average (MA) and EWMA charts. On the one hand CUSUM charts represent data at the scale of all measurements or at the coarsest scale and directly incorporate all of the information in the sequence of sample values by plotting the cumulative sums of the deviations of the sample values from a target value. MA and EWMA charts, which fall in between these two extremes viz. Shewhart and CUSUM, are very effective in detecting small mean shifts. The MA chart monitors the process location over time based on the average of the current subgroups and one or more prior subgroups and hence it gives equal importance to past data within its moving window. On the other hand, in EWMA the average of the samples is computed in a way that gives less and less weight to data as they are further removed in time from the current measurement.

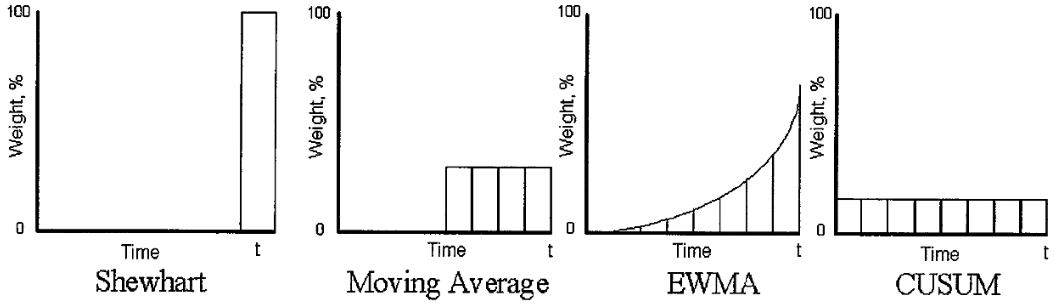


Figure 3.1. The Traditional Multivariate Control Charts.

3.1.2. Overview of Principal Component Analysis (PCA)

The data-based PCA method, initially proposed by Pearson (1901) and later developed by Hotelling (1947), is one of the popular Statistical Process Control (SPC) methods which relates to its conceptual simplicity. In a PCA method, a number of related variables are transformed to a smaller set of uncorrelated variables. PCA is a useful multivariate analysis technology, which can be used for data compression, reduction of the data dimension, feature extraction, and image compression. The method produces a lower dimensional representation in a way that preserves the correlation structure among the process variables, and is optimal in terms of capturing the variability in the data (Russell, 2000).

Kourti (2005) provided a good explanation for PCA method. When using a PCA method, Principle Components (PC) can be extracted by linearly combining the original input variables. A simple schematic interpretation of PCA is illustrated in Figure 3.2. Suppose that there are five variables in a process. Notice those variables x_1, x_3 and x_4 exhibit the same pattern; they are correlated with each other for this time period. Therefore, two PCs can be used in this example. The first PC is a weighted average of x_1, x_3 and x_4, while the second PC is a weighted average of x_2 and x_5. Again, the main purpose of PCA method is to find factors that have much lower dimensions than the original data set but can still properly describe the major trends in the original data set.

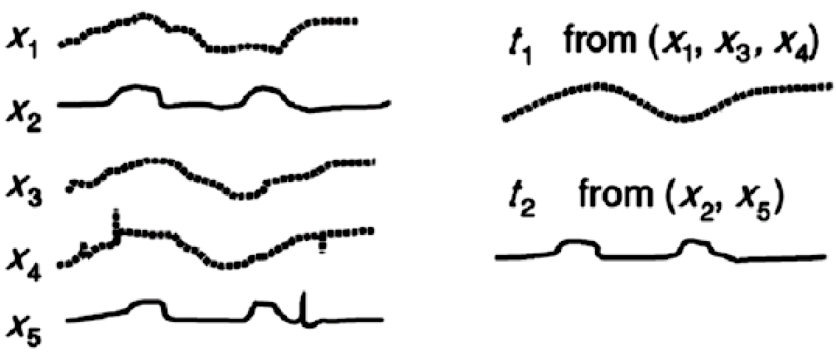


Figure 3.2. A Simplified Representation of PCA.

PCA is based on the assumption that process operates at a steady-state operating condition and each of the variables is uncorrelated in time. In practice, industrial processes exhibit dynamic behavior and, therefore, in addition to being cross correlated, variables exhibit some degree of autocorrelation arising from, for example, throughput changes, controller feedback and the presence of unmeasured disturbance. Moreover, the high sampling frequency relative to the dominant process time constant and process inertia may lead to incorrect decisions due to false alarms when using PCA. To address these and other drawbacks, several extensions of PCA have been developed to account for non-Gaussianity, autocorrelation and nonlinearity in observed data. These are introduced in the current sections (Shun, 2009).

3.1.2.1. Dynamic PCA (DPCA)

Ku et al. (1995) showed that a linear time-series relationship can be incorporated into the conventional PCA analysis. Dynamic PCA (DPCA) model can be extracted from the data arranged to represent an ARX model structure. For a dynamic system, the current values of the variables depend on the past values. The steady state PCA approach may be extended to model and monitor dynamic systems by augmenting the data matrix to include time-lagged variables:

$$X = \left[x_1(t)\ x_1(t-1),\ ...,\ x_2(t)\ x_2(t-1),...,\right] \tag{3.1}$$

The augmented data matrix uses appending lagged time-series modeling to extract the time-dependent relations in the measurements. DPCA extends the capability of conventional static PCA to be used for dynamic multivariate system and was proven to be effective at small disturbance detection.

3.1.2.2. Moving PCA (MPCA)

Moving PCA (MPCA), proposed by Kano et al. (1999), is based on the idea that a change of operating condition can be detected by monitoring directions of PCs. In the following sections, principal component is abbreviated as PC.

In order to detect a change of PCs, the reference PCs representing a normal operating condition should be defined, and the differences between the reference PCs and the PCs representing a current operating condition should be used as indexes for monitoring. The index A_i can be used for evaluating the change of PCs,

$$A_i(t) = 1 - \left| w_i(t)^T w_{i0} \right| \tag{3.2}$$

where, $w_i(t)$ denotes the ith PC calculated at time t, and w_{i0} denotes the reference of ith PC. Both w_i and w_{i0} are unit vectors. The index A_i is based on the inner product, i.e. the angle of PCs. When the ith PC representing a current operating condition is equivalent to its reference, A_i becomes 0. On the contrary, A_i becomes 1 when w_i is orthogonal to w_{i0}.

3.1.2.3. Nonlinear PCA

Conventional PCA has been found to perform well when applied to steady-state linear processes without serious dynamics. For more complicated cases with particularly nonlinear characteristics, PCA performs poorly due to its assumption that the process data are linear. A nonlinear PCA technique, called Kernel Principal Component Analysis (KPCA) has emerged in recent years as an effective approach to solve the problem of nonlinear data (Cho et al., 2005). The basic idea of KPCA is to map input vectors into a high-dimensional feature space via the appropriate kernel function, which helps to relate input space by some nonlinear mapping. Then PCA is performed in the projected feature space.

Yoo and Kyoo (2006) proposed a new dynamic nonlinear monitoring method that combined KPCA and an exponentially weighted moving average (EWMA) for biological wastewater treatment processes. Biological wastewater treatment processes have several features similar to Air Handling Unit (AHU) operations. First, most process changes occur slowly and continuously. Second, the processes exhibit strong non-stationary and dynamic characteristics. The kernel functions of KPCA can capture the nonlinearity of bioprocesses and EWMA can catch the dynamics of bioprocesses. The monitoring results on bioprocesses showed that this method was better at detecting small shifts than existing static, linear and nonlinear monitoring methods. Therefore, the results indicated that this method is an appropriate tool to supervise process stability and to analyze nonlinear bioprocesses, yielding a fast and robust monitoring system.

3.1.2.4. Multi-Block PCA (MBPCA)

When using multi-block PCA (MBPCA), a large data matrix is decomposed into smaller matrices of blocks to allow easier modeling and interpretation of a large data matrix. The PCA model is then developed for each block, as well as for multiple blocks together to capture the relationship between the sub-blocks. The blocks are defined based on physical knowledge about the system that being modeled, such as variables measured on distinct equipment or corresponding to different regimes of operation. The approach provides greater insight into the data than conventional PCA.

Qin et al. (2001) explored the orthogonal properties of MBPCA and Partial Least Squares (PLS) algorithms. The use of MBPCA and PLS for monitoring and diagnosis is derived in terms of regular PCA and PLS scores and residuals, which can be identified in the contribution plot. While the multi-block analysis algorithms are basically equivalent to regular PCA and PLS, blocking of process variables in a large-scale plant based on process knowledge helps to localize the root cause of the fault. New definitions of block and variable contributions to Q and T^2 are proposed and successfully applied by Qin et al. (2001).

3.1.2.5. Recursive PCA (RPCA)

A major limitation of PCA is that PCA model, once built from the data, is time-invariant, while most real processes are time-varying. Frequent external condition changes can cause process fluctuations and result in variables that have (i) changes in the mean value (ii) changes in the variance (iii) changes in the correlation structure among variables. As most industrial processes experience slow and normal time-varying behaviors, recursive PCA (RPCA) method is expected to have a broad applicability. RPCA efficiently updates the model by recursively calculating the correlation matrix, determining the number of PCs, and confidence level for Q and T^2, which are indices for fault detection.

Li et al. (2000) presented a monitoring strategy that built a RPCA model with a moving time window. PCA model was updated at fixed time intervals to overcome the problem of changing operation conditions, which commonly demonstrated slow time varying behaviors. The updated elements of RPCA included sample-wise update, recursive determination of number of principal components, and confidence limits for Q and T^2 in real time to facilitate adaptive monitoring.

3.1.2.6. Multi-Scale PCA (MSPCA)

Another way to handle changing process conditions is to use the wavelet transform method. The method can be used to decompose a signal into different scales of decreasing level of detail or resolution. Multi-scale PCA (MSPCA) combines the ability of PCA to extract the cross-correlation relationship between the variables with the ability of wavelet transform method to extract the auto-correlation features in the measurements. MSPCA monitors process measurement at different time-scale by decomposing measurement data into separate frequency bands. This method increases the sensitivity of fault detection, which makes it possible to detect small but significant events in data displaying large variations.

As an application of this approach to fault detection, MSPCA has been proposed by Bakshi (1998). Wavelet analysis partitions data set into frequency intervals (scales) and each scale is modeled locally by PCA method. MSPCA extracted relationships between the variables such as supply air temperature and humidity by PCA, and between the samples by wavelet analysis. It is similar to multi-block PCA in that both methods decompose the overall monitoring statistics. However, the multi-block methods block the information according to variables, whereas the multi-scale methods block the data with respect to the wavelet coefficients at different scales.

3.2. PCA Based Fault Detection Method

Modern industrial processes are large-scale interconnected systems. Thus, efficiency of any data-driven monitoring scheme depends upon its ability to compress a huge amount of process data and extract the meaningful information within. One of the most common multivariate statistical process control (MSPC) methods used for this purpose is principal component analysis. PCA has been used for various multivariate data analysis techniques such as process monitoring, quality control, sensor and process fault diagnosis (Wang et al., 2004; Xiao et al., 2006; Youming and Lili, 2009; Wang et al., 2010; Youming and Lili, 2010). In the present section, the general principle of using PCA for fault detection is presented.

3.2.1. Data Reduction and Information Extraction

Principal component analysis is a data-based multivariate statistical technique that was developed primarily to explain the variance-covariance structure of a set of correlated variables through a few linear combinations of these variables. Usually, data collected from industrial processes contains redundancy due to multiple measurements of the same variable or due to constraints or linear relationships between variables. PCA separates this redundancy by decomposing the data into a few key independent components, which will describe the major trends in the processes, thereby reducing the number of variables to be monitored.

Let X , $X \in \Re^{\mathrm{nxm}}$ represent a data matrix, *n* denotes number of measurements, *m* denotes number of physical variables, and *n*>*m* (He et al., 2006). The data matrix must be normalized to zero mean and unity variance with the scale parameter vectors $\overline{x}$ and ϵ as the mean and variance vectors respectively. Using PCA, the data matrix X can be decomposed as shown in Figure 3.3(a):

$$X = \hat{X} + E \tag{3.3}$$

$$\hat{X} = \hat{T}\hat{P}^{T} \tag{3.4}$$

$$E = \tilde{T}\tilde{P}^{T} \tag{3.5}$$

where, X is Principal Component Subspace (PCS), it represents the correct direction of the measured vectors. *E* represents residual subspace (RS), it is the direction of faulty measurements. *E* is noise or uncertain disturbance mostly, when the measurements are fault free. $\hat{T}$ is score matrix, $\hat{T} \in \Re^{\mathrm{nxa}}$, $\hat{T} = X\hat{P}$. $\hat{P}$ is *loading* matrix, $\hat{P} \in \Re^{\mathrm{nxa}}$. "*a*" is principal components (PCs) number of the model. The columns of $\hat{P}$ are eigenvectors of the correlation matrix associated with the "*a*" largest eigenvalues and the columns of $\tilde{P}$ are the remaining *m – a* eigenvectors.

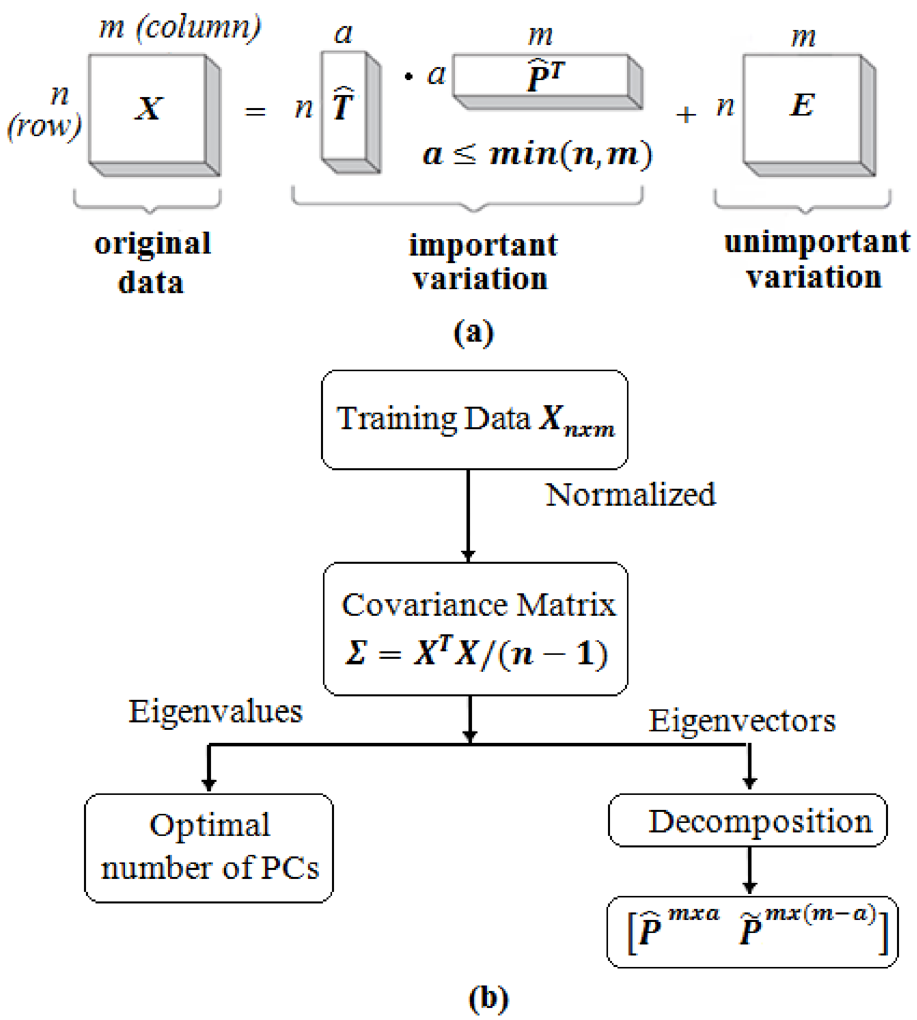


Figure 3.3. a) Reduced-Subspace Matrix, b) PCA Modeling Process.

As shown in Figure 3.3(b) (Youming and Lili, 2009), the PCA modeling process is composed of the following steps:

1. The normalization of the original variables.

   It is necessary to normalize the data before using the PCA. In our application, the data is normalized to zero mean, unit variance to eliminate the undesired weight difference in the variation of different variables with different magnitudes. The correlation PCA approach treats every variable equally and gives each variable an equal weight in the total variance.

2. Calculation of the covariance matrix $\sum$ (Xiao et al., 2009):

$$\Sigma = \frac{1}{n-1} X^T X \tag{3.6}$$

3. Singular Value Decomposition (SVD) is performed as (Bin and Yang, 2008):

$$\Sigma = V \Lambda V^T \tag{3.7}$$

   where Λ is a diagonal matrix that contains the eigenvalues (λ_i) of the covariance matrix Σ sorted in decreasing order $(\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m \geq 0)$ in the diagonal locations. Columns of the matrix V are the eigenvectors of Σ.

4. Determine the optimal number "*a*" of PCs.

   To select the proper number of principal components "*a*", most used approaches in literature are given of the end of this section.

   The *Cumulative Percentage Variance (CPV)* approach (Wold et al., 1987) which is commonly used in FD problems is used in present case such that the cumulative variance percent (=cumulative sum of the variances captured by each PC) is considered to select the PC for which, say, over 90% of the cumulative variance is captured. The variances captured by the PC's are calculated using the eigenvalues computed from (3.7) that is related to the data matrix *X*:

$$CPV(a) = \frac{\sum_{i=1}^{a} \lambda_i}{trace(\Sigma)} 100 \tag{3.8}$$

5. Choose loading matrix $\hat{P}$ according to the PCs number "*a*".
6. The projection matrix Ω and $\tilde{\Omega}$ are calculated using the loading matrix $\hat{P}$ as:

$$\Omega = \hat{P}\hat{P}^T \qquad \tilde{\Omega} = \tilde{P}\tilde{P}^T = (I - \Omega) \tag{3.9}$$

The original *m* dimension data space is substituted by the "*a*" PCS and "*m-a*" RS, and then the correlations of variables are removed. After the PCA model has been built, when new measurements are collected, the PCA model can be used for the fault detection.

3.2.2. Statistics Associated with PCA Models

Statistic PCA models, that are built using historical data under normal operating condition, can be used to monitor correlations among sensor measurements in a dimension-reduced subspace. Statistics are used as fault indexes, which will increase significantly to abnormal level if sensor faults occur. There are two monitoring statistics that are the Hotelling's T^2-statistic and *Q*-statistic or Square Prediction Error (SPE).

3.2.2.1. T^2-Statistic

Normal operations can be characterized by employing T^2-statistic method proposed by Hotelling (1947):

$$T^2 = X^T \hat{P} \Lambda_a^{-1} \hat{P}^T X \tag{3.10}$$

where Λ_a is a squared matrix formed by the first "*a*" rows and columns of Λ.

The process is considered *normal* for a given confidence level $(100(1-\alpha)\%)$ if:

$$T_\alpha{}^2 = \frac{(n^2-1)a}{n(n-a)} F_\alpha(a, n-a) \tag{3.11}$$

where $F_\alpha(a, n-a)$ is the critical value of the Fisher-Snedecor distribution with *n* and $n - a$ degrees of freedom and α the level of significance. α takes the values between

95% and 99% as recommended by Antory (2007). The T^2-statistic with (3.11) defines the normal process behavior, and an observation vector outside this region indicates that a fault has occurred.

3.2.2.2. *Q*-Statistic

The portion of the measurement space corresponding to the lowest $m - \alpha$ singular values can be monitored by using the Q-statistic method, developed by Jackson and Mudholkar (1979) as**:**

$$Q = r^T r \qquad r = (I - PP^T)x \tag{3.12}$$

The threshold for the Q-statistic method can be calculated using its approximate distribution:

$$Q_\alpha = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \tag{3.13}$$

with

$$\theta_i = \sum_{j=a+1}^{m} \lambda_j^i \qquad h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \tag{3.14}$$

where c_α is the value of the normal distribution and α is the level of significance. When an unusual event produces a change in the covariance matrix structure of the model, it will result in a high value of Q.

3.2.3. Geometrical Interpretation of the Monitoring Statistics

The geometrical interpretation of the Hotelling's T^2- and the Q-statistics is illustrated in Figure 3.4 for a two-dimensional (2D) plane formed by the first and second PCs. Point A shows the orthogonal deviation of a new sample perpendicular to the ellipse plane model, while point B shows the horizontal deviation of a new sample from the center of the ellipse plane model. The deviation represents a serious effect of the abnormal situation to the process. The further away this deviation is from the ellipse plane model the more serious the effect of the fault which has occurred (Antory, 2007).

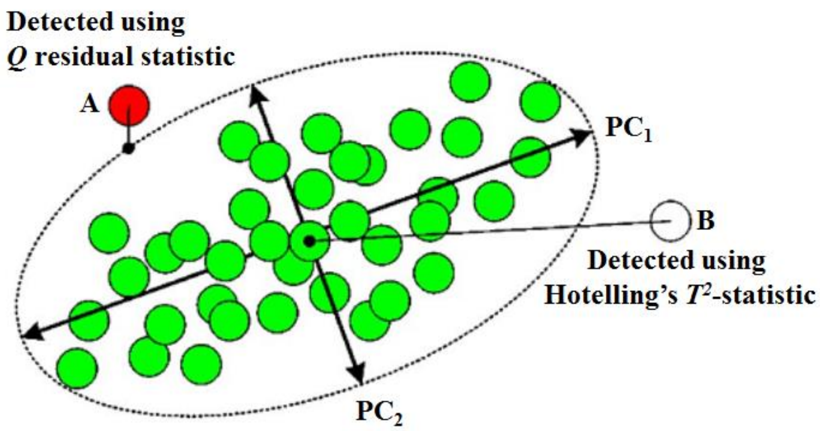


Figure 3.4. Geometric Interpretation of the Monitoring Statistics.

3.2.4. Proposed Threshold for T^2-Statistics

Classical PCA methods for FD use a data collected from a steady-state process to monitor T^2- and Q-statistics with fixed threshold which are calculated in (3.11) and (3.13). For the systems where the transient values of the processes must also be taken into account, the usage of fixed threshold in a PCA method causes the false alarms which significantly compromise the reliability of the monitoring systems.

To overcome this problem a new threshold is proposed in section 3.2.4.2 as a ***first contribution*** of this research. The proposed threshold is generated as a combination of fixed and adaptive threshold (combined threshold, T_{comb}).

However, missing alarm signal appeared due to the high variance occurred as a result of high noise components in magnitude, since the complete raw signal (unfiltered) is used in the research.

In section 3.2.4.3, to overcome the issues pointed in section 3.2.4.2, the variance sensitive adaptive algorithm is proposed as a ***second contribution***. Now in the present section the issue caused to the missing data is cleared using a proposed algorithm being in a straightforward fashion.

In literature the algorithms developed have been used only for regulation systems not for the servo systems or variable set point conditions. The proposed algorithm is valid for both the regulation and servo systems, for variable set point conditions.

3.2.4.1. Adaptive Threshold (T_{adp})

Considering the variation of the T^2 signal defined in (3.10) with the data matrix X which consist of all measured data, the mean and variance of the T^2 signal can be expressed according to the stochastic theory (Bhattacharya and Waymire, 1990):

$$\mu\left(X,t\right)=\frac{1}{n}\sum_{i=1}^{n}T_i^2(t) \tag{3.15}$$

$$\sigma^2\left(X,t\right)=\frac{1}{n-1}\sum_{i=1}^{n}T_i^2(t)-\mu\left(X,t\right)^2 \tag{3.16}$$

where μ, σ^2 and n are the mean, variance and number of data sample respectively. From the statistical theory (Wang, 2003) the confidence limits of the mean that represent a confidence of $(1-\alpha)$ is

$$P\{\bar{\mu} - z\sigma < \mu < \bar{\mu} + z\sigma\} = 1 - \alpha \tag{3.17}$$

where α is the confidence level, and z is the coefficient related to the confidence level. z is calculated for zero mean $(\bar{\mu} = 0)$ and standard deviation is equal to unity $(\sigma = 1)$. Therefore (3.17) can be rewritten as:

$$P\{-z < \mu < z\} = 1 - \alpha \tag{3.18}$$

The coefficient z is obtained from the cumulative distribution function $\varnothing(z)$:

$$\varnothing(z) = P\{\mu < z\} = 1 - \frac{\alpha}{2} \tag{3.19}$$

$$z = \varnothing^{-1}(1 - \frac{\alpha}{2}) \tag{3.20}$$

The confidence $(1-\alpha)$ is typically selected to be 95–99% in practice. From (3.17), the adaptive threshold of the mean for the T^2 signal can be calculated as (Wang, 2003):

$$T_{adp} = \mu(T^2, t) \pm z\sigma(T^2, t) \tag{3.21}$$

3.2.4.2. Combination of Fixed and Adaptive Threshold (T_{comb})

Fixed threshold (T_α) in (3.11) provides the chosen confidence limit. However it causes the false alarms in the transient state of the system. The adaptive threshold (T_{adp}) eliminates the false alarm but eliminates confidence limit. The fixed (T_α) and adaptive (T_{adp}) thresholds are combined as T_{comb} to overcome the false alarms and to provide reasonable confidence limit as:

$$T_{comb} = \begin{cases} T_\alpha & if \quad T_\alpha \geq T_{adp} \\ T_{adp} & if \quad T_\alpha < T_{adp} \end{cases} \tag{3.22}$$

where, T_{comb} eliminates the false alarms arising from the transient state. But the high variance which occurs during the variations of the states in transient conditions and measurements noise results in very high T_{comb}. This causes to produce the missing fault signal components.

3.2.4.3. Variance Sensitive Adaptive Threshold (T_{vsa})

To overcome the outlined drawbacks the threshold should be sensitive to the size of the variance. High variance is obtained if $\sigma(T^2,t) \geq \mu(T^2,t)$. Instead, if the variance $\sigma(T^2,t)$ is taken to be equal to the mean $\mu(T^2,t)$ the high variance will be reduced to a reasonable level:

$$T_{adp} = \mu\left(T^2,t\right) \pm z\mu\left(T^2,t\right) = \mu\left(T^2,t\right)\left(1 \pm z\right) \tag{3.23}$$

New threshold T_{vsa} called "variance sensitive adaptive threshold" can be given as:

$$T_{vsa} = \begin{cases} T_\alpha & if \quad T_\alpha \geq T_{adp} \\ T_{adp} = \mu(T^2,t) \pm z\sigma(T^2,t) & if \quad T_\alpha < T_{adp} \ and \ \sigma < \mu \\ T_{adp} = \mu\left(T^2,t\right)\left(1 \pm z\right) & if \quad T_\alpha \leq T_{adp} \ and \ \sigma \geq \mu \end{cases} \tag{3.24}$$

where the parameter T_α in (3.24) provides the confidence limit, the relation $\mu(T^2,t) \pm z\sigma(T^2,t)$ overcomes the false alarm in the transient state and the last relation $\mu\left(T^2,t\right)\left(1 \pm z\right)$ reduces high variance effect and eliminates missing fault signal component. The overall fault detection issues described in the present section are illustrated in Figure 3.5.

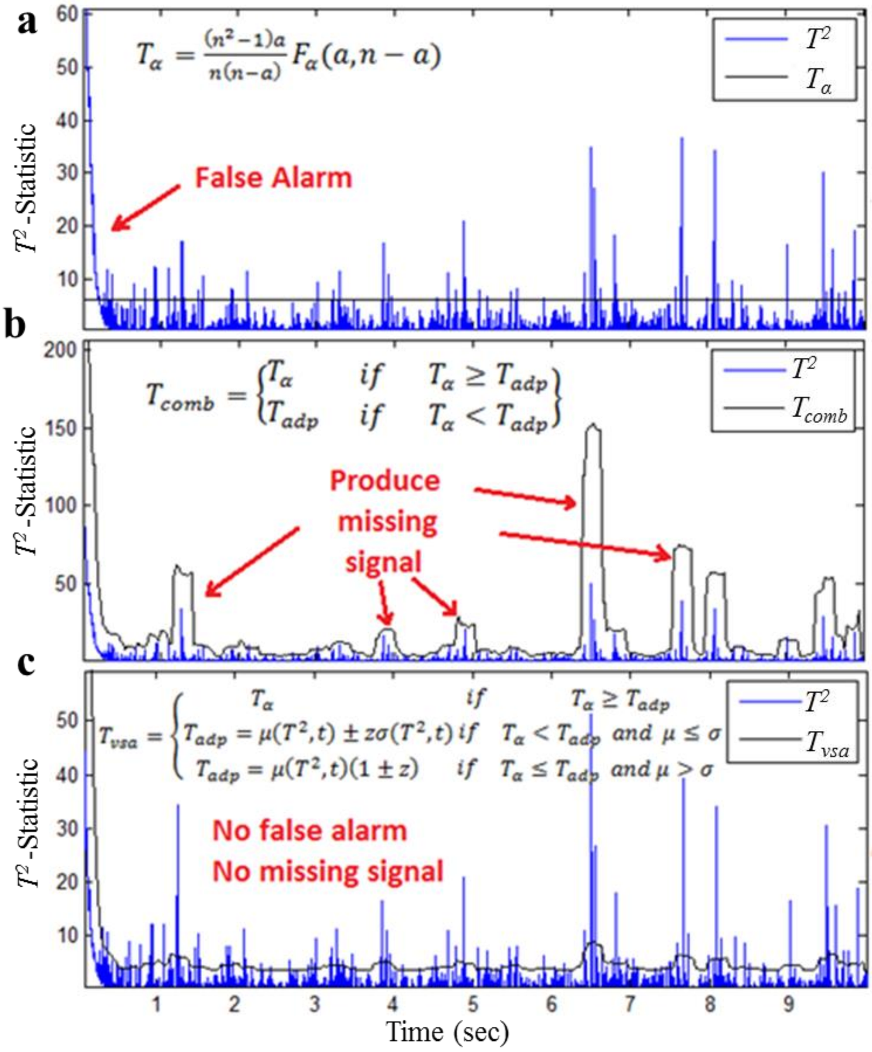


Figure 3.5. a) Fixed Threshold T_α , b) Combination of Fixed and Adaptive Threshold T_{comb} , c) Variance Sensitive Adaptive Threshold T_{vsa} .

3.2.5. Fault Diagnosis Using Contribution Plots

When a fault has been detected using the T^2- and Q-statistics, it is important to identify the cause of the out-of-control status. This can be achieved using contribution plots. In a PCA model two types of contribution plots are needed to identify the fault since two types of multivariate control charts are used, i.e., by Q-statistic for residuals and Hotelling's T^2-statistic for systematic variations within the model structure (Teppola et al., 1998). PCA contributions plots are defined as the contribution of each process variable to the individual score of the T^2- or Q-statistic. Note that the role of variable contribution plots in fault identification is to show which of the variables are related to the fault rather than to reveal the actual size of it. The variables with high contribution to the contribution plots are simply the signature of such faults (Kourti, 2005).

3.2.5.1. Contribution Plots: Hotelling's T^2-Statistic

For the T^2 statistic value of an observation, the variable contributions to an out-of-limits value are obtained as a bar plot of the mean of the absolute value of $T\sqrt{\Lambda^{-1}}P'$ which shows how each variable is involved in the calculation of T^2 value at that point. T is the matrix containing the score values of all the variables at that scale and P is the corresponding loading matrix. The matrix Λ is a diagonal matrix of the eigenvalues. The inverse of this matrix normalizes the score values of different PCs. In order to decide whether the individual variable contribution to the T^2 value is significant or not, one can either compute control limits for the contribution plots or one can compare the size of the variable's contribution under faulty conditions with the size of the same variable's contribution under normal operating conditions. In other words variables with the largest contribution to the T^2 value often indicate the source of the fault. The control limit for individual variable contribution will be the length of T^2 interval that is the square root of the T^2-limit (Jackson, 1991; Johnson and Wichern, 1992; Teppola et al., 1998).

3.2.5.2. Contribution Plots: Q-Statistic

When an out-of-control situation is detected using the Q-statistic, bar graphs of the ratio of residual variance of each variable in the testing and training set show the variations of each process variable in the residual space. This is computed by generating the residual matrix E_{new} and E_{old} of the testing and training data set by the following equation:

$$E_{new} = X_{new}\left(I - PP^T\right) \tag{3.25}$$

where X_{new} is the new data matrix (testing data) and P is the loading matrix containing the retained PCs in the PCA model.

Similarly,

$$E_{old} = X_{old}\left(I - PP^T\right) \tag{3.26}$$

where X_{old} is the old data matrix (training set). Then finding the ratio of residual variance, that is $var\left(E_{new}\right)/var\left(E_{old}\right)$, can assist in identifying the variables responsible for the variations in the residual space. '*var*' denotes variance of the signal. Variables with large variation in the residual space will show a large value of the residual variance and will be also be out of the control limits of the Q-statistic.

3.2.6. Selecting the Proper Number of Principal Components (PCs)

A key issue in developing a PCA model is to choose an adequate number of principal components to represent the process in an "optimal" way (Valle et al., 1999). If fewer principal components are selected than required, a poor model will be obtained which has an incomplete representation of the process. On the contrary, if more principal components than necessary are selected, the model will be over-parameterized and will include a significant amount of noise. Based on the available literature, a simple but reliable CPV (Cumulative Percent Variance) method has been chosen for our application. The other approaches are given in the current sections.

3.2.6.1. Prediction Sum of Squares (PRESS)

The PRESS (**PRE**diction **S**um of **S**quares) procedure was introduced by Wold (1994). As summarized by Fourie (2000), the procedure is as follows:

If m is the number of variables and n is the number of samples, then the data can be represented in $n \times m$ matrix X. These data are then randomly divided time-wise in g groups. The first group is removed from the sample and the PCA is performed on the remaining samples. The first principal component and then the first two (continuing until all m components are used) are then used to predict values of the deleted sample. For each predicted observation of the deleted sample, obtain the SPE (Q) statistic.

After this has been completed, the removed sample should be replaced and another group should be removed. The calculation of the Q-statistic is then calculated as before again.

After this has been repeated for all g groups, then Q-statistic should then be summed for each a type of model (e.g. for the two component model). The *PRESS* statistic is then formed as follows:

$$PRESS_a = \frac{1}{nm}\sum_{i=i}^{n} SPE_{ai} \tag{3.27}$$

Now, to check if the addition of the a^{th} principal component is warranted, Θ is calculated as follows:

$$\Theta = \frac{\left(PRESS_{a-1} - PRESS_a\right).D_R}{D_M.PRESS_a} \tag{3.28}$$

with

$$\begin{cases} D_M = n + p - 2a \\ D_R = p\left(n-1\right) - \sum_{i=1}^{a}\left(n + p - 2i\right) \end{cases} \tag{3.29}$$

If $\Theta > 1$, then the a^{th} principal component should then be retained and testing of the $(a + 1)^{th}$ component should be tested in the same way. If $\Theta < 1$ that component need not be included. It is possible that after the first occurrence of $\Theta < 1$ that later values of a will produce occurrences of $\Theta > 1$. This may be due to outliers (Fourie, 2000).

This technique is overly complex as compared to the other techniques. It does have the advantage of being quantitative, making it suited for computer calculation.

3.2.6.2. The Broken Stick Method

This method makes use of the amount of variance explained by each component (Martinez and Martinez, 2004). If a line is *randomly* divided into p (corresponding to the maximum number of components or original variables) segments, then the expected length of the a^{th} longest piece is:

$$g_a = \frac{1}{p}\sum_{i=a}^{p}\frac{1}{i} \tag{3.30}$$

If the proportion of variance explained by the a^{th} component is greater than g_a, then the amount of variance that the component explains is greater than expected by pure chance. It would then be useful to keep this component.

3.2.6.3. The Size of Variance Technique

Using the correlation technique, principal components with variance greater than 1 ($l_a \geq 1$) would be retained (Martinez and Martinez, 2004). For the covariance technique, the component would be kept if its variance was greater than 70% of the average of all the variances, i.e.

$$\begin{cases} l_a \geq 0.7\bar{l} \\ \quad or \\ l_a \geq 1.0\bar{l} \end{cases} \tag{3.31}$$

Occasionally, a cut of value of 100% rather than 70% is used. This method may be preferred due to its simplicity and robustness (Lee et al., 2004). The justification is simply that the principal components contributing less than the average variance are probably insignificant.

3.2.6.4. The Scree Plot Method

The scree plot is a graphical method to gauge the amount of variance contributed by each component. It is a bar plot of l_a against a (the index of the component). The cumulative variance is also plotted as a line. To use the plot, the point where the line or the slope of a line between the blocks represents the value of l_a levels off. An example is shown in Figure 3.6. In this example, between 2 and 4 principal components would be selected. A variant of this plot is the log-eigenvalue plot. This plot is used when the first few eigenvalues are much larger than the rest (Martine z & Martinez, 2004).

Other methods include the Akaike Information Criteria (AIC) mentioned in Lee et al. (2004).

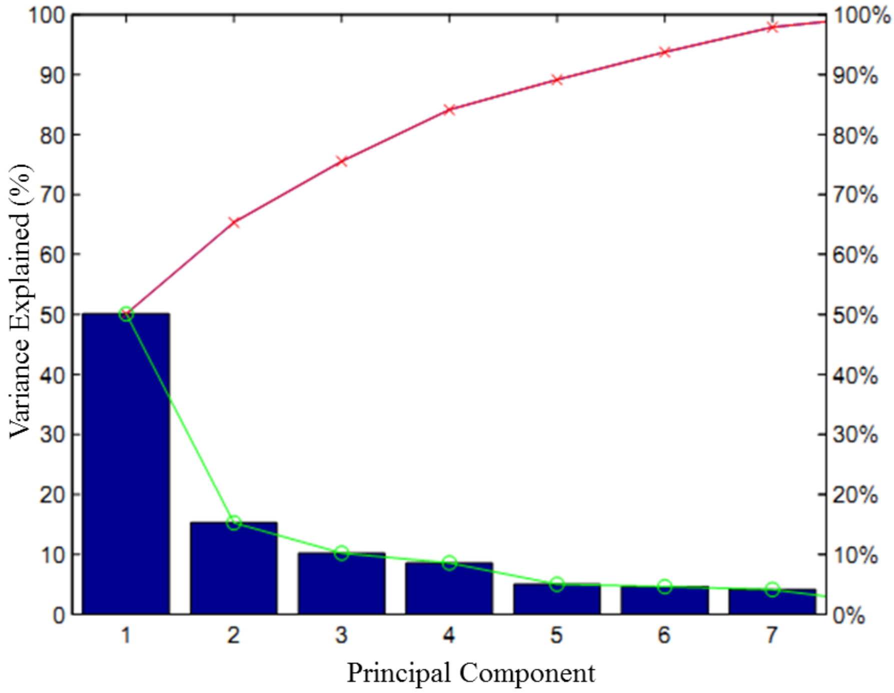


Figure 3.6. An Example of a Scree Plot.

3.3. Wavelet Based Fault Detection Method

Fault detection and diagnosis (FDD) is useful for ensuring the safe running of machines. Signal analysis is one of the most important methods used for FDD, whose aim is to find a simple and effective transform to the original signals. Therefore, the important information contained in the signals can be shown; and then, the dominant features of signals can be extracted for fault diagnostics. Hitherto, many signal analysis methods have been used for fault diagnostics, among which the Fourier Transform (FT) is one of the most widely used and well-established methods. Unfortunately, the FT-based methods are not suitable for non-stationary signal analysis and are not able to reveal the inherent information of non-stationary signals. However, various kinds of factors, such as the change of the environment and the faults from the machine itself, often make the output signals of the running machine contain non-stationary components. Usually, these non-stationary components contain abundant information about machine faults; therefore, it is important to analyses the non-stationary signals (Pan and Sas, 1996).

To resolve this deficiency, in 1946 the physicist Gabor (1964) motivated by quantum mechanics, modified the Fourier transform to analyze only a small section of the signal at a time. Gabor's adaptation, called the *Short-Time Fourier Transform* (STFT), maps a signal into a two-dimensional function of time and frequency (Figure 3.7). It provides some information about both when and at what frequencies a signal event occurs, but its precision is limited by the size of the time window used. Its other weakness is that once one chooses a particular window size, which remains the same for all frequencies. The time-frequency window of any STFT is rigid; in many applications we need a more flexible approach where we can vary the window size to examine an event more accurately either in time or frequency.

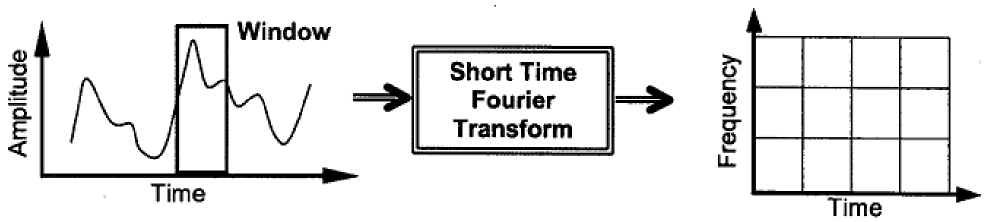


Figure 3.7. Short Time Fourier Transform

However, STFT uses a fixed tiling scheme, i.e., it maintains a constant aspect ratio (the width of the time window to the width of the frequency band) throughout the analysis (Figure 3.8). As a result, one must choose multiple window widths to analyze different data features localized in the time and frequency domains. Hence, the STFT is badly adapted to signals where patterns with different frequencies appear. It also fails to efficiently resolve short time phenomena associated with high frequencies (Li, 2002). In recent years, time-frequency methods, such as wavelet-based Multi Resolution Analysis (MRA) have gained popularity in the analysis of both stationary and non-stationary signals. These methods provide excellent time-frequency localized information, which is achieved by varying the aspect ratio as shown in Figure 3.9 Hence, wavelet-based multiscale methods analyze time and frequency localized features simultaneously with a high resolution, and also is adaptable to transient signals.

Other advantages of the wavelet-based methods include the following:

(i) The ability to de-noise signals;

(ii) At each scale, the wavelet coefficients are de-correlated even if the input data is auto-correlated;

(iii) The wavelet coefficients are normally distributed irrespective of the input data distribution;

(iv) The wavelet coefficients are stationary even if the input data is non-stationary.

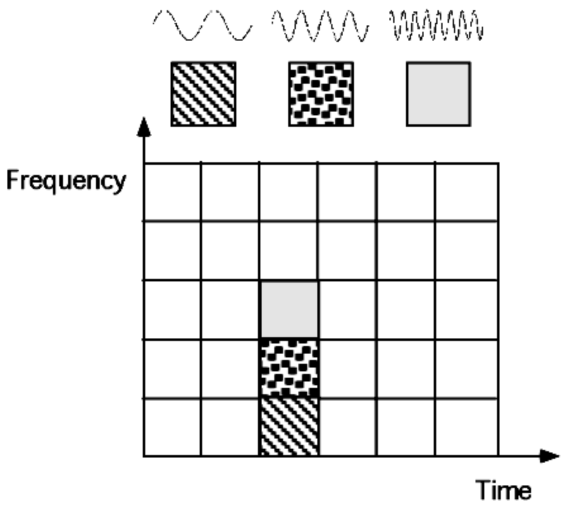


Figure 3.8. STFT with Fixed Aspect Ratio.

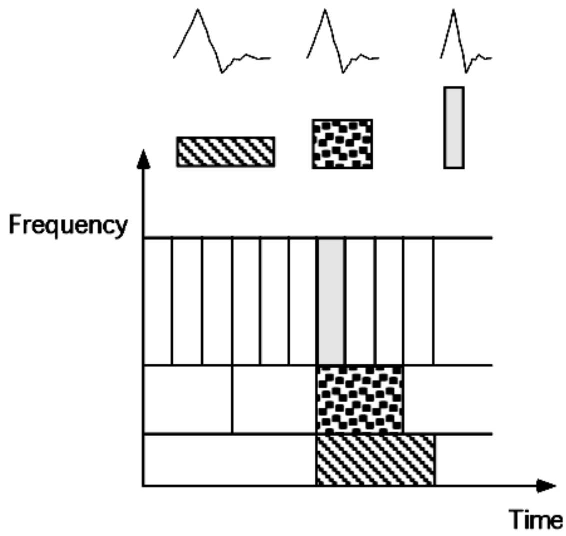


Figure 3.9. Wavelet Transform with Variable Aspect Ratio (Daubechies Functions).

3.3.1. History of Wavelet and Application Area

The wavelet was first mentioned by Alfred Haar in 1909 in his research. In the 1930's, Paul Levy found the scale-varying Haar basis function superior to Fourier basis functions. The transformation method of decomposing a signal into wavelet coefficients and reconstructing the original signal again is derived in 1981 by Jean Morlet and Alex Grossman. In 1986, Stephane Mallat and Yves Meyer developed a multiresolution analysis using wavelets. They mentioned the scaling function of wavelets for the first time; it allowed researchers and mathematicians to construct their own family of wavelets using the derived criteria. Around 1998, Ingrid Daubechies used the theory of multiresolution wavelet analysis to construct her own family of wavelets.

Transient nature of machine signals and search for a particular time-frequency behavior for diagnostic purposes render wavelets as highly suitable for the analysis of such signals. Some of the reasons for the use of wavelets in machine diagnosis applications are given (Tafreshi, 2005):

- **Wavelets as Time-Frequency Analysis Tools.** Wavelets are mainly time-frequency analysis tools. They are highly suitable candidates for machine data analysis as information about a given machine operation lie both in time and frequency behavior of the signal.

- **Wavelets and Localized Signal Analysis.** As stated earlier, machine data utilized for diagnosis are highly transient where information about a given machine condition reside in local behavior of the signal; i.e., changes occurring in part or the entire segment of the signal. Wavelets are highly suitable to capture localized changes and behavior.
- **Wavelet Coefficients as Feature Variables.** Signal expansion by wavelets often leads to a few wavelet coefficients of large magnitude and large number of coefficients of small magnitude. This leads to signal approximation with limited number of large amplitude coefficients used as feature variables. Considerable reduction of dimensionality is achieved in this manner.
- **Wavelets as Unconditional Bases.** Signal information lies in coefficient values obtained from wavelet signal decomposition. Wavelets are unconditional bases (Burrus et al., 1998) which imply a very robust basis in which the coefficients drop off fast independent of the sign of the coefficients. Therefore, in orthogonal signal decomposition, absolute values of the coefficients carry the necessary information about the signal. This allows using absolute values of wavelet coefficients for feature extraction.
- **Noise Reduction using Wavelets.** Wavelets are used for noise reduction in which wavelet coefficients of small amplitude (below a given threshold) are set to zero. Often such coefficients belong to noise content of the signal at the highest frequency band. De-noising is different from the commonly used high frequency filtering, as it can be carried-out at all frequencies. On the other hand, we utilize de-noising scheme for reduction of machine background noise corresponding to overall acceleration observed in machine data. Often it is composed of white noise. Careful selection of threshold level in de-noising can successfully reduce machine background noise

3.3.2. Wavelet Theory

Wavelets are classes of wave-like functions that are often irregular, non-symmetric, and with no analytical/mathematical expression. They have finite number of oscillations and an effective length of finite duration. Wavelets are used as basis functions for signal decomposition and signal processing. They allow function expansion in orthogonal, non-orthogonal or redundant structures. Wavelets are considered as unconditional bases with properties that allow efficient information extraction and coding (Chui, 1992).

Wavelets are a family of basis functions which are localized in the time and frequency domains, and are obtained from a single prototype wavelet, called mother wavelet or basic function $\Psi(t)$, by scaling (frequency) and translation (shifting in Time). The wavelet family can be defined as

$$\Psi_{\eta,\tau}(t)=\frac{1}{\sqrt{\eta}}\Psi\left(\frac{t-\tau}{\eta}\right) \tag{3.32}$$

where η is the so-called scaling parameter, τ is the time localization parameter, $\tau\in\Re$ and $\eta>0$, $\Re$ denotes set of real numbers (Qiu et al., 2006). In the discrete case, the scale and translation parameters are discretized as $\eta=2^j$ and $\tau=k2^j$. (3.32) can be rewritten as:

$$\Psi_{\eta,\tau}(t)=\Psi_{j,k}(t)=2^{-j/2}\psi(2^{-j}t-k) \tag{3.33}$$

where j and k denote the scale and translation parameters, respectively. The translation parameter determines the location of the wavelet in the time domain, while the scale parameter determines the location of the wavelet in the frequency domain.

The wavelet transform of a finite energy signal $g(t)$ with the analyzing wavelet $\Psi(t)$ is the convolution of $g(t)$ with a scaled and conjugated wavelet:

$$W(\eta,\tau)=\frac{1}{\sqrt{\eta}}\int_{-\infty}^{\infty} g(t)\Psi^{*}\left(\frac{t-\tau}{\eta}\right)dt \tag{3.34}$$

where $\Psi^{*}(t)$ stands for the complex conjugation of $\Psi(t)$.

The wavelet transform $W(\eta,\tau)$ can be considered as functions of translation τ with each scale η. Equation (3.34) indicates that the wavelet analysis is a time-frequency analysis, or a time-scaled analysis. Different from the Short Time Fourier Transform (SFFT), the wavelet transform can be used for multi-scale analysis of a signal through dilation and translation so it can extract time-frequency features of a signal effectively (Peng et al., 2004).

In general, wavelet analysis uses the wavelet functions which can be stretched and translated with a flexible resolution in both frequency and time. The flexible windows are adaptive to the entire time-frequency domain, which narrows while focusing on high frequency signals and widen while searching the low-frequency background. In this way, wavelet analysis allows the wavelets to be scaled to match most of the high and low frequency signal so as to achieve the optimal resolution with the least number of base functions.

Several wavelet basis function types are available in the literature (Genesan et al., 2004). Some of these are the Haar's, Daubechies', coiflets, symlets, bi-orthogonal wavelets, etc. The Haar basis was known even before the wavelets were developed. Although Haar's has a compact support, it does not have good time frequency localization. Moreover, it is unsuitable for representing classes of smoother functions due to its discontinuities. Some of the desirable properties of the basic functions are good time-frequency localizations, various degrees of smoothness (number of continuous derivatives), and large number of vanishing moments (ensures maximum number of zeros of the polynomial at the highest discrete frequency).

The most widely used wavelet is the Daubechies' basis function which is shown by DbN, where N is the order; the greater the N, the more oscillating and smooth the wavelet. Two examples of Daubechies family, Db4 and Db8, are shown in Figure 3.10 The Haar's filter is best suited to represent step signals or piecewise constant signals, whereas the Daubechies' filter is better for smoother signals (Shao et al., 1999). Therefore, Daubechies' wavelet function is used in our applications.

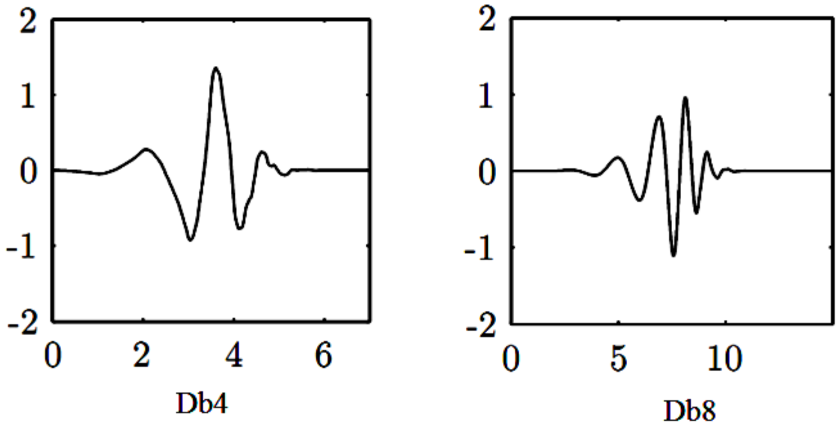


Figure 3.10. Two Examples of Daubechies Family of Wavelets.

3.3.2.1. Multi Resolution Analysis (MRA)

Any process is characterized by its parameters. For an in-control process operating under chance causes of variation will typically consist of a single feature, i.e., data collected from the process will have a stable probability distribution. However, most real life processes do not behave in an ideal way. For example, the measurements representing a process may be contaminated with noise and/or various other features due to tool failure, faults, sensor failure and machine parts degradation. This implies that the state of a process in general would not consist of a single feature but rather would have multiple features (Ganesan et al., 2004).

Thus, to identify a change in a process parameter, it is necessary to analyze the features of the data relevant to the change in both the time and frequency domains. For example, a step change in a signal (i.e., a change in the process mean) is more localized in time but not in frequency, whereas a change in the variance is more localized in the frequency domain than in the time domain. It is clear that different process features are better represented at different domains and hence should be examined accordingly. Thus, a useful approach for analyzing a process should be a time-frequency approach, which would describe the time localization as well as the frequency localization of data. One such approach is wavelet-based MRA (Ganesan et al., 2004).

Mallat (1989) combined the idea of an image pyramid structure and Wavelet analysis to develop a technique in signal processing, which is called the multi-resolution analysis (MRA). Previous research on the application of MRA in fault detection focused on the analysis of the abrupt change in signal (Mallat & Hwang, 1992; Nobuyuki et al., 1996; Paul, 1994). Using MRA, the original signals are decomposed into multi-scale components. Since the abrupt changes in signal produce significant peaks in some scales, a fault in signal can therefore easily be detected.

As shown in Figure 3.11, the wavelet transform can be used to decompose multivariate signals s into approximations a_1 and details d_1 coefficients at the first level. The difference between the first approximation a_1 and the original signal S is the detail d_1 (this is the high frequency components that are filtered out). Application of the same transform on the approximations a_1 causes them to be decomposed further into approximations a_2 and details d_2 coefficients at the second level. The decomposition process can continue to a level L as long as the length of approximation coefficients in a_l is more than the length of coefficients in the wavelet filter. The wavelet transforms work like a filter. After passing a signal through a wavelet transform filter, wavelet coefficients are generated. The wavelet transform contains a low pass filter (only obtaining low frequencies), which is denoted by L_0, and a high pass filter (only obtaining high frequencies), which is denoted by H_0. At each level, the original signal, s, passes through two both low and high pass filters and emerges as two signals, which is detail coefficients d_n, and approximation coefficients a_n. In Figure 3.11 "↓2" means the number of coefficients is halved through the filters.

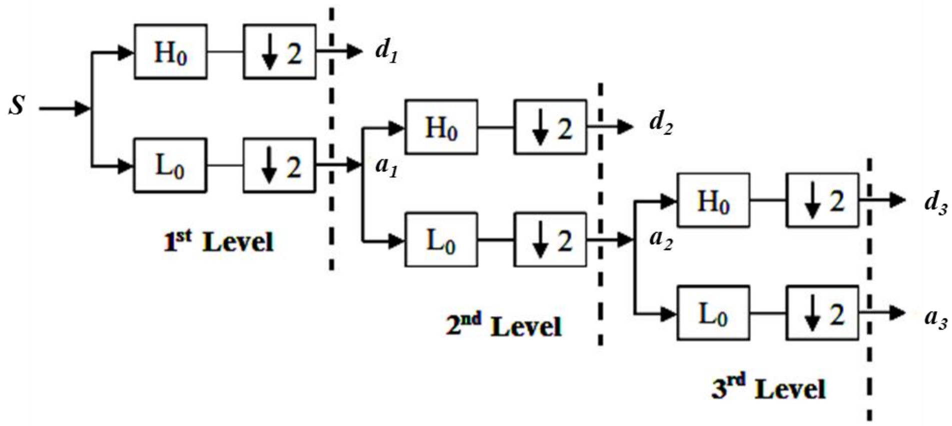


Figure 3.11. Three Level Wavelet Decomposition Tree.

3.3.2.2. Wavelet De-noising

Industrial data is also synonymous with process measurement 'noise'. Noise associated with the process measurements is known to impact upon the robustness of the process model. It is therefore desirable to extract the 'true' signal from the noise corrupted data prior to carrying out any detailed statistical analysis (Shao et al., 1999)**.** The traditional approach to filtering is to remove the high frequency components above a certain level since they are associated with noise. Small wavelet coefficients at low scales (high frequency area) are usually expected to be mainly due to noise components.

The discrete wavelet transform (DWT) is found to yield a fast computation of the wavelet transform (Xinhua et al., 2008)**.** It is computed by a successive low pass and high pass filtering of the discrete digital signal.

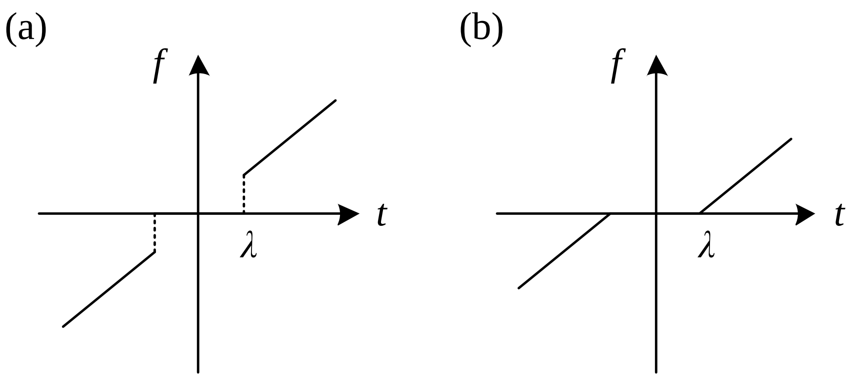
Specifically, applying the DWT $\Gamma = Wf(t)$ to the data $f(t)$, one obtains $\Gamma = \zeta + e$, where Γ, ζ, e are the collections of all coefficients, parameters and errors transformed from data $y(t_i)$ to the true data $f(t_i)$, respectively. Because smaller coefficients are usually contributed from data noises, thresholding out these coefficients has an effect of 'removing data noises' (Kunpeng et al., 2009). Donoho (1994)**,** and Donoho and Johnstone (1995) developed several wavelet-based thresholding techniques to find a smooth estimate $(\hat{f})$ of f from the 'noisy' data y.

In wavelet thresholding, after setting some coefficients to zeros, the reconstructed (denoised) signal is obtained by inverse transformation. Figure 3.12 shows hard and soft thresholding of signal $y = f(t)$. The hard thresholding method consists in setting all the wavelet coefficients below a given threshold value equal to zero, while in soft thresholding the wavelet coefficients are reduced by a quantity equal to the threshold value (Pasti et al., 1999). The resulting coefficients are then used for selective reconstruction of an estimate of the inverse of DWT:

$$\hat{f}(t) = W^{-1}\hat{\Gamma} \tag{3.35}$$

The general de-noising procedure involves three steps:

1. *Signal decomposition.* Choose a wavelet basis function. Compute the wavelet decomposition of the signal. (i.e. calculating all the wavelet coefficients such as approximation and detail coefficients.).
2. *Threshold detail coefficients.* Select threshold (soft or hard) and apply it to the detail coefficients.
3. *Signal reconstruction.* Compute wavelet reconstruction using the original approximation coefficients and recover the de-noised signal.



$$y_{hard}(t) = \begin{cases} f & |f| \geq \Upsilon \\ 0 & |f| < \Upsilon \end{cases} \qquad y_{soft}(t) = \begin{cases} \operatorname{sgn}(f)(|f| - \Upsilon) & |f| \geq \Upsilon \\ 0 & |f| < \Upsilon \end{cases}$$

Figure 3.12. Thresholding Schemes: a) Hard Thresholding, b) Soft Thresholding.

where *sgn* denotes the usual sign function. The choice of Υ is obviously crucial: small/large thresholds will produce estimates that tend to overfit/underfit the data.

3.4. Conclusion

Classical PCA methods for FD use a data collected from a steady-state process to monitor T^2-statistic with fixed threshold. For the systems where the transient values of the processes must also be taken into account, the usage of the fixed threshold in a PCA method causes the false alarms which significantly compromise the reliability of the monitoring systems. A combined algorithm (Combination of Fixed and Adaptive Threshold (T_{comb})) is proposed to overcome the problems raised from the fixed threshold and provide required confidence limit. However, the data collected from industrial process often contain measurement noise that causes to produce the missing fault signal components when the combined threshold method is used.

Two methods are proposed to overcome the outline drawbacks. The first one is variance sensitive adaptive threshold methods that sensitive to the high variance which occurs due to noise. The second one is wavelet based combined PCA method that removes the noise before combined PCA analysis.

The proposed methods overcome the false alarms which occur in the transient states according to changing process conditions and the missing data problem according to noisy signals.

4. EXPERIMENTAL SET-UP

In the following section, experimental tests are presented in order to verify the fault detection methods studied in Section 3. Experimental tests are performed on an electromechanical system and process control system. A data acquisition card (DAQ-National Instruments, NI, Model: PCI-6229, 250 kHz in speed, 16 bit) is used to communicate between the plant and the computer. Operating range of the card is ±10V for input data and control outputs. A computer (Pentium IV, 2 GHz in speed, 1 GB RAM) is used to practice the methods which are implemented in Simulink of MATLAB software (Eva, 1996).

4.1. Description of Electromechanical System

Diagram of the first experimental set-up illustrated in Figure 4.1 is used in the experiments. Output shaft speed is measured from an optical sensor (as rev/sec) and a tachogenerator (as volts) connected to the motor shaft. The slotted opto-sensor consists of a gallium-arsenide infra-red L.E.D and a silicon phototransistor mounted in a special plastic case which is transparent to light of the wavelength. A series of pulses is generated when the slotted disk that is mounted on the motor shaft, is rotated. When the shaft of the dc motor is turned, a voltage is induced at the tachogenerator terminals and it is directly proportional to the shaft speed. The specifications of the DC motor are given in Table 4.1. The motor also drives a shaft that carries disks that operate various transducers, and a tachogenerator. Speed of the dc motor corresponding to the different input armature voltages is measured in Table 4.2 to obtain the tachogenerator characteristics. It is determined that the tachogenerator has linear characteristics and gain of the tachogenerator is calculated as 2.02 volt/rad/sec. The 1Ω resistor is fitted in series with the armature to allow monitoring of the armature current by measurement of the voltage dropped across it. Since the resistor is 1Ω, the voltage measured across it in mV will directly correspond to currents in mA.

Table 4.1. DC Motor Specification

Armature Resistance	R = 6.2 ohm
No Load Current	I = 120 mA (V=12 Volt)
Stall Current	I = 1,93 A (V=12Volt)
Starting Torque	T = 7 Ncm/A
Torque Constant	K = 3.5 Ncm/A
Time Constant	t = 19.6 ms
Efficiency	Ef = 70% - 82%
Shaft Speed at No Load	2400 rpm (max)

Table 4.2. Electromechanical System Responses to Different Input Signals.

Applied input (volt) (V_a)	Speed (rev/sec)	Speed (rad/sec)	Tacho output (volt)
2.0	7	0.7330	1.46
3.0	11	1.1519	2.30
4.0	15	1.5708	3.14
5.0	19	1.9896	3.98
6.0	24	2.5132	5.04
7.0	28	2.9320	5.87
8.0	32	3.3510	6.70
9.0	37	3.8746	7.75
10.0	41	4.2935	8.60

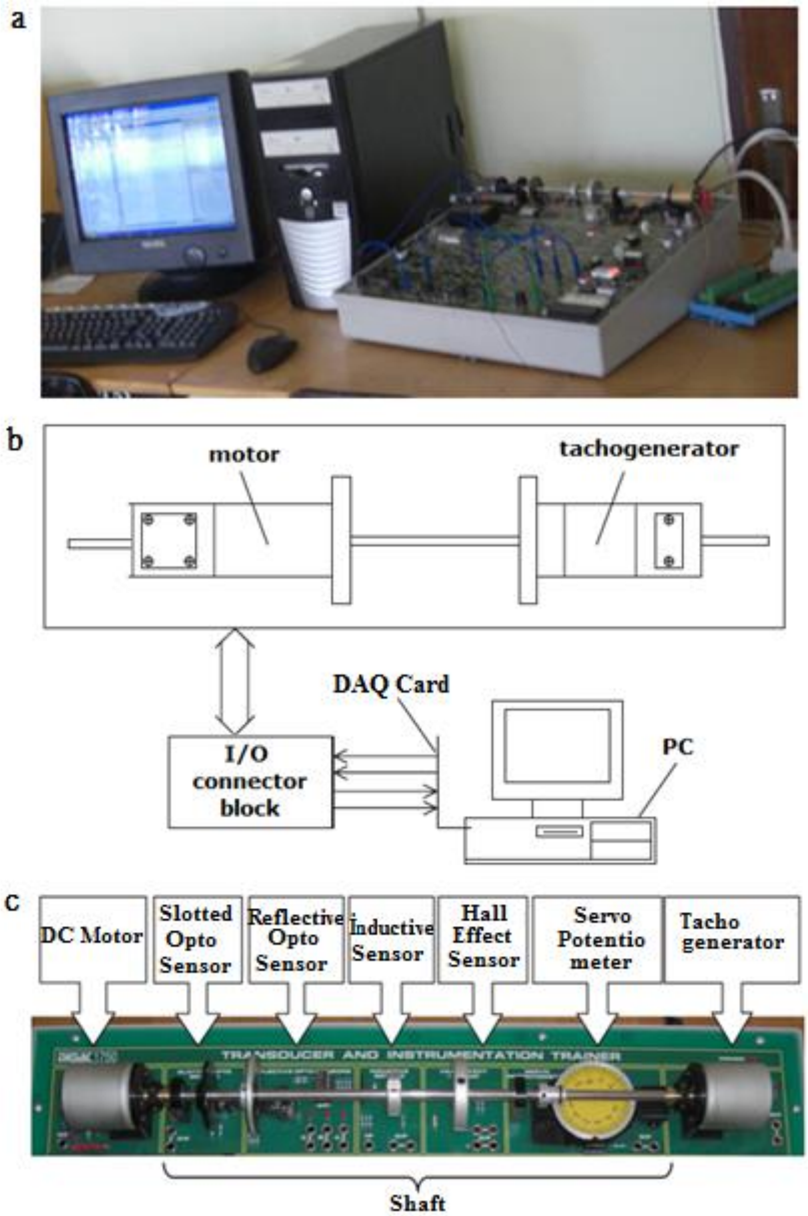


Figure 4.1. a) A Scene from the Laboratory, b) General Diagram of the Laboratory Equipment, and c) A Scene from an Electromechanical Plant.

The electrical and mechanical equations for the electromechanical plant, shown in Figure 4.2, consisting of a dc motor connected to a load via a long shaft can be given as:

$$v_a(t) = L_a \frac{d}{dt} i_a(t) + R_a i_a(t) + K_m \omega_m(t) \tag{4.1}$$

$$J_m\left(\frac{d\omega_m(t)}{dt}\right)=T_m(t)-T_s(t)-R_m\omega_m(t)-T_f(\omega_m) \tag{4.2}$$

$$J_L\left(\frac{d\omega_L(t)}{dt}\right)=T_s(t)-R_L\omega_L(t)-T_d(t)-T_f(\omega_L) \tag{4.3}$$

$$T_s(t)=k_s\left(\theta_m(t)-\theta_L(t)\right)+B_s\left(\omega_m(t)-\omega_L(t)\right) \tag{4.4}$$

$$\frac{d\theta_m(t)}{dt}=\omega_m(t) \quad \frac{d\theta_L(t)}{dt}=\omega_L(t) \tag{4.5}$$

where v_a is the motor armature voltage, R_a and L_a are the armature coil resistance and inductance respectively, i_a is the armature current, K_m is the torque coefficient, T_m is the generated motor torque, ω_m and ω_L are the rotational speeds of the motor, J_m and J_L are the moments of inertia, R_m and R_L are the coefficients of viscous-friction, T_d is the external load disturbance, T_f is the nonlinear friction, and T_s is the transmitted shaft torque, t is the independent time variable. Model of the nonlinear friction T_f can be obtained by considering an asymmetrical characteristic as (Jang and Jeon, 2000):

$$T_f(\omega)=\left(\alpha_0+\alpha_1 e^{-\alpha_2|\omega|}\right)sgn1(\omega)+\left(\alpha_3+\alpha_4 e^{-\alpha_5|\omega|}\right)sgn2(\omega) \tag{4.6}$$

where α_i denote friction constants and positive, $\alpha_i > 0$, $i=0,\ldots,5$, $\alpha_0 \neq \alpha_3$, $\alpha_1 \neq \alpha_4$, $\alpha_2 \neq \alpha_5$ and the functions sgn1 and sgn2 are defined as:

$$sgn1(\omega)=\begin{Bmatrix} 1 & \omega \geq 0 \\ 0 & \omega < 0 \end{Bmatrix} \quad sgn2(\omega)=\begin{Bmatrix} 0 & \omega \geq 0 \\ -1 & \omega < 0 \end{Bmatrix} \tag{4.7}$$

The electromechanical plant to implement theoretical results, a dc motor connected to a load via a shaft, is shown in Figure 4.2.

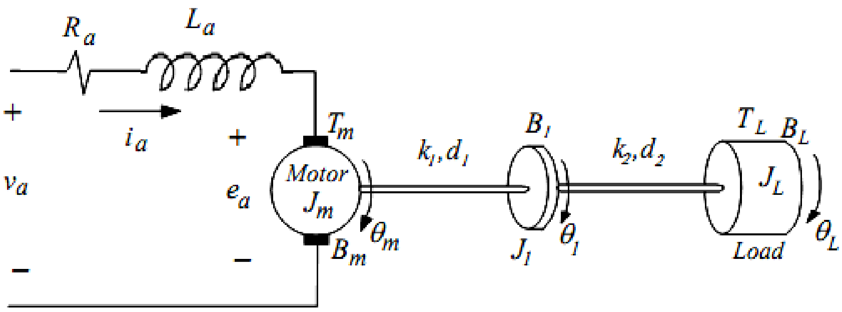


Figure 4.2. Diagram of the Electromechanical Plant.

Block diagram of the electromechanical plant is illustrated in Figure 4.3 (Eker, 2010).

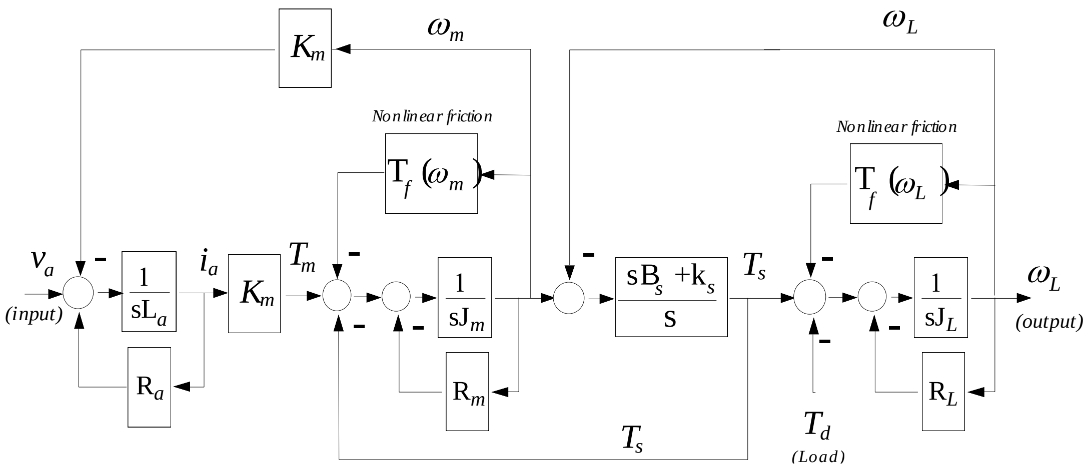


Figure 4.3. Block Diagram of the Electromechanical Plant.

4.2. Description of Process Control System

The second experimental setup used in this research is a process control set (Figure 4.4) whose components are listed in Table 4.3. It features a pilot scale process that consists of a 5 liters pressurized vessel equipped with the temperature, continuous level, and pressure sensors. The vessel liquid temperature may be controlled using an electric resistance. The feedline is equipped with a variable speed pump, a flow meter, and a proportional control valve. The outlet line has a needle valve which is adjusted as a disturbances and kept untouched for the whole set of experiments. The leaving liquid from the vessel is then collected into a sump tank, from where the feedline begins. Also, the pilot has a set of manual valves. The level and flow sensor characteristics are presented in Figure 4.5 and Figure 4.6, respectively.

Besides the process itself, the trainer includes a control module which contains the interface circuits for the sensors and the actuators, and on-off, proportional, integral and derivative control circuits.

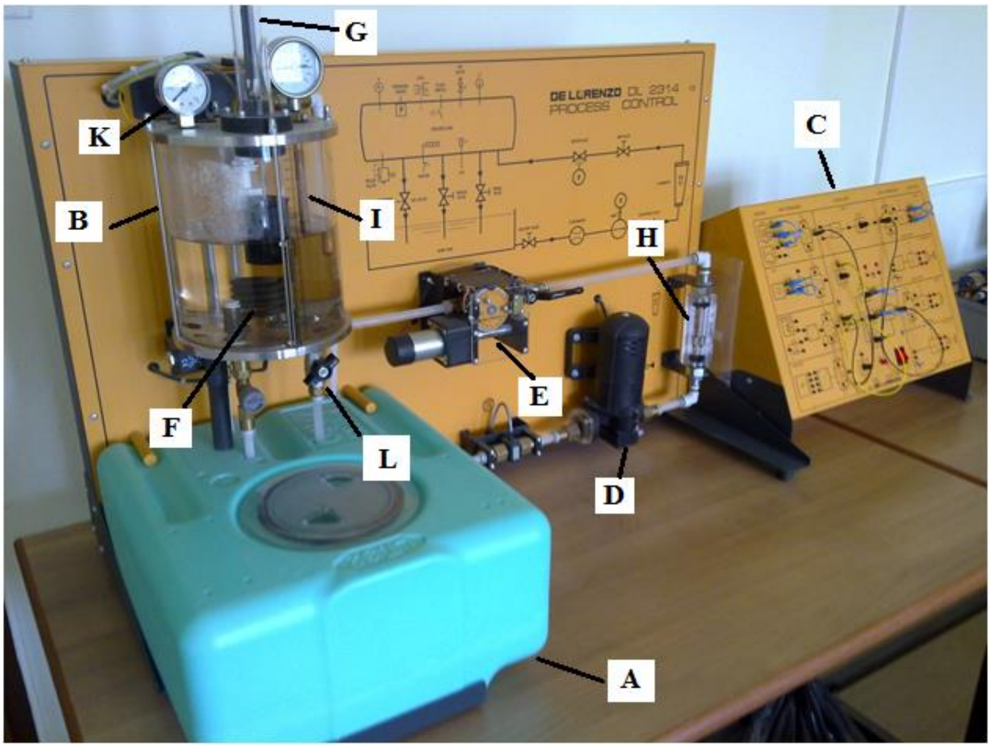


Figure 4.4. Process Control Experimental Setup.

Table 4.3. Parts of the Process Control Experimental Setup Shown in Fig. 4

A	Sump Tank (20 liters approx..)
B	Pressurized Vessel (5 liters approx..)
C	Control Module (On/Off, P, PI, PID)
D	Pump (6 liters/minute, 12 V, 1.5 A.)
E	Motor Driven Valve (4 manual valve)
F	Water Heating Resistance (48V, 200W)
G	Level Sensor (LVDT)
H	Flow Sensor (8000 pulses/liter)
I	Temperature Sensor (Pt. 100)
K	Pressure Sensor (strain gauge)
L	Needle Valve (manual)

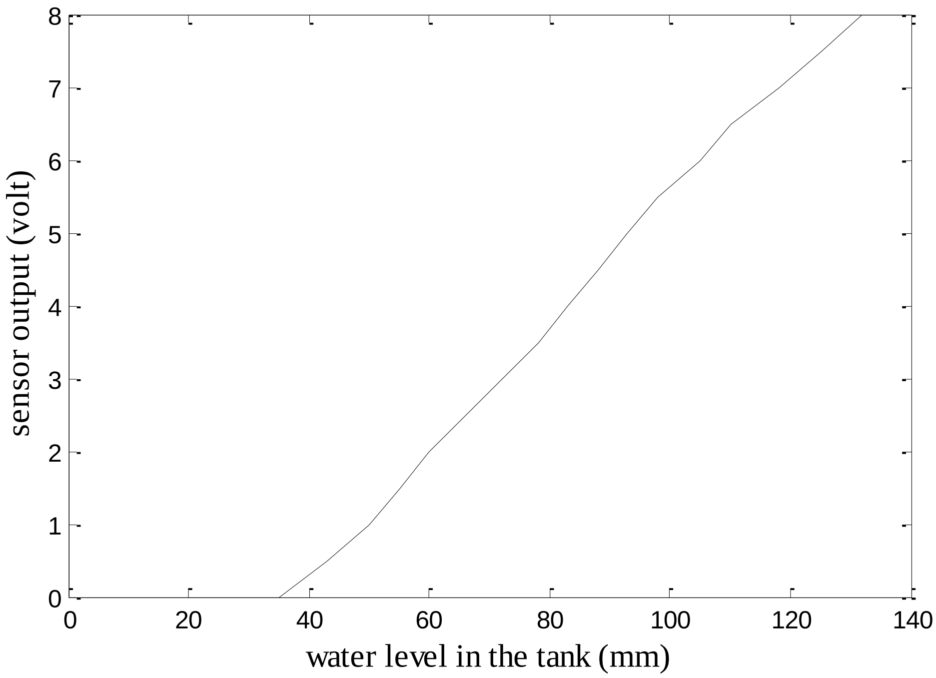


Figure 4.5. Level Sensor Characteristics.

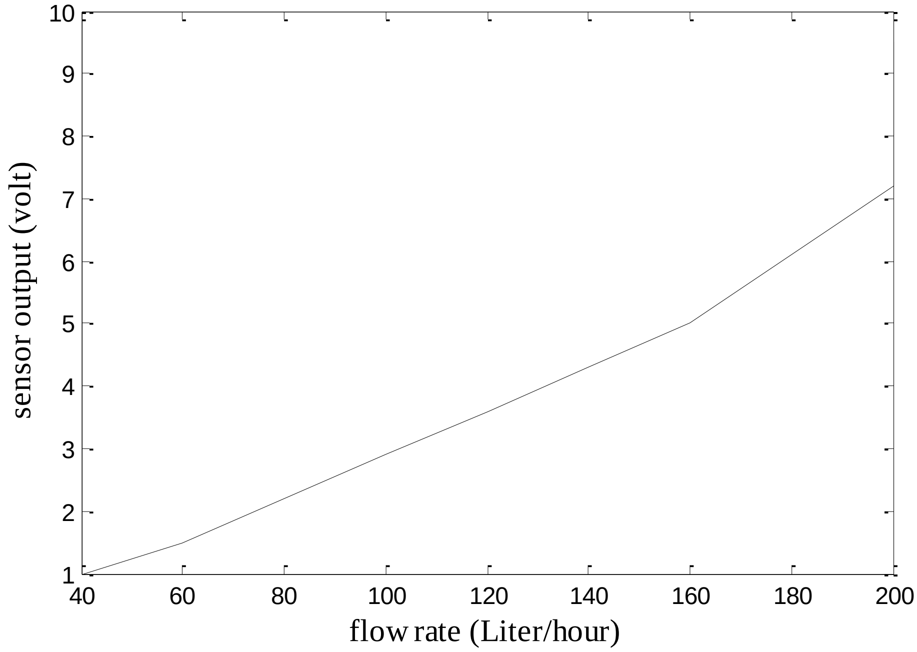


Figure 4.6. Flow Sensor Characteristics.

4.3. Data Acquisition (DAQ) Card

The same data acquisition card (DAQ-National Instruments (NI), Model: PCI-6229, 250 kHz in speed, 16 bit) is used to communicate between the plant and the computer. Figure 4.8 shows a typical DAQ system, which includes sensors, transducers, signal conditioning devices, cables that connect the various devices to the accessories, programming software, and PC. The features of the DAQ card used in the experimental applications are given in table 4.4. The data acquisition card used in experiments is illustrated in Figure 4.7.

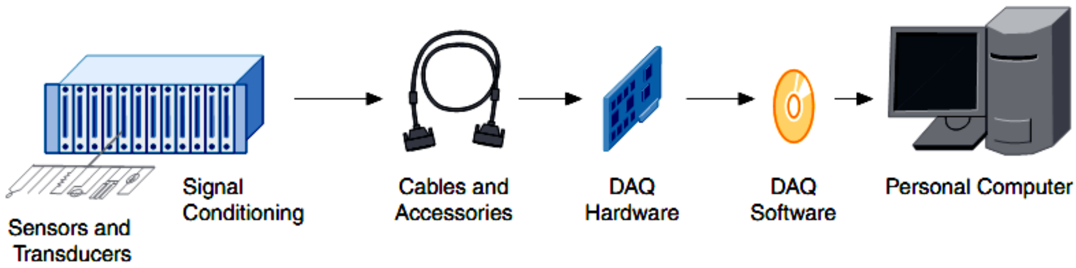


Figure 4.7. Components of a Typical DAQ System.

Table 4.4. Specifications of NI PCI-6229 DAQ Card

ANALOG INPUT	
Operating System/Target	Real-Time, Linux, Mac OS, Windows
Measurements Type	Digital, frequency, voltage
Channels	32, 16
Single-Ended Channels	32
Differential Channels	16
Resolution	16 bits
Sample Rate	250 kHz
Maximum Voltage	10 V
Maximum Voltage Range	-10 V, 10 V
Maximum Voltage Range Accuracy	3100 μV
Maximum Voltage Range Sensitivity	97.6 μV
Minimum Voltage Range	-200 mV, 200 mV
Minimum Voltage Range Accuracy	112 μV
Minimum Voltage Range Sensitivity	5.2 μV
ANALOG OUTPUT	
Channels	4
Resolution	16 bits
Maximum Voltage	10 V
Maximum Voltage Range	-10 V, 10 V
Maximum Voltage Range Accuracy	3230 μV
Minimum Voltage Range	-10 V, 10 V
Minimum Voltage Range Accuracy	3230 μV
Update Rate	833 kHz
Current Drive Single	5 mA
PHYSICAL SPECIFICATIONS	
Length	15.5 cm
Width	9.7 cm
I/O Connector	68-pin VHDCI female
TRIGGERING/SYNCHRONIZATION	
Triggering	Digital
Synchronization	Yes

Figure 4.8. Data Acquisition Card (NI PCI-6229).

4.4. Conclusion

In this chapter, the plants and equipment that will be used in experimental tests are described in detail. These are Digiac 1750 process control set, DeLorenzo CL 2314 process control set, Data acquisition card (NI, model: PCI-6229) and computer. Mathematical model of the electromechanical plant and properties of the process control systems and specifications of DAQ card are presented. This equipment is located in Control Systems Laboratory at Electrical and Electronics Engineering Department, Çukurova University, Adana, Turkey.

5. EXPERIMENTS AND RESULTS

5.1. Model-Based FD Applications

The validity of the observer based fault detection method was tested through an experimental system as shown in previous section Figure 4.1. For this propose, a series of experiments were carried out by using a DIGIAC 1750 process control set. All the experiments are performed with 6 V input voltage and sampling time T_s = 5 ms. Measured signal and calculated residuals are employed without filters to show the robustness of the designed observer model.

DC motor model is required for observer based fault detection methods. DC motor model is built in the Matlab/Simulink© environment. A functional diagram of the overall observer-based fault detection method is demonstrated in Figure 5.1 After the model has been built, the residuals are generated by feeding input data into the full order estimator, monitoring the corresponding estimated outputs from the model and comparing the model outputs with the actual measured values. The errors are considered as the residuals. Thresholds for detection and diagnosis are set by considering the maximum values reached by the residuals over a range of tests. Once the residual crosses over a certain threshold, an alarm will be triggered, indicate a fault.

The fault types that have been considered in this work are the following:

• **Drift or additive-type sensor fault:** This is a very common fault in analog sensors. Due to internal temperature changes or calibration problems, the sensor output has an added constant term (abrupt and intermittent).

• **Multiplicative-type sensor fault:** In this fault type, a multiplicative factor is applied to the sensor nominal value (incipient).

• **Sensor Failure:** This is a catastrophic fault, at a given time the sensor fails and gives a constant zero output after the failure. The fault can be due to electrical or communication problems (disconnecting the speed sensor).

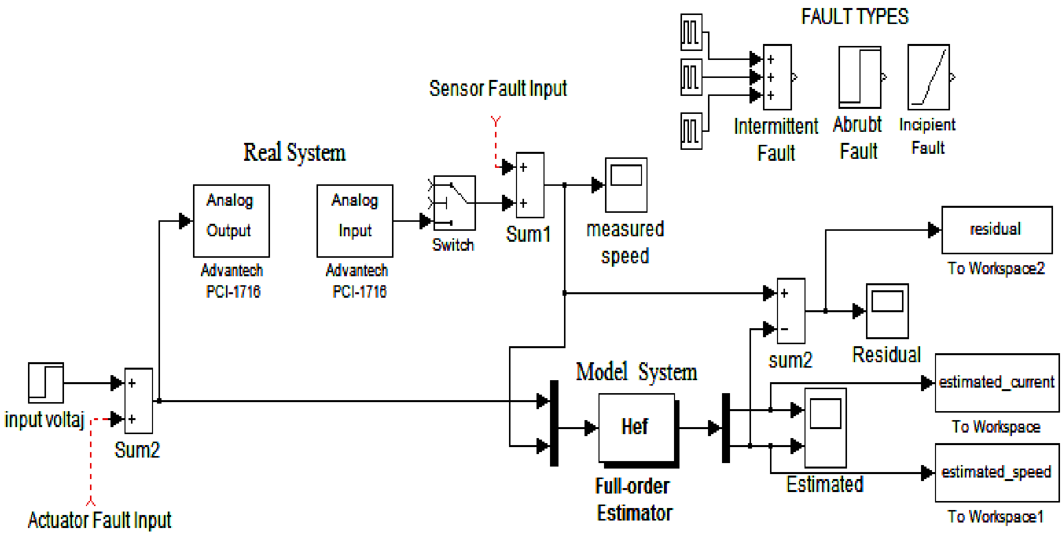


Figure 5.1. Simulink Model of the Observer Based Fault Detection Method.

5.1.1. Fault Free Case

Measured and estimated speed output of the DC motor is shown in Figure 5.2(a) and the calculated residual is illustrated in Figure 5.2(b) when there is no fault in the system.

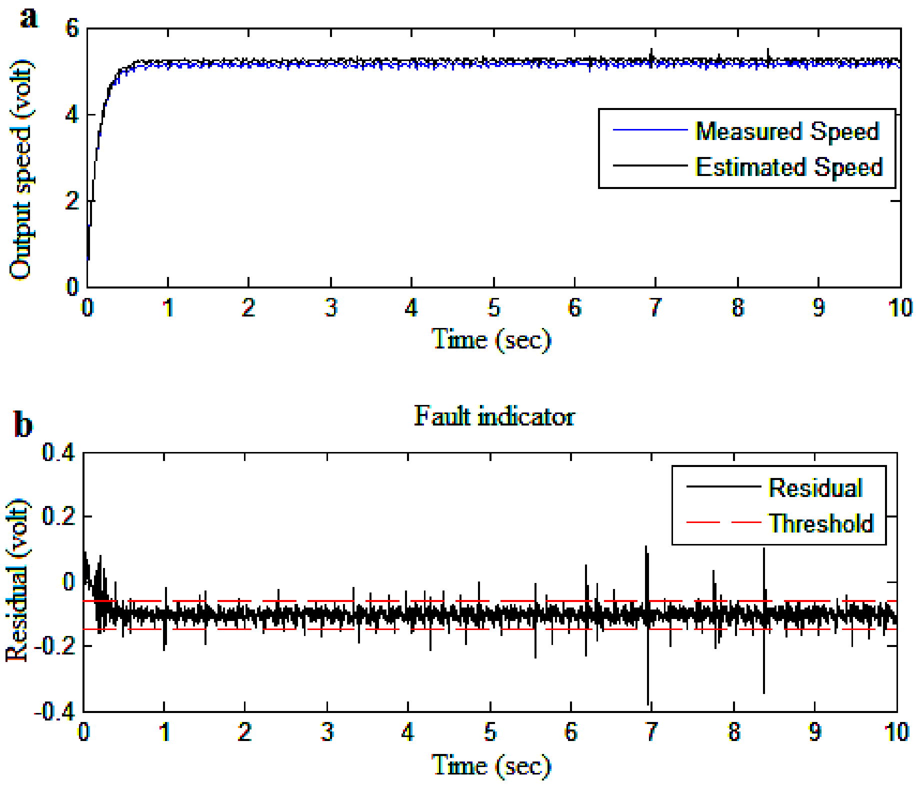


Figure 5.2. Fault Free Case a) Measured and Estimated Outputs, b) Residual.

Figure 5.2(b) shows the magnitude of the residual need not be exactly zero due to the presence of noise in the instruments and also due to errors in the observer design parameters. Hence, to avoid false alarms, a threshold of upper magnitude -0.06 and the lower magnitude -0.15 was selected for residual.

5.1.2. Abrupt Fault

Fault was modeled as a stepwise function and applied to the sensor fault input as shown in Figure 5.1. Abrupt fault was applied to the measured sensor output at 3 sec. Response of the measured and estimated output is illustrated in Figure 5.3(a) and fault indicator residual in Figure 5.3(b).

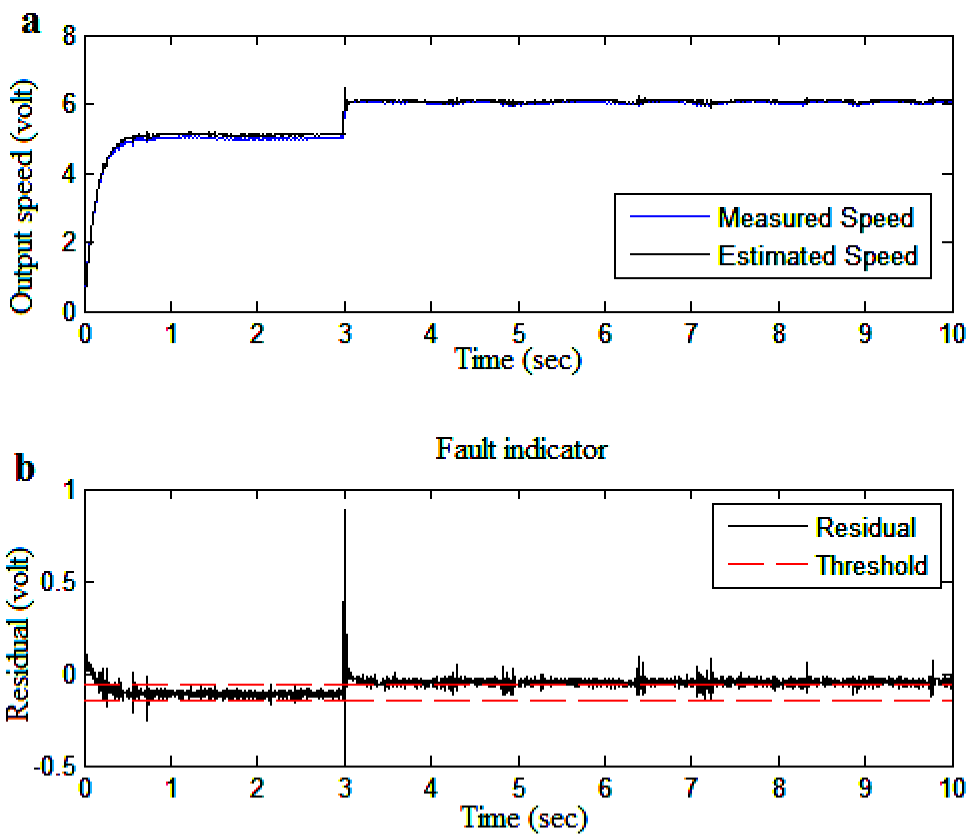


Figure 5.3. Abrupt Fault a) Measured and Estimated Outputs, b) Residual.

In this case, it is observed that the magnitude of residual shown in Figure 5.3(b) increases above the threshold value signaling the fault of the speed sensor. The time scale of Figure 5.3 is expanded around the region of occurrence of the fault for clarity and it is shown in Figure 5.4.

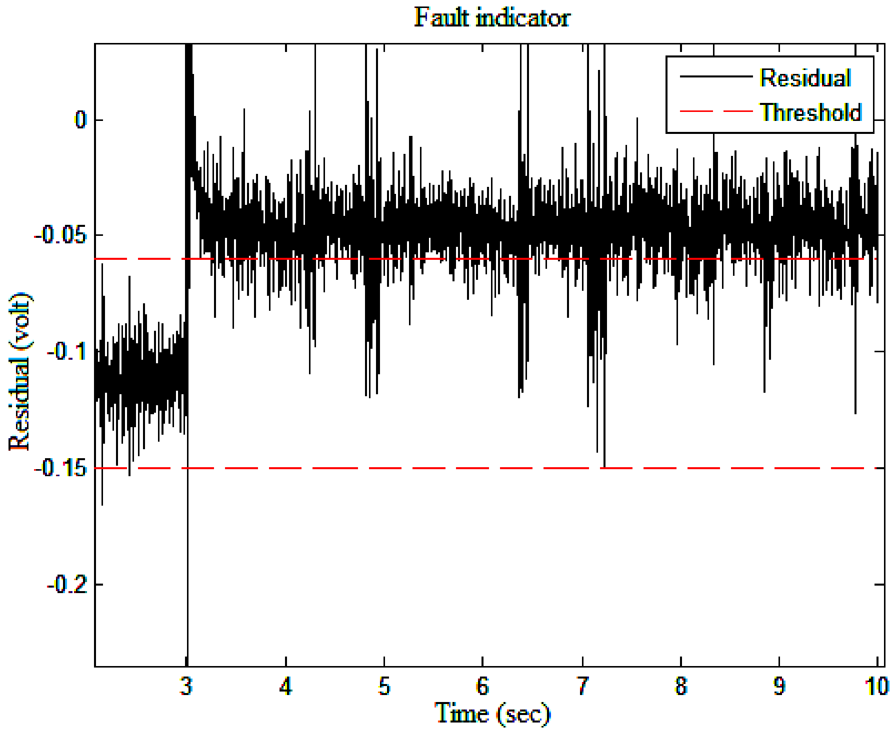


Figure 5.4. Abrupt Fault Residual.

5.1.3. Incipient Fault

Fault was modeled as a ramp function and applied to the sensor fault input as shown in Figure 5.1. Incipient fault was applied to the measured sensor output at 3 sec. Response of the measured and estimated output is shown in Figure 5.5(a) and fault indicator residual in Figure 5.5(b).

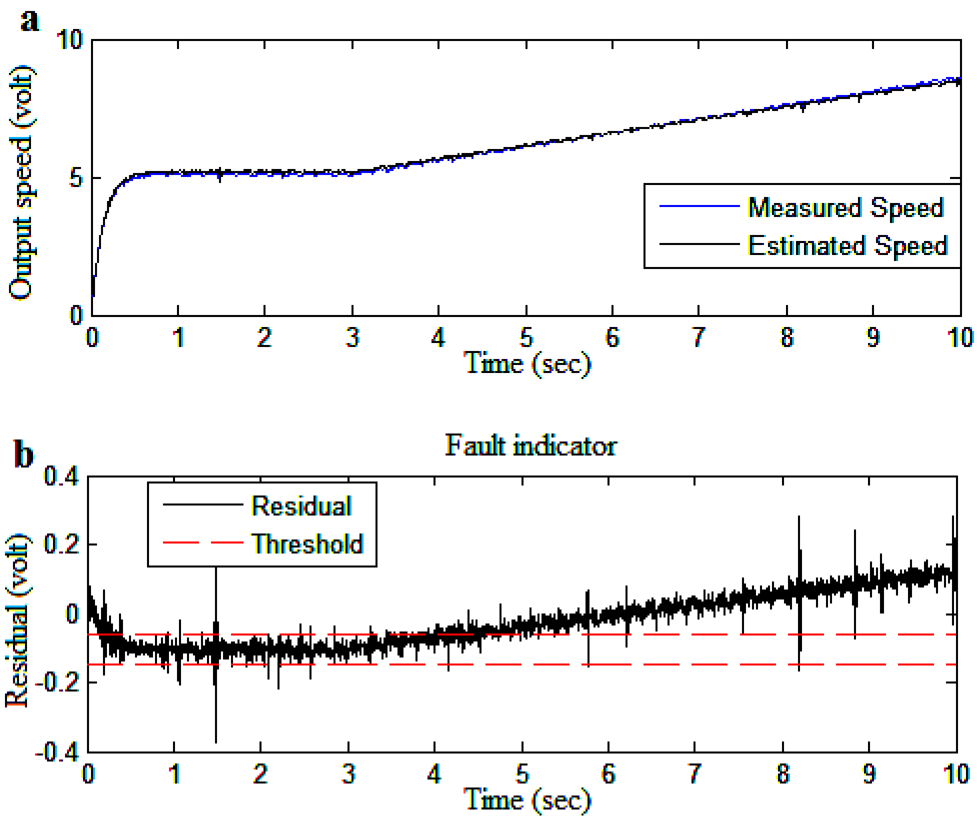


Figure 5.5. Incipient Fault a) Measured and Estimated Outputs, b) Residual.

In this case, it is observed that the magnitude of residual shown in Figure 5.5(b) incipiently increases at 3.5 sec. above the threshold value signaling the fault of the speed sensor. Hence, it is assumed that better performance would be achieved with calculating the accurate model parameters.

5.1.4. Intermittent Fault

Intermittent fault is generated as combination of impulses in different amplitudes and applied to the sensor fault input as shown in Figure 5.1. The fault is applied to the measured sensor output at 3, 5, and 7 sec with amplitude 1, 1.5 and 2 respectively in each time during 0.5 sec. Responses of the measured and estimated outputs are shown in Figure 5.6(a) and fault indicator residual in Figure 5.6(b).

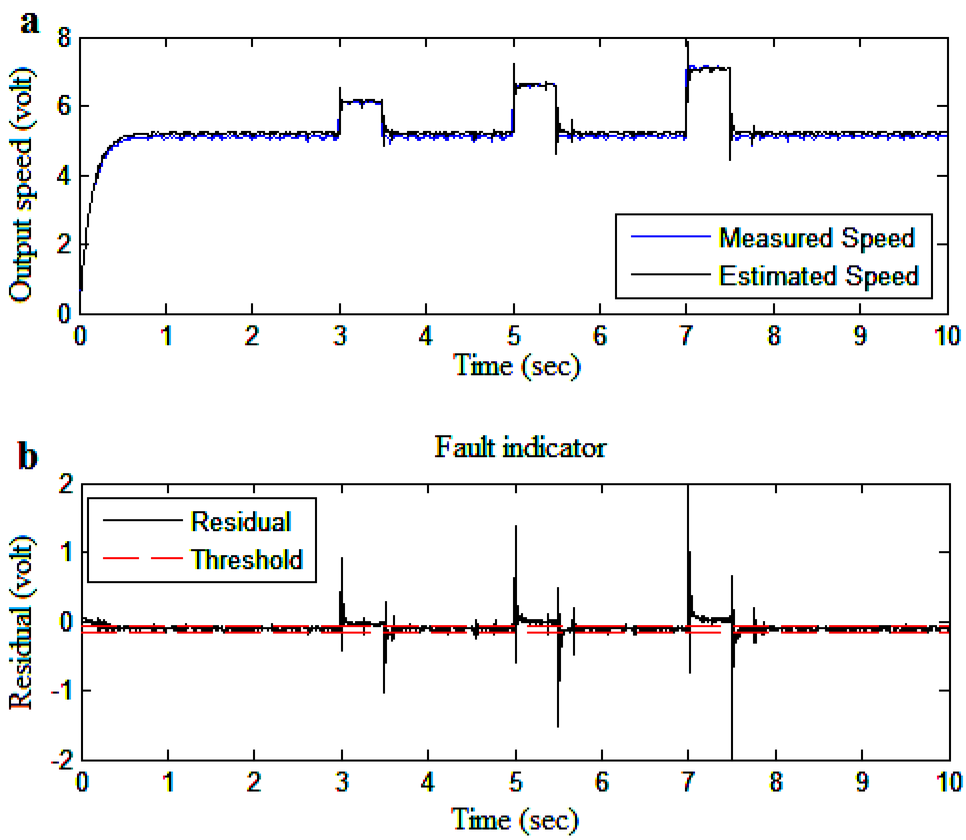


Figure 5.6. Intermittent Fault a) Measured and Estimated Outputs, b) Residual.

It is observed that the magnitude of residual shown in Figure 5.6(b) increases above the threshold value at applied fault time and size. The time scale of Figure 5.6 is expanded around the region of occurrence of the fault for clarity and it is illustrated in Figure 5.7.

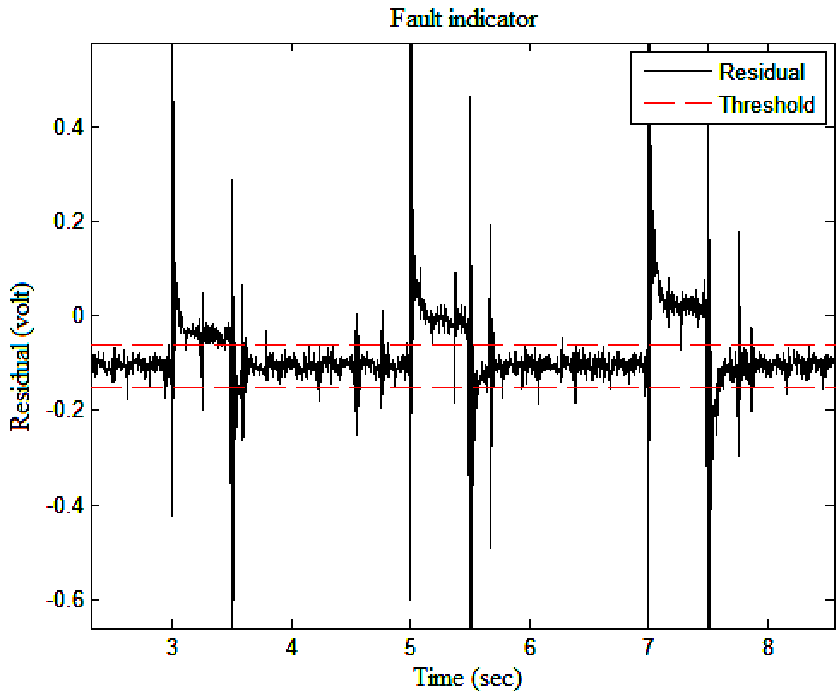


Figure 5.7. Intermittent Fault Residual.

5.1.5. Sensor Failure

In the first scenario, a fault is initially injected by disconnecting the speed sensor at 3 sec under the steady state condition. The sensor is reconnected after 5 seconds. In the second scenario, disconnecting and reconnecting operation is injected to the system under the transient condition. These two scenarios are implemented as on-line. Response of the measured and estimated outputs is shown in Figure 5.8(a) and Fault indicator residual in Figure 5.8(b) for the first scenario. The second scenario is illustrated in Figure 5.9.

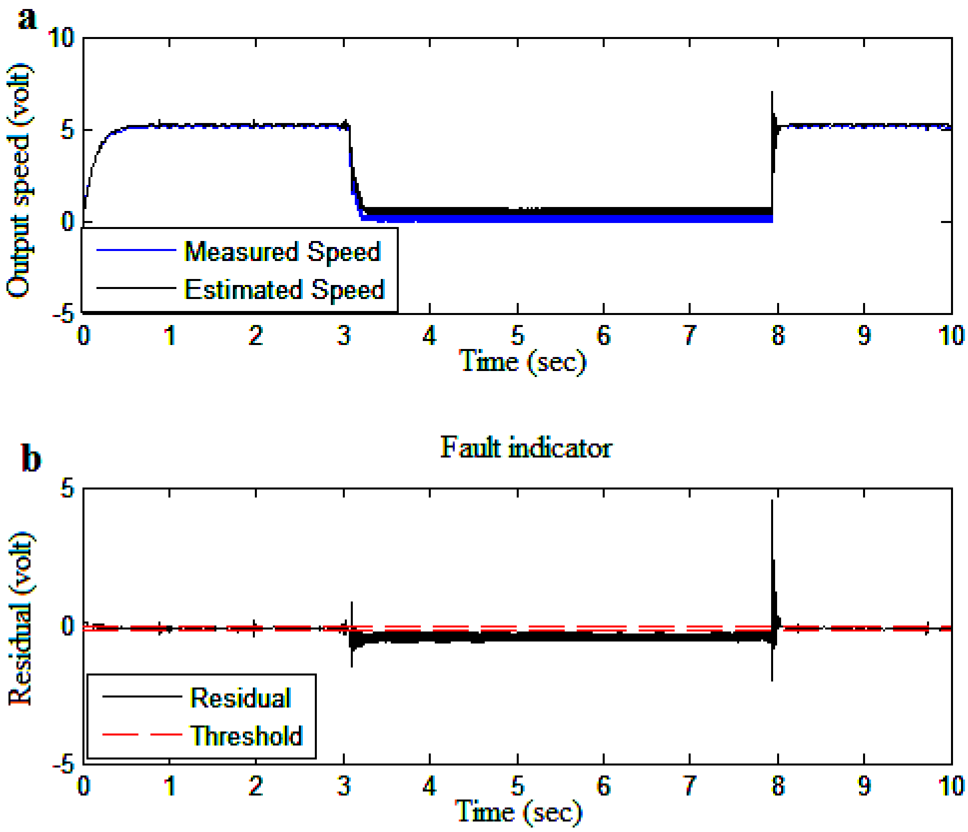


Figure 5.8. Disconnection Fault Scenario 1 a) Measured and Estimated Outputs, b) Residual.

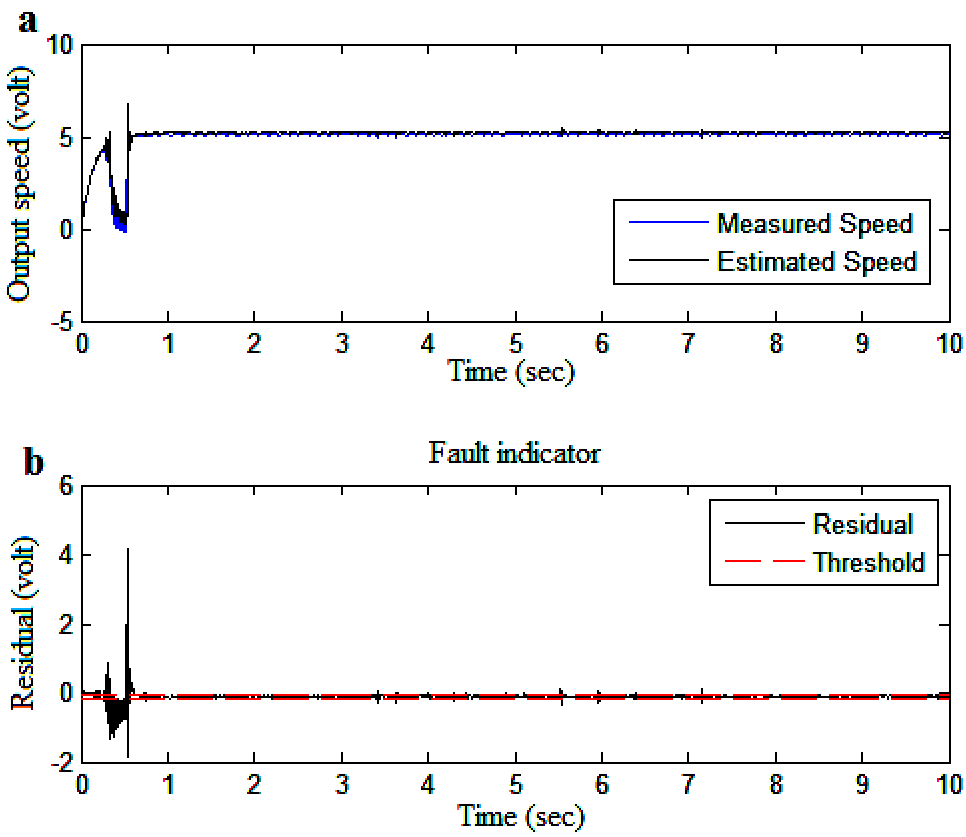


Figure 5.9. Disconnection Fault Scenario 2 a) Measured and Estimated Outputs, b) Residual.

The magnitude of residual shown in Figure 5.8(b) and 5.9(b) decreases under the threshold value at applied fault time. The time scale of Figure 5.8 and Figure 5.9 are expanded around the region of occurrence of the fault for clarity and these are given in Figure 5.10 and Figure 5.11, respectively. From Figure 5.10 and 5.11 it can be observed that the generated residual signals are sensitive to the faults under consideration.

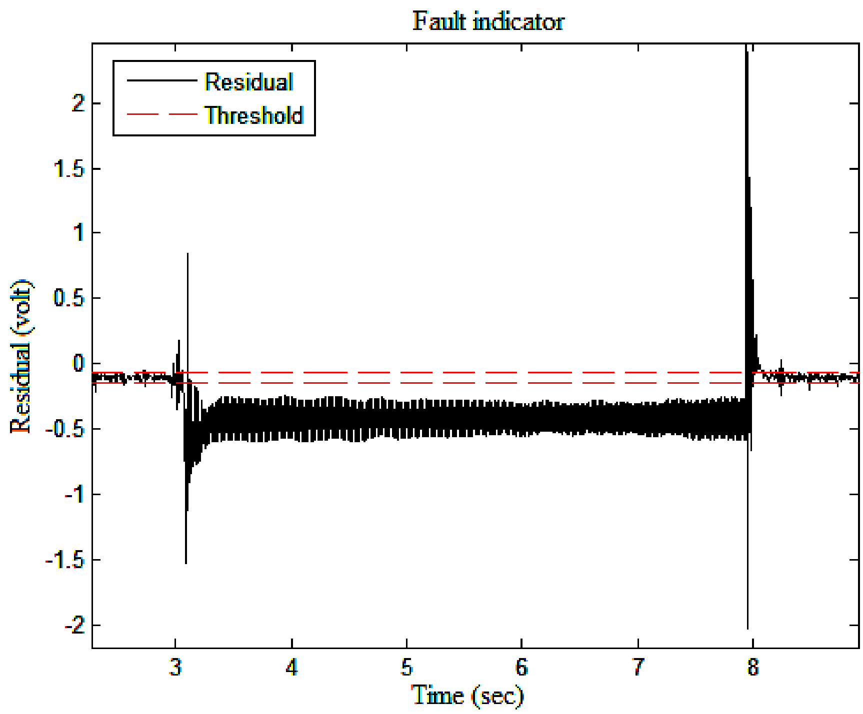


Figure 5.10. Disconnection Fault Scenario 1 Residual.

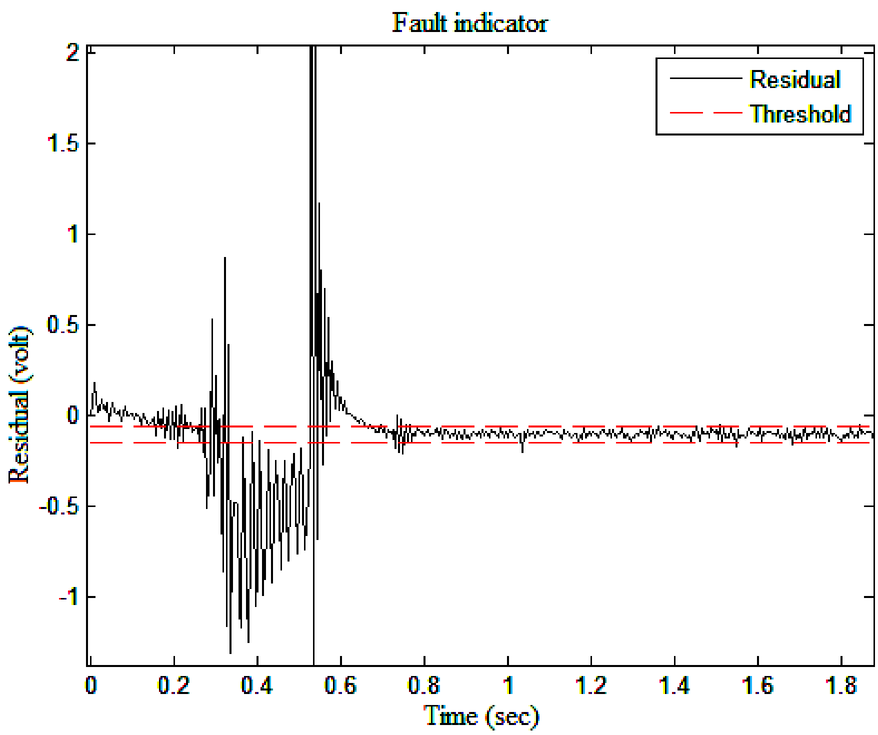


Figure 5.11. Disconnection Fault Scenario 2 Residual.

5.2. Data-Driven Based FD Applications

5.2.1. Variance Sensitive Adaptive Threshold Based PCA Method

In order to demonstrate the efficacy of the proposed PCA method, some experimental tests are performed. The experiments are carried out in open–loop and closed–loop conditions to test both the steady-state and transient operating conditions for the actuator and sensor faults under the constant load. The results of experimental tests using the fixed threshold-based Q-statistic (Q_α), T^2-statistic (T_α), combination of fixed and adaptive threshold-based T^2-statistic (T_{comb}) and proposed variance sensitive adaptive threshold-based (T_{vsa}) methods are presented for PCA monitoring. The alarm signals are computed for each method.

A functional diagram of the overall PCA fault detection method is demonstrated in Figure 5.12. Two measurements are chosen for calculating the PCA detection algorithm: the motor shaft speed and the motor armature current Using the Cumulative Percent Variance (CPV) approach, one principal component (a = 1) is found adequate to capture major correlations (%98) in the process variable. Q-statistic and T^2-statistic are computed in order to monitor behavior of the process. Then, the fixed threshold T_α and Q_α are calculated for %95 confidence limit. The combination of fixed and adaptive threshold T_{comb} and the proposed variance sensitive adaptive threshold T_{vsa} are calculated. For the 97% confidence level (i.e. $\alpha = 0.03$) (100(1-0.03) %=97%), and the coefficient z is calculated as 2.17.

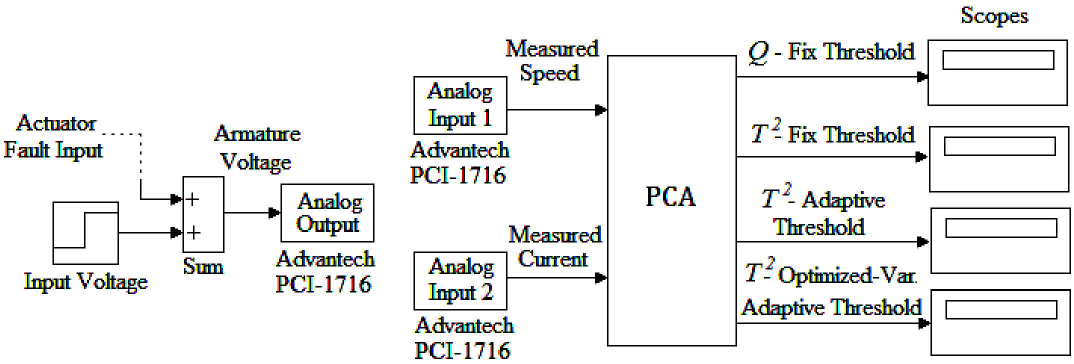


Figure 5.12. Simulink Model of the PCA Fault Detection Method.

5.2.1.1. Open-Loop Experiments

The open-loop diagram for the experiments is shown in Figure 5.13. Three fault experiments in the open-loop condition are carried out as actuator faults, sensor fault and transient tests.

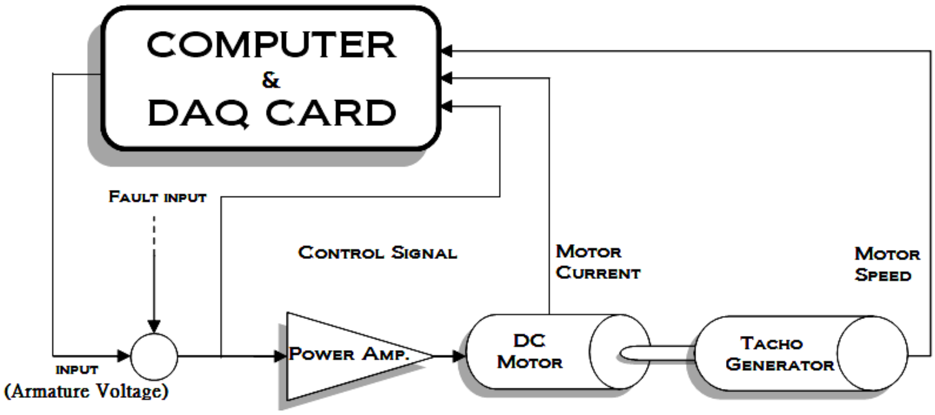


Figure 5.13. Block Diagram of the Open Loop System.

Experiment 1: Actuator Fault

Experimental application for the actuator fault is performed that 6 V of the input voltage (corresponds to 1700 rpm) is applied to the open loop dc motor system. The measured output speed and the armature current of the motor are shown in Figure 5.14 and Figure 5.15, respectively.

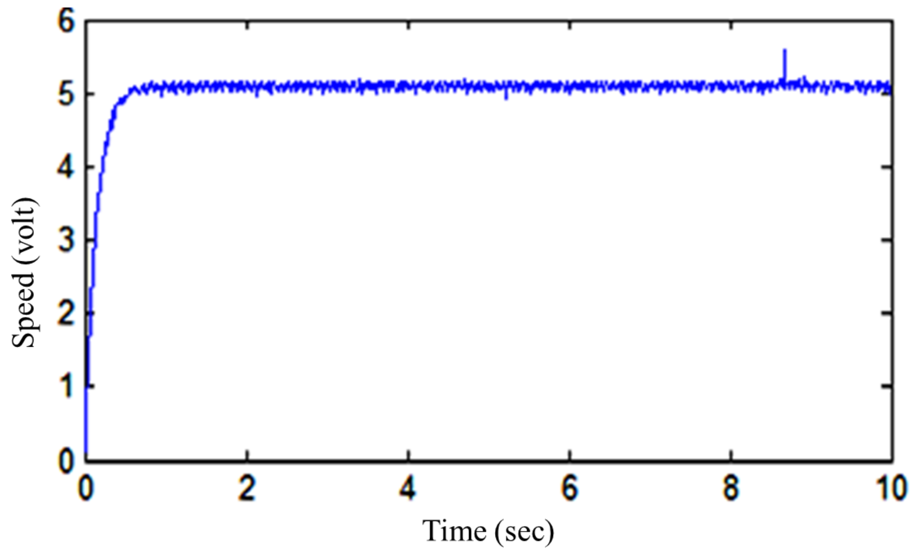


Figure 5.14. Fault-Free Measured Speed.

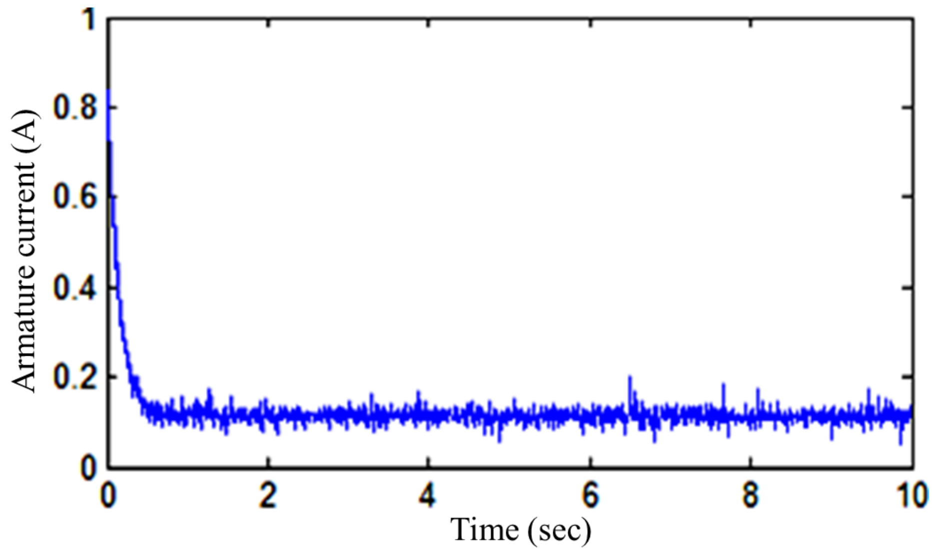


Figure 5.15. Fault-Free Measured Current.

A signal shown in Figure 5.16 is applied to the system as an actuator fault during the plant is running. After the actuator fault is applied, the measured speed and current signals are illustrated in Figure 5.17 and Figure 5.18, respectively.

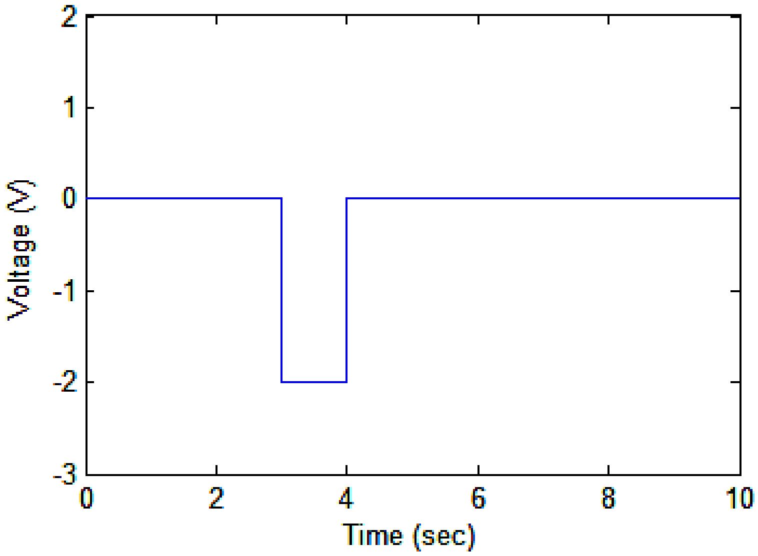


Figure 5.16. Applied Fault Signal.

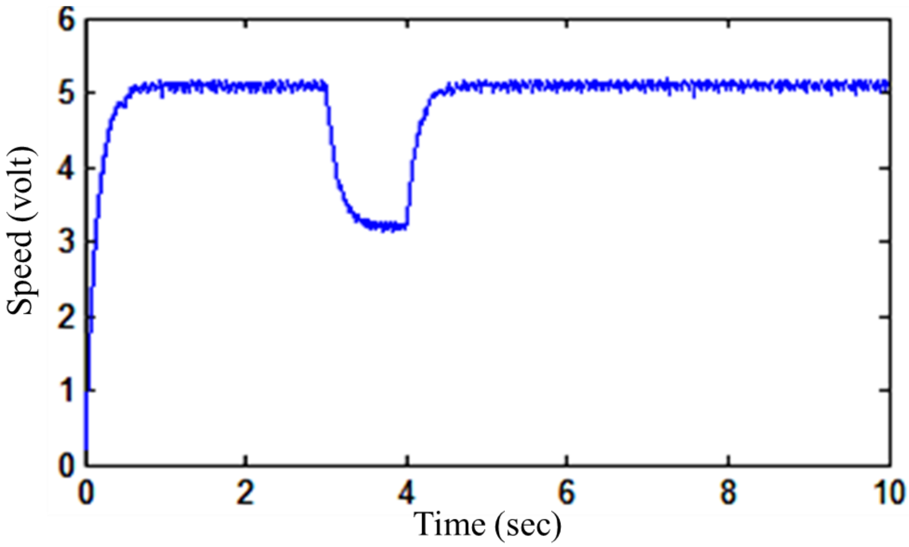


Figure 5.17. Measured Speed.

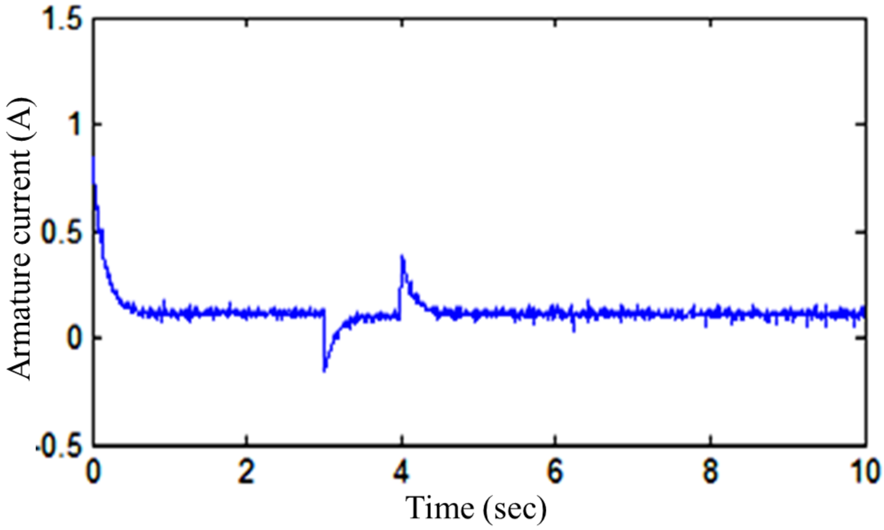


Figure 5.18. Measured Current.

The results are plotted including alarm signals in Figure 5.19 in which the left column plots show PCA method with their threshold values and the right column plots show alarm signals produced.

Q-statistic and T^2-statistic methods based on the fixed thresholds illustrated in Figure 5.19(a) and Figure 5.19(b) produce false alarm signals during the transient state of the fault signal applied. The missing fault signal components appear in the combination of fixed and adaptive threshold-based T^2-statistic method as shown in Figure 5.19(c). The alarm signal produced from the proposed variance sensitive adaptive threshold method (T_{vsa}) is illustrated in Figure 5.19(d) such that both the false alarm caused during the transient state and the missing data are eliminated.

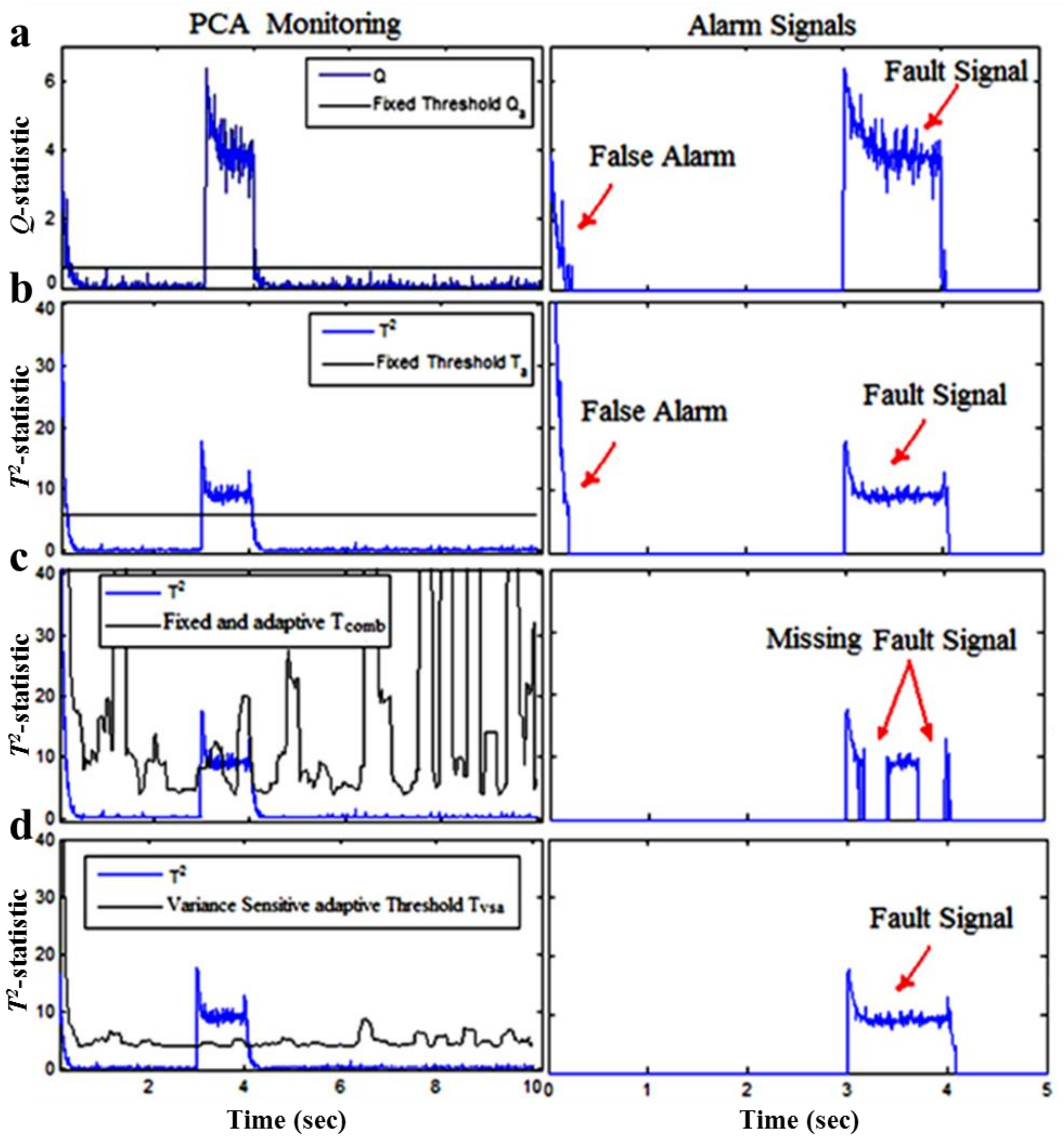


Figure 5.19. PCA Monitoring Charts with Alarm Signal a) Q_{α}, b) T_{α}, c) T_{comb}, and d) T_{vsa} .

Experiment 2: Sensor Fault

When the system is under the steady-state operation condition a fault is injected by disconnecting and reconnecting the speed sensor for duration of 1.6 seconds at 3 sec. After the applied fault, the measured speed is shown in Figure 5.20. Applied sensor fault does not efffect the input armatur current.

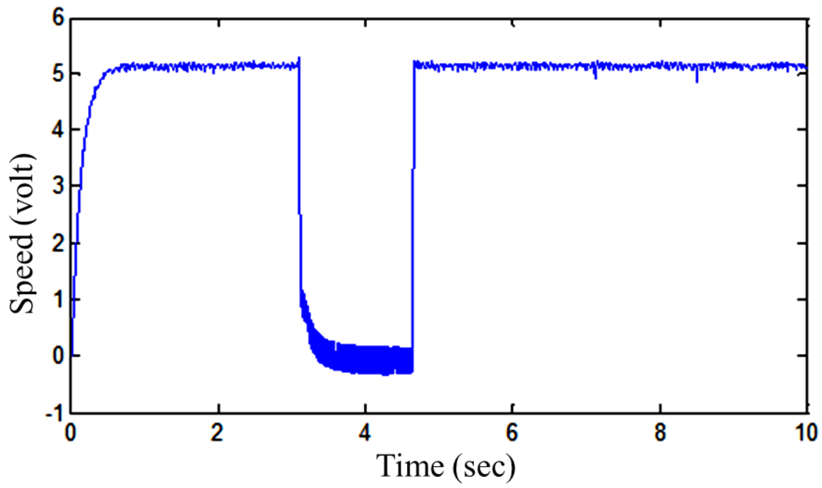


Figure 5.20. Measured Speed.

The results are plotted in Figure 5.21. The false alarm signal is produced in the Q-statistic and T^2-statistic methods as shown in Figure 5.21(a) and Figure 5.21(b), respectively. Interrupted fault signal (missing fault signal component) is produced in the combination of fixed and adaptive threshold-based T^2-statistic method in Figure 5.21(c). Proper fault signal without missing component is produced and the false alarm is eliminated in the proposed method as illustrated in Figure 5.21(d).

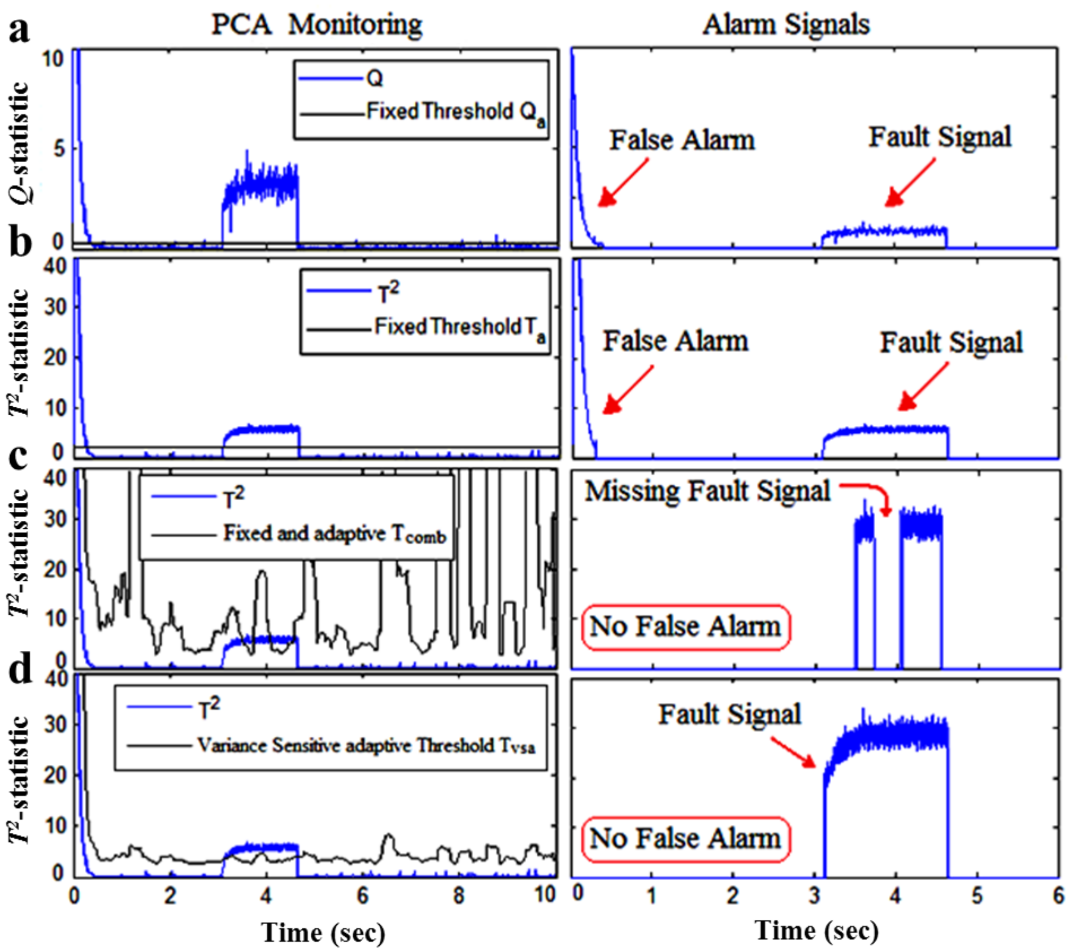


Figure 5.21. PCA Monitoring Charts with Alarm Signal a) Q_{α}, b) T_{α}, c) T_{comb}, and d) T_{vsa}.

Experiment 3: Transient Test (No Fault)

A square wave input armature voltage is applied to test the transient condition such that the motor is permitted to run between 420 rpm – 1440 rpm. When 2 volts of the armature input voltage is applied to the plant the output speed is measured to be 420 rpm. The input armature voltage is increased from 2 volts to 6 volts, and the output speed is measured to be 5.18 volts from the tachometer as shown in Figure 5.22 such that 5.18 volts of the output corresponds to 1440 rpm of speed. The measured speed and armature current of the motor are plotted in Figure 5.23 and Figure 5.24, respectively.

False alarm signals are produced in the Q-statistic and T^2-statistic methods as shown in Figure 5.25(a) and Figure 5.25(b). No alarm signal is produced in the combination of fixed and adaptive threshold-based T^2-statistic method and proposed variance sensitive adaptive method. Although no fault is applied to the system, the transient states are evaluated as a fault in the usual Q-statistic and T^2-statistic methods. This confirms the fact that usual Q-statistic and T^2-statistic fault detection methods are not applicable to the systems with variable set-points. On the other hand, the proposal method prevents the false alarm as shown in Figure 5.25(c) and Figure 5.25(d).

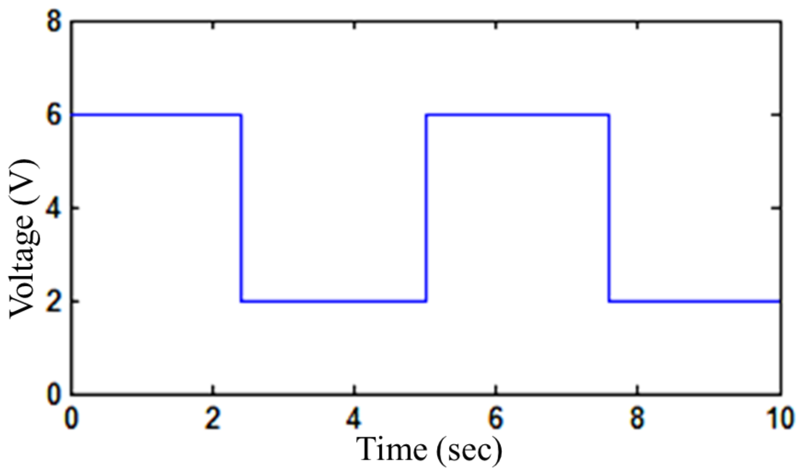


Figure 5.22. Applied Input Voltage.

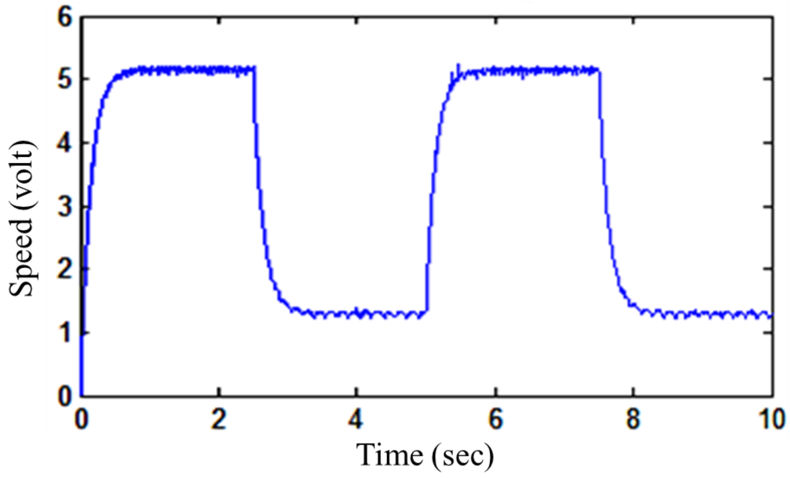


Figure 5.23. Measured Speed.

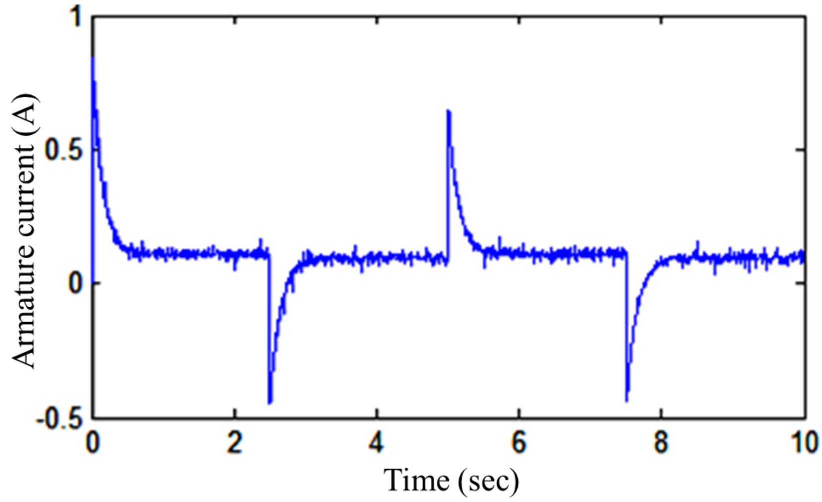


Figure 5.24. Measured Current.

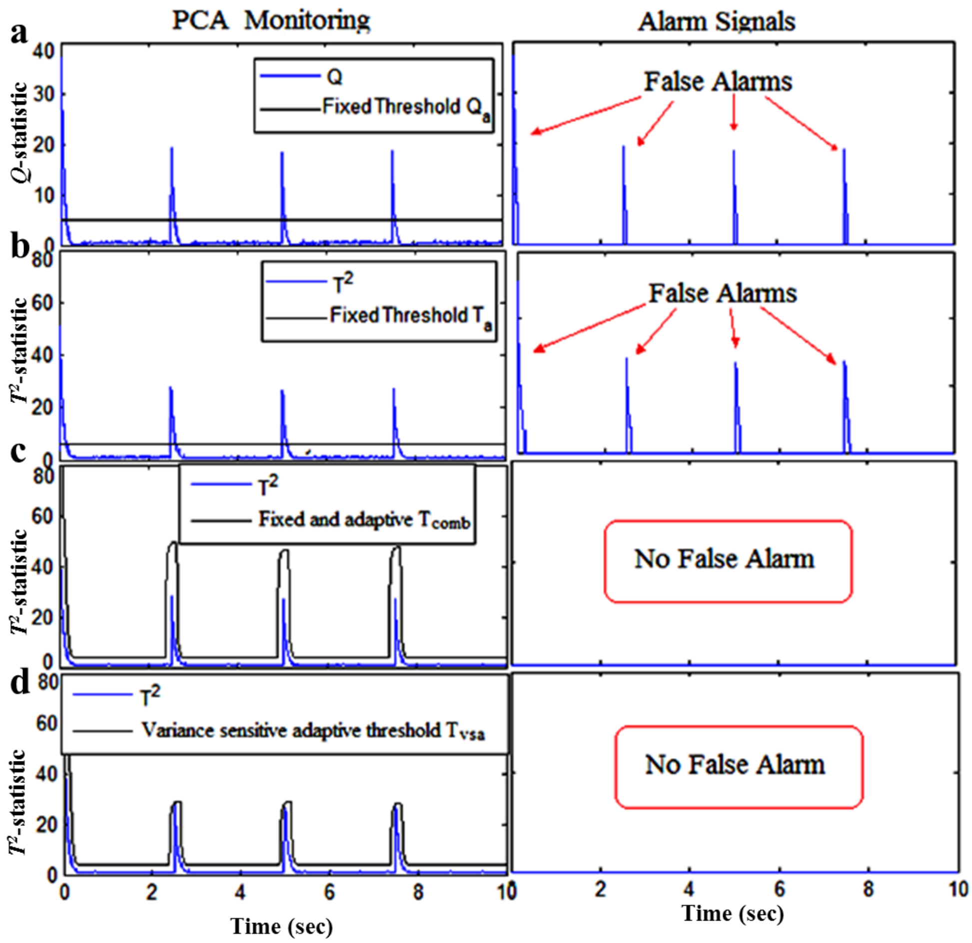


Figure 5.25. PCA Monitoring Charts with Alarm Signal a) Q_α, b) T_α, c) T_{comb}, and d) T_{vsa}.

5.2.1.2. Closed-Loop Experiments

Closed-loop experimental tests are performed using a Proportional and Integral (PI) control structure illustrated in Figure 5.26 to test the system for the actuator and sensor faults.

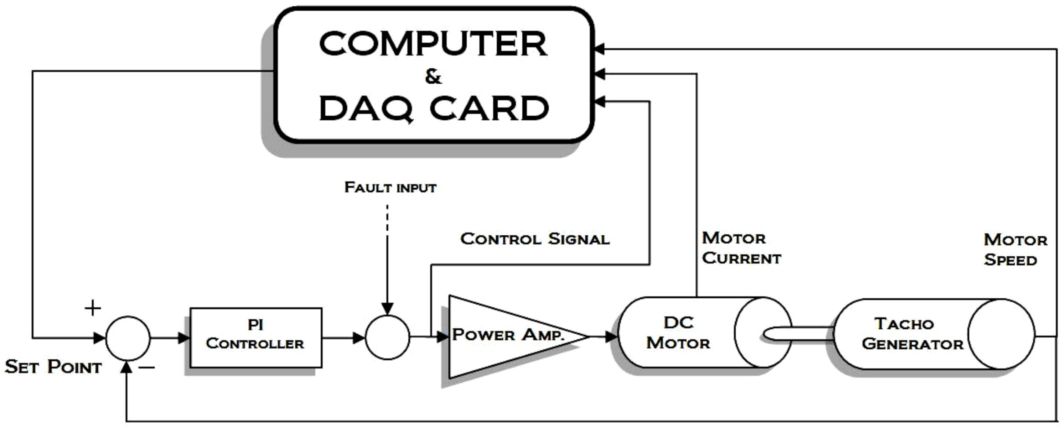


Figure 5.26. Experimental Closed-Loop Diagram.

Experiment 4: Actuator Fault

The experimental test is performed that 4 V of the set-point voltage (corresponding to 1130 rpm) is applied to the closed loop system. Actuator fault is applied to the system while it is running in the steady-state operation such that the same fault signal performed in experiment 1, Figure 5.16, is applied to test response of the closed-loop system. The measured speed and current signals are illustrated in Figure 5.27 and Figure 5.28, respectively.

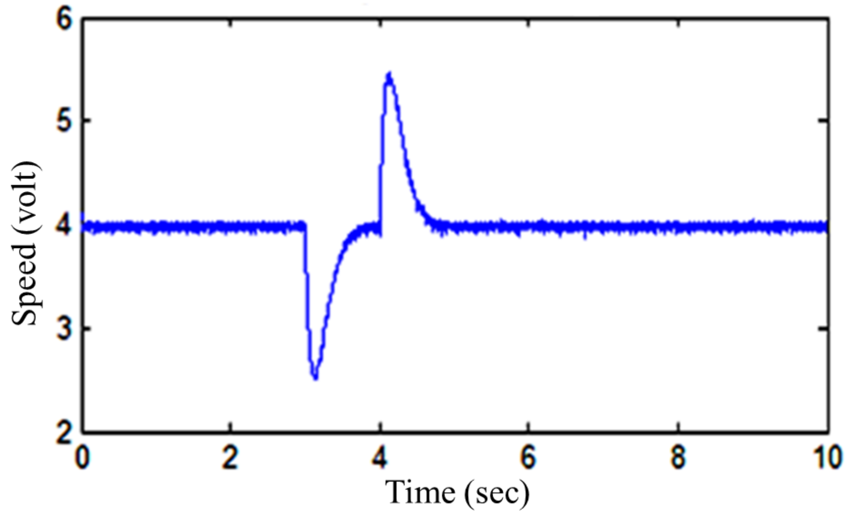


Figure 5.27. Measured Speed.

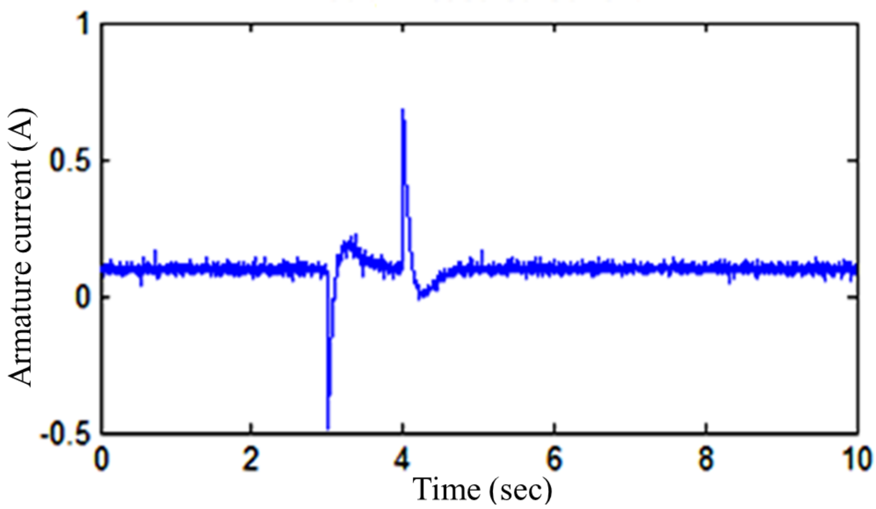


Figure 5.28. Measured Current.

Fault alarm signals are indicated in all methods as shown in Figure 5.29 accounted without any fault alarm since the system operates in the steady-state condition. However a missing fault signal is observed with the combination of fixed and adaptive threshold based T^2-statistic method that is illustrated in Figure 5.29(c).

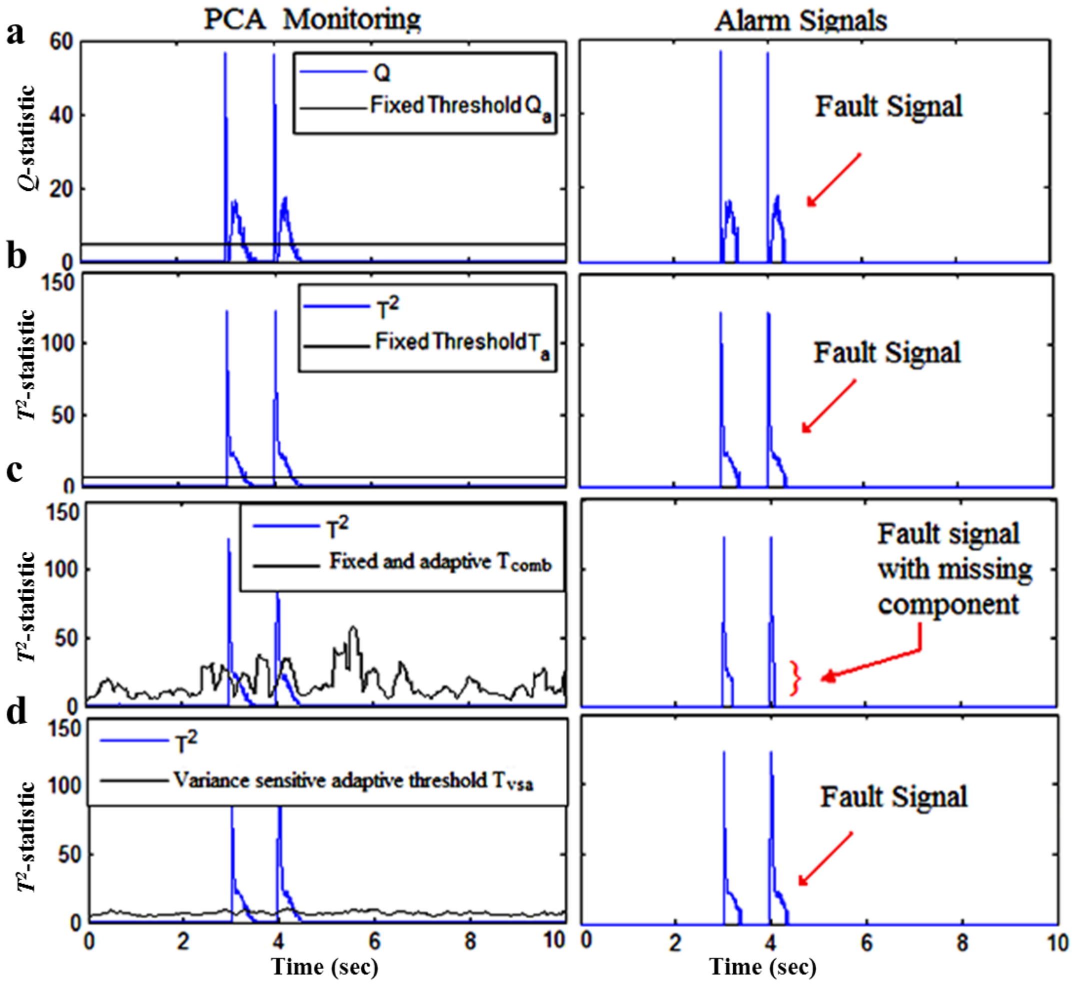


Figure 5.29. PCA Monitoring Charts with Alarm Signal a) Q_α, b) T_α, c) T_{comb}, and d) T_{vsa}.

Experiment 5: Sensor Fault

In the current experiment, the fault is applied to the speed sensor. When the system is under the steady-state operation condition the fault is injected by disconnecting and reconnecting the speed sensor for duration of 0.5 seconds at 3.5 sec after the start-up. After the applied sensor fault, the measured speed is shown in Figure 5.30 and the measured current is illustrated in Figure 5.31.

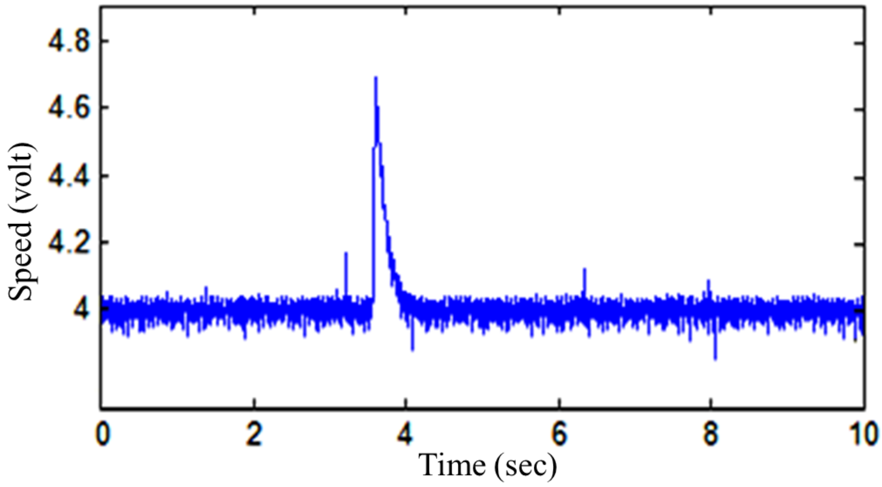


Figure 5.30. Measured Speed.

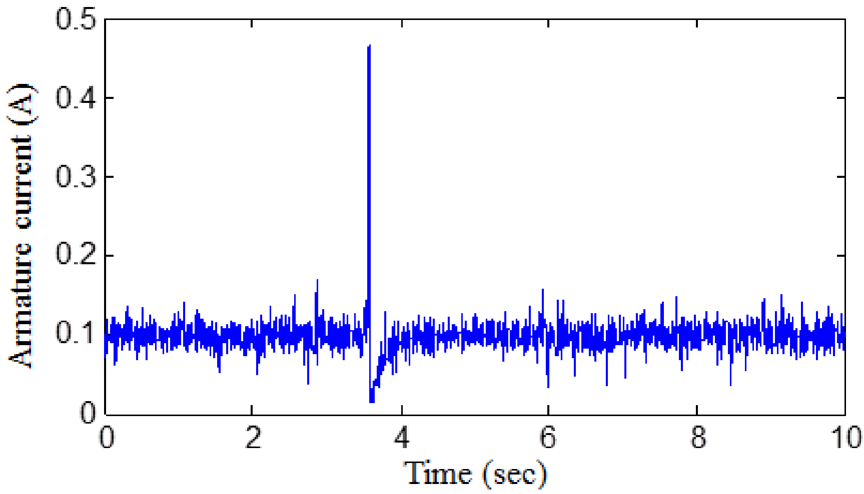


Figure 5.31. Measured Current.

The results are plotted in Figure 5.32. The alarm signals are observed in all methods without any fault alarm signal since the system operates in the steady-state condition. However a missing fault signal is observed with the combination of fixed and adaptive threshold-based T^2-statistic method.

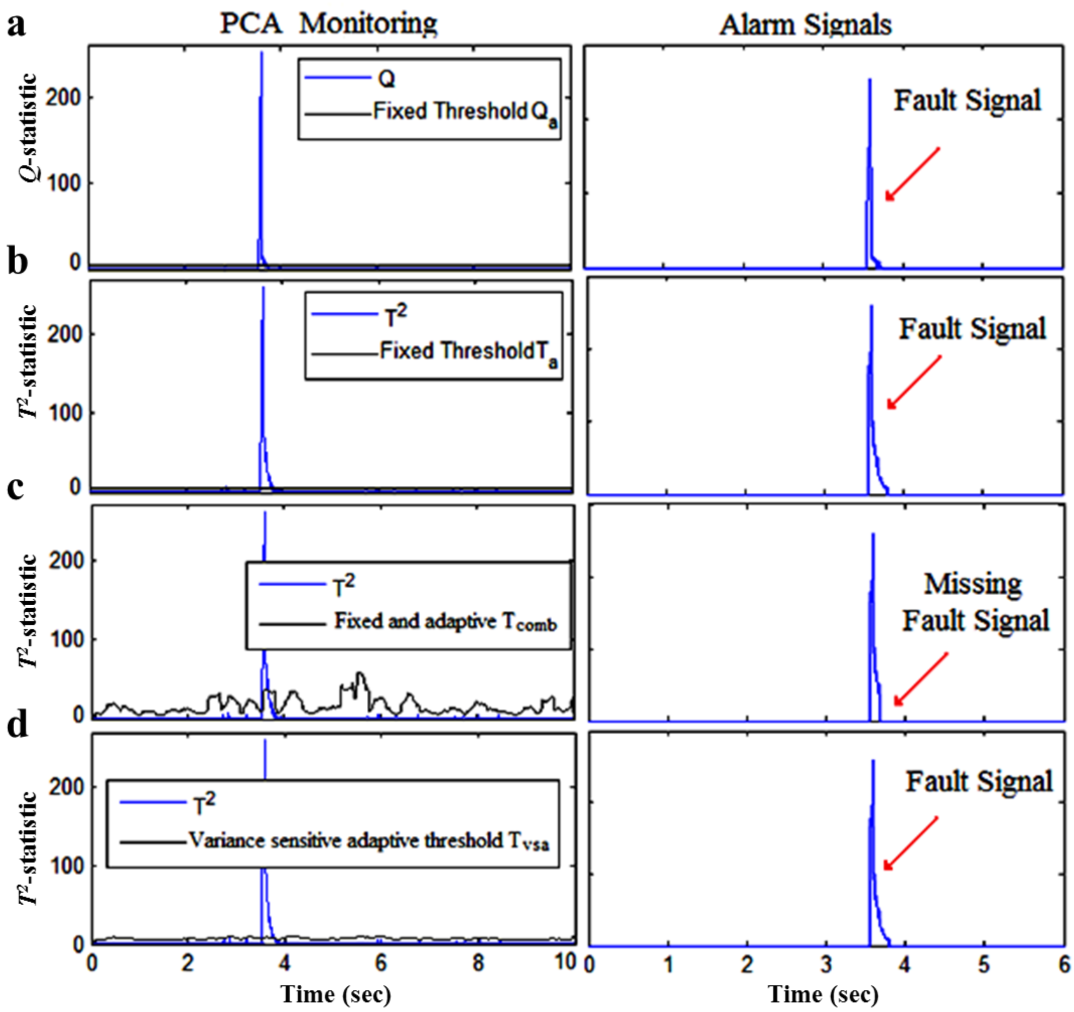


Figure 5.32. PCA Monitoring Charts with Alarm Signal a) Q_{α} , b) T_{α} , c) T_{comb} , and d) T_{vsa} .

Experiment 6: Transient Test (Sensor Fault)

A variable set-point signal shown in Figure 5.22 is applied to the closed loop system. The measured speed and the armature current of the motor are shown in Figure 5.33 and Figure 5.34, respectively.

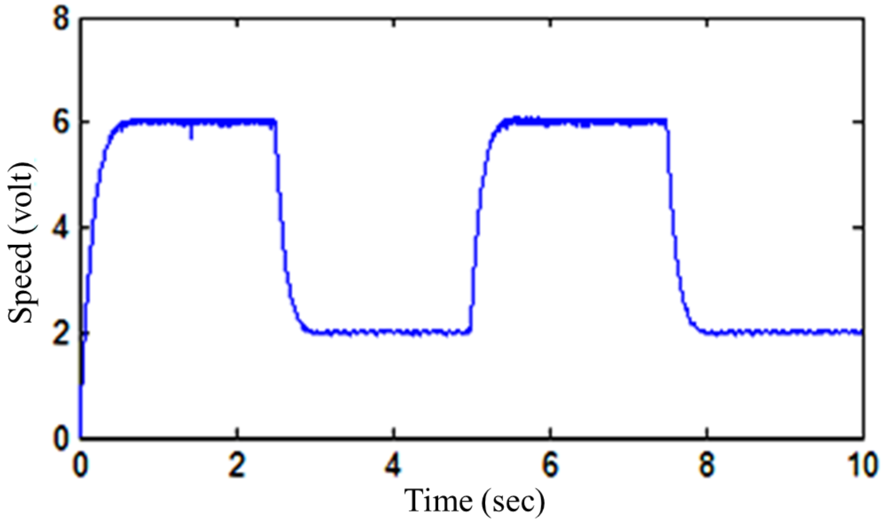


Figure 5.33. Measured Speed.

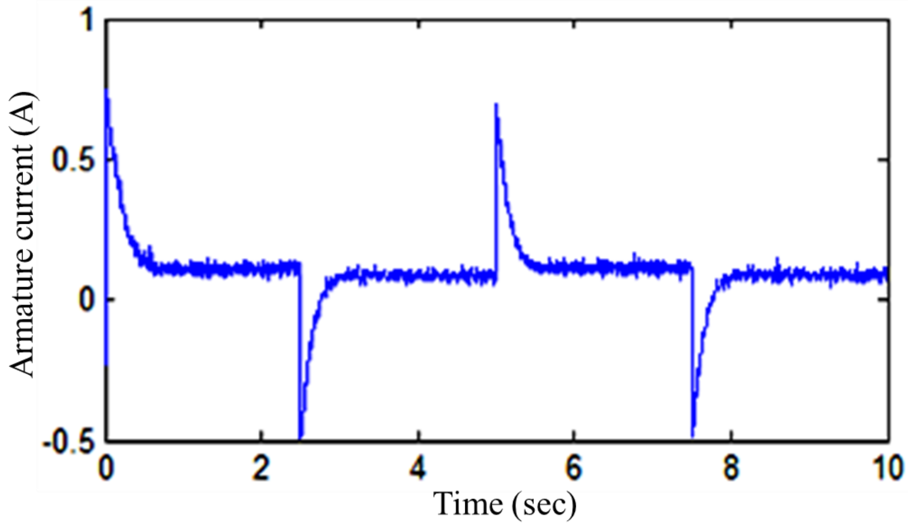


Figure 5.34. Measured Current.

While the system operates under the servo (tracking) condition as shown in Figure 5.33, a fault is injected by disconnecting and reconnecting the speed sensor for duration of 0.7 seconds at 1.75 sec. After the sensor fault is applied, the measured output speed and the measured current are plotted in Figure 5.35 and in Figure 5.36.

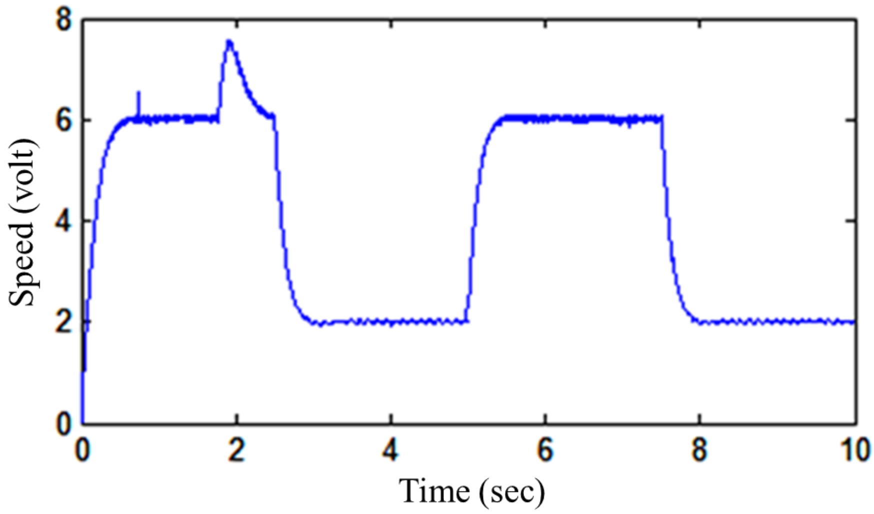


Figure 5.35. Measured Speed After the Sensor Fault.

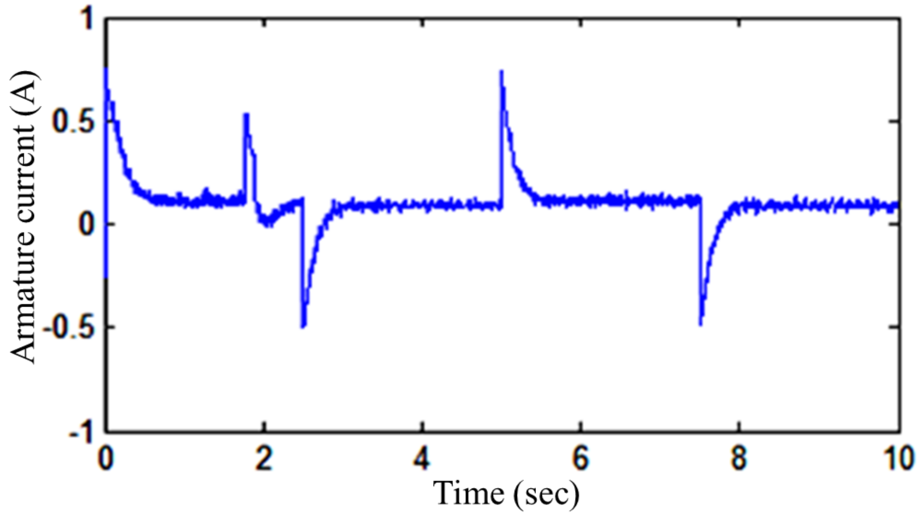


Figure 5.36. Measured Current After the Sensor Fault.

False alarms are produced in the Q-statistic (Figure 5.37(a)) and T^2-statistic (Figure 5.37(b)) method. False alarm is not indicated in the combination of fixed and adaptive threshold-based T^2-statistic method but the fault signal appears with the missing component. False alarm is not present in the proposed variance sensitive adaptive threshold-based method and the produced fault signal does not include any missing component.

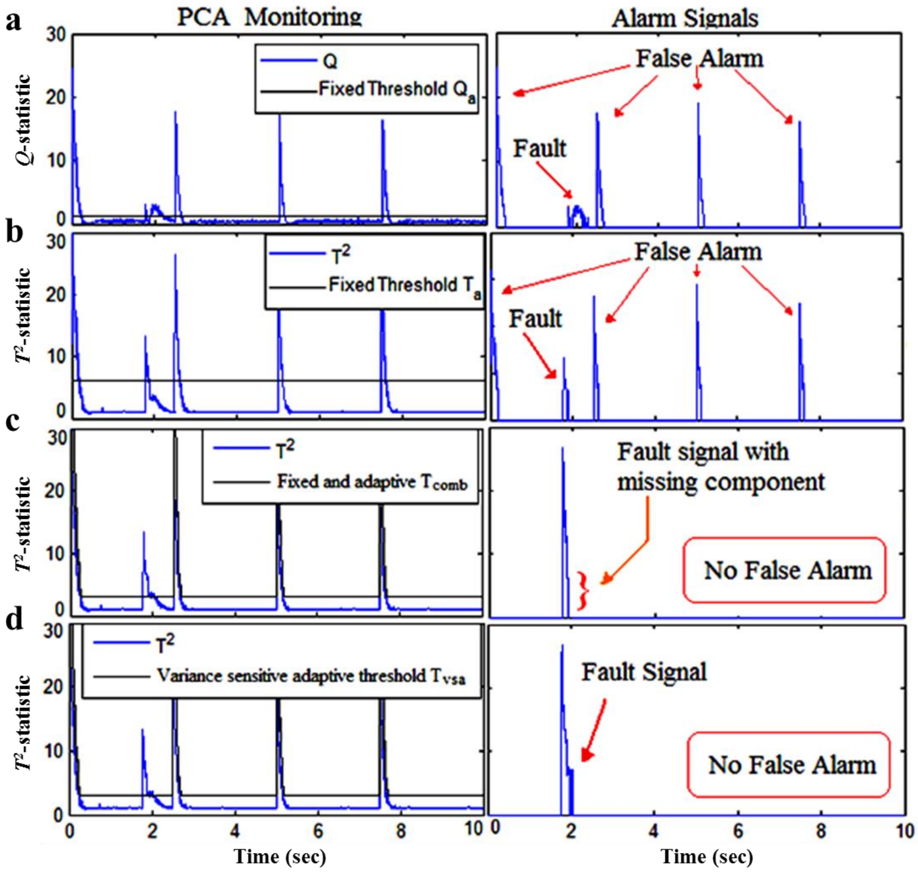


Figure 5.37. PCA Monitoring Charts with Alarm Signal a) Q_α , b) T_α , c) T_{comb} , and d) T_{vsa} .

Experiment 7: Actuator and Sensor Faults

Experimental application in which the actuator and sensor faults occur simultaneous is performed. While the plant runs in open-loop conditions with the applied of 6 V the input armature voltage, the output speed is 1440 rpm corresponding to 5.18 volts of the measured voltage from the tachometer. After 1.6 seconds of starting, as an actuator fault the input armature voltage cable is disconnected manually, and the sensor cable is also disconnected manually at 1.8 seconds. The sensor cable is re-connected at 2.0 seconds manually and the actuator cable is re-connected at 2.35 seconds. The measured output and the armature current are illustrated in Figure 5.38 and Figure 5.39, respectively.

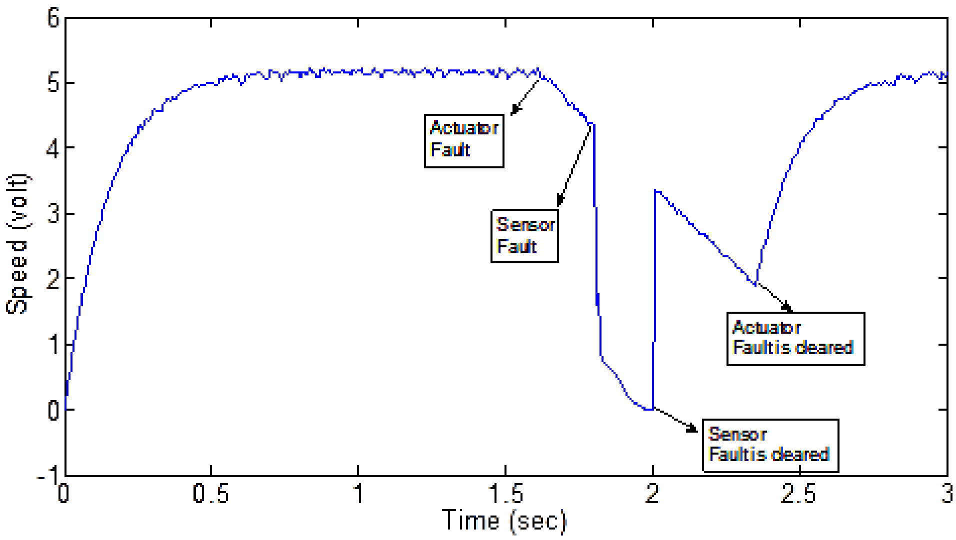


Figure 5.38. Measured Output Speed.

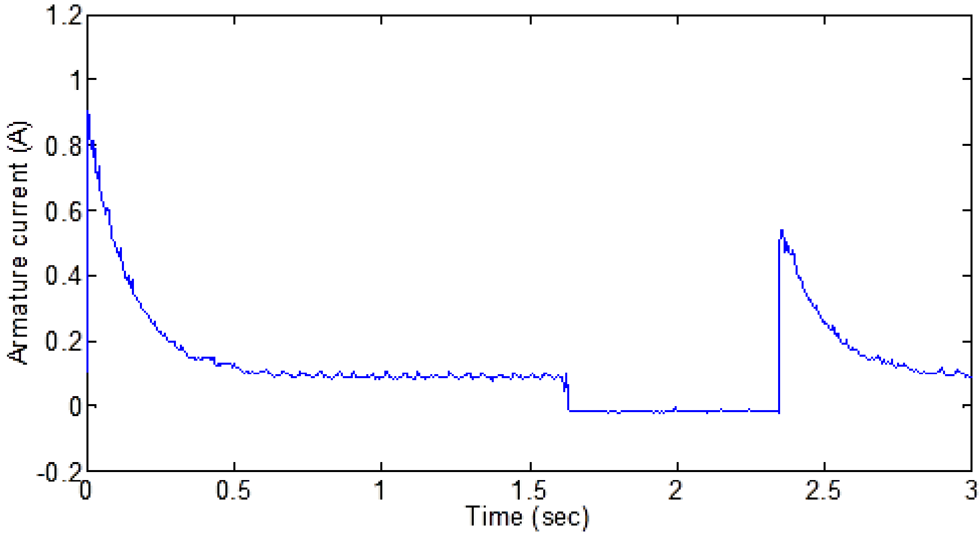


Figure 5.39. Measured Armature Current.

The results are plotted with their alarm signals in Figure 5.40. *Q*-statistic and T^2-statistic methods based on the fixed thresholds illustrated in Figure 5.40(a) and Figure 5.40(b) produce false alarm signals during the transient state of the fault signal applied. The missing fault signal components appear in the combination of fixed and adaptive threshold-based T^2-statistic method as shown in Figure 5.40(c). The alarm signal produced from the proposed variance sensitive adaptive threshold method (T_{vsa}) is illustrated in Figure 5.40(d) such that both the false alarm caused during the transient state and the missing data are eliminated.

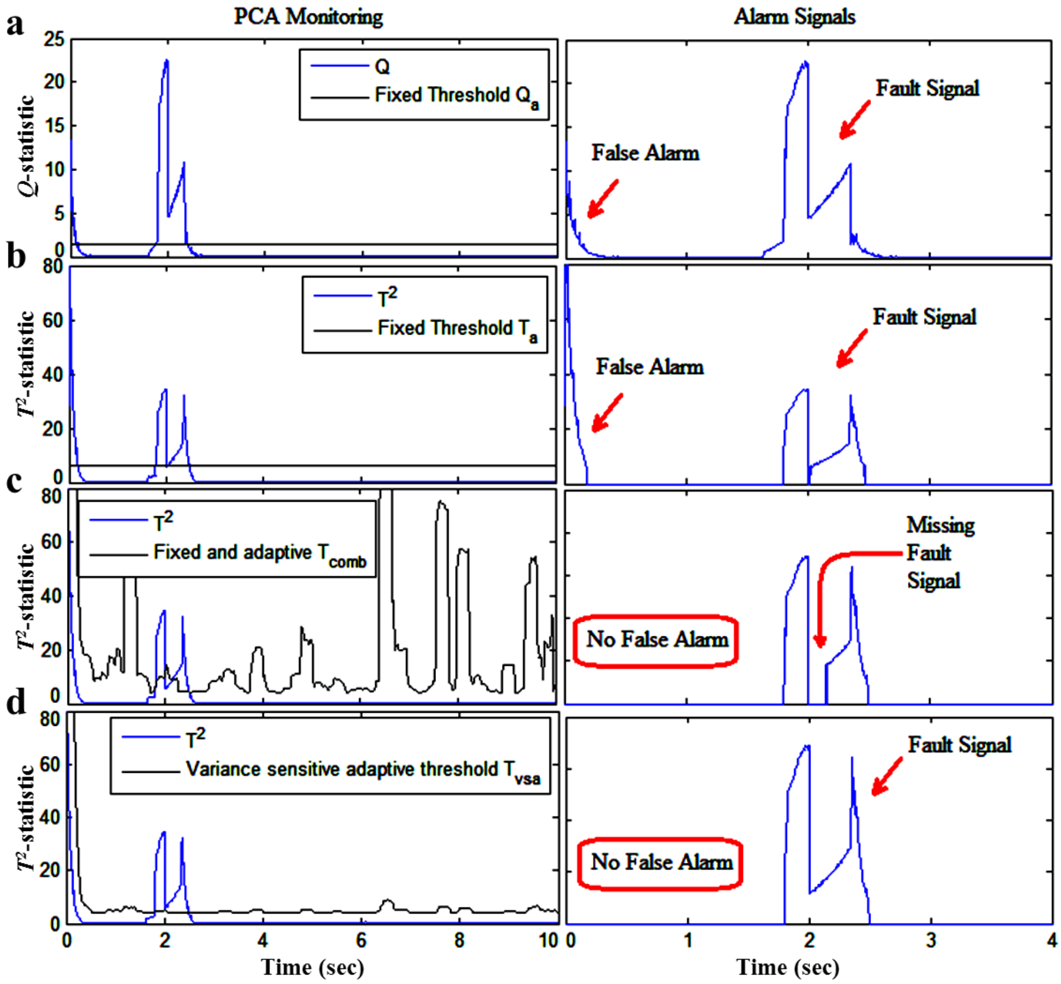


Figure 5.40. PCA Monitoring Charts for Simultaneous Actuator and Sensor Faults with Alarm Signal a) Q_α, b) T_α, c) T_{comb}, and d) T_{vsa}.

The overall experimental results based on the test performed in the open-loop and closed-loop operating conditions are summarized in Table 5.1 for a complete comparison. The fixed threshold-based Q-statistic and T^2-statistic methods produce false alarms during the variable input in the open-loop or variable set-point signals.

Fault signal with the missing components is obtained with the combination of fixed and adaptive threshold-based T^2-statistic method. The false alarm and missing alarm signal problems are eliminated in the proposed variance sensitive adaptive threshold-based T^2-statistic method.

Table 5.1. Comparison Results Between the Standard and Proposed Thresholds.

			Thresholds for PCA Monitoring			
			Q_a	T^2_a	T^2_{comb}	T^2_{vsa}
Experimentals Tests	Open-loop Experiment	Actuator Fault	✗ False Alarm ✓ No Missing Data	✗ False Alarm ✓ No Missing Data	✓ No False Alarm ✗ Missing Data	✓ No False Alarm ✓ No Missing Data
		Sensor Fault	✗ False Alarm ✓ No Missing Data	✗ False Alarm ✓ No Missing Data	✓ No False Alarm ✗ Missing Data	✓ No False Alarm ✓ No Missing Data
		Transient test (No Fault)	✗ False Alarm	✗ False Alarm	✓ No False Alarm	✓ No False Alarm
		Actuator and Sensor fault	✗ False Alarm ✓ No Missing Data	✗ False Alarm ✓ No Missing Data	✓ No False Alarm ✗ Missing Data	✓ No False Alarm ✓ No Missing Data
	Closed-loop Experiment	Actuator Fault (steady state)	✓ No False Alarm ✓ No Missing Data	✓ No False Alarm ✓ No Missing Data	✓ No False Alarm ✗ Missing Data	✓ No False Alarm ✓ No Missing Data
		Sensor Fault (steady state)	✓ No False Alarm ✓ No Missing Data	✓ No False Alarm ✓ No Missing Data	✓ No False Alarm ✗ Missing Data	✓ No False Alarm ✓ No Missing Data
		Transient test Sensor Faults	✗ False Alarm ✓ No Missing Data	✗ False Alarm ✓ No Missing Data	✓ No False Alarm ✗ Missing Data	✓ No False Alarm ✓ No Missing Data

✓ Desired result ✗ Undesired result

5.2.2. Wavelet Based Combined PCA Method

In order to validate the feasibility and reliability of the proposed combined PCA method using wavelet, open loop and closed-loop experimental tests are performed on an electromechanical system (PI controller) and process control system (PID controller).

The results of the experimental tests obtained from fixed threshold-based T^2-statistic (T_α), combination of fixed and adaptive threshold-based T^2-statistic (T_{comb}), and combination of fixed and adaptive threshold-based T^2-statistic (T_{comb}) with wavelet based de-noised signal are presented for PCA monitoring. The alarm signals produced for each method are computed. The sampling period is taken to be 5 ms for all experimental tests.

The measured noisy data are de-noised by using discrete wavelet transform. The fourth order of Daubechies wavelet (Db4) having four decomposition levels (level 4) and soft thresholding is applied to minimize the negative influence of the noise. After the wavelet de-noising, the combined PCA method is performed to detect the faults in the system. The flow chart of the process of the proposed method is demonstrated in Figure 5.41.

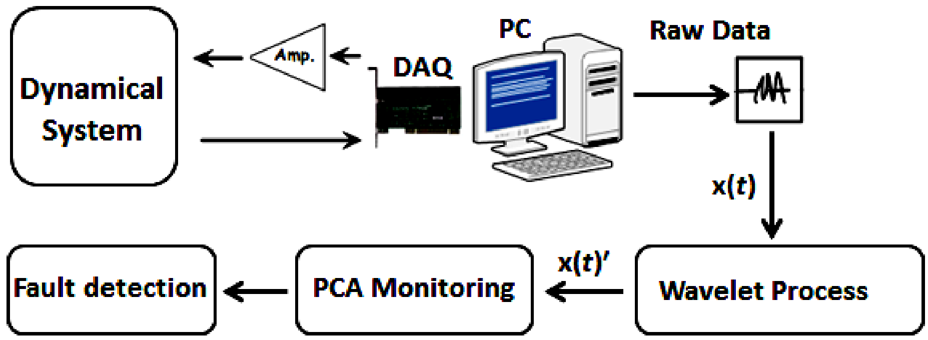


Figure 5.41. Flow Chart of the Proposed Method.

5.2.2.1. Experiments on Electromechanical System

5.2.2.1. (a) Open-Loop Tests

Experiment 1: Actuator Fault

Experimental application for the actuator fault is performed that 6 V of the input voltage (corresponds to 1700 rpm) is applied to the armature of the open-loop dc motor systems. Original noisy and de-noised signals of the measured output speed and the armature current of the motor are shown in Figure 5.42 and Figure 5.43, respectively.

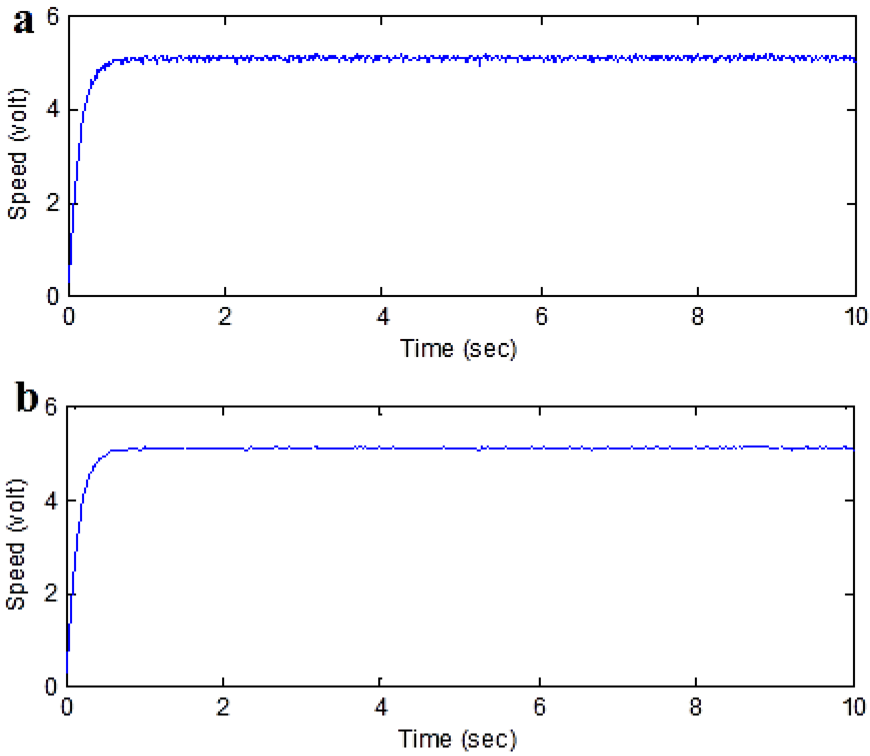


Figure 5.42. Fault-Free Measured Speed a) Noisy, b) De-noised by Wavelet.

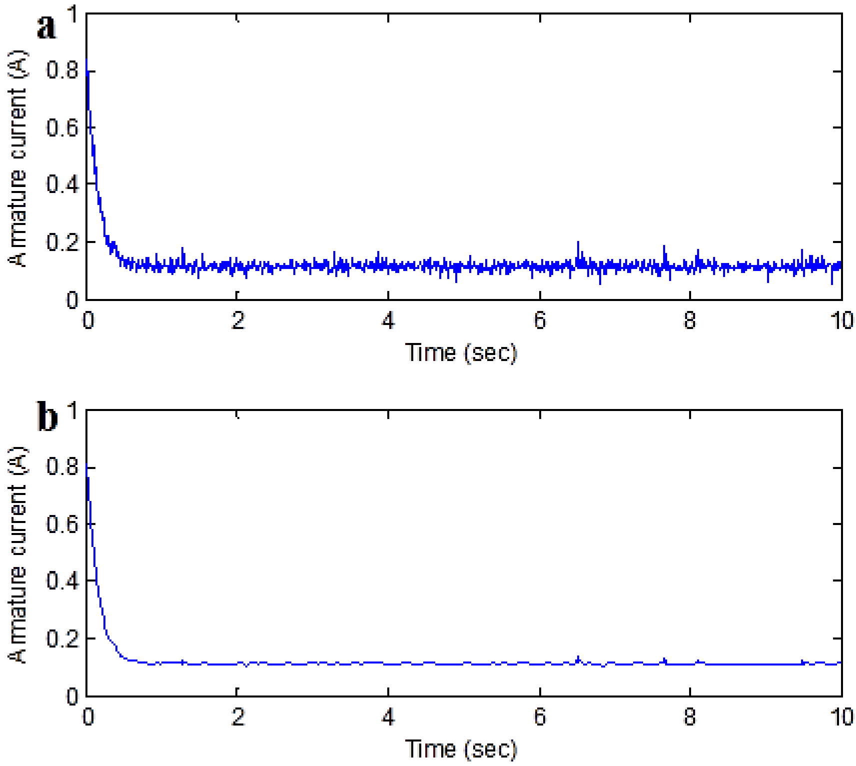


Figure 5.43. Fault-Free Measured Current a) Noisy, b) De-noised by Wavelet.

The signal shown in Figure 5.44 is applied to the system as an actuator fault during the plant is running. After the actuator fault is applied, the measured speed and current signals are illustrated in Figure 5.45 and Figure 5.46, respectively.

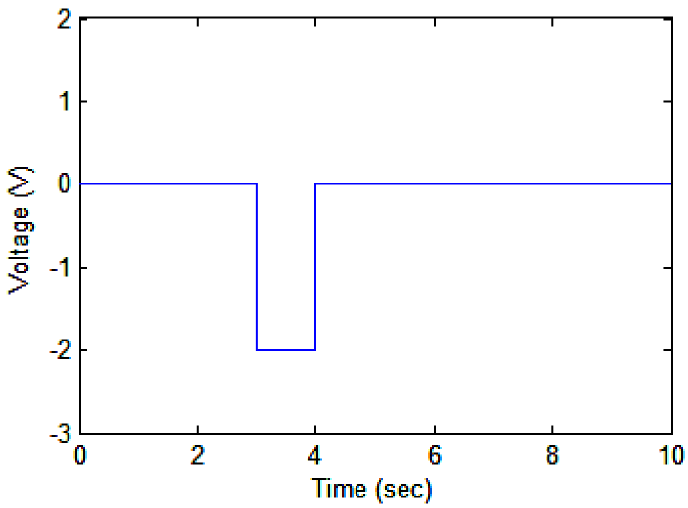


Figure 5.44. Applied Fault Signal.

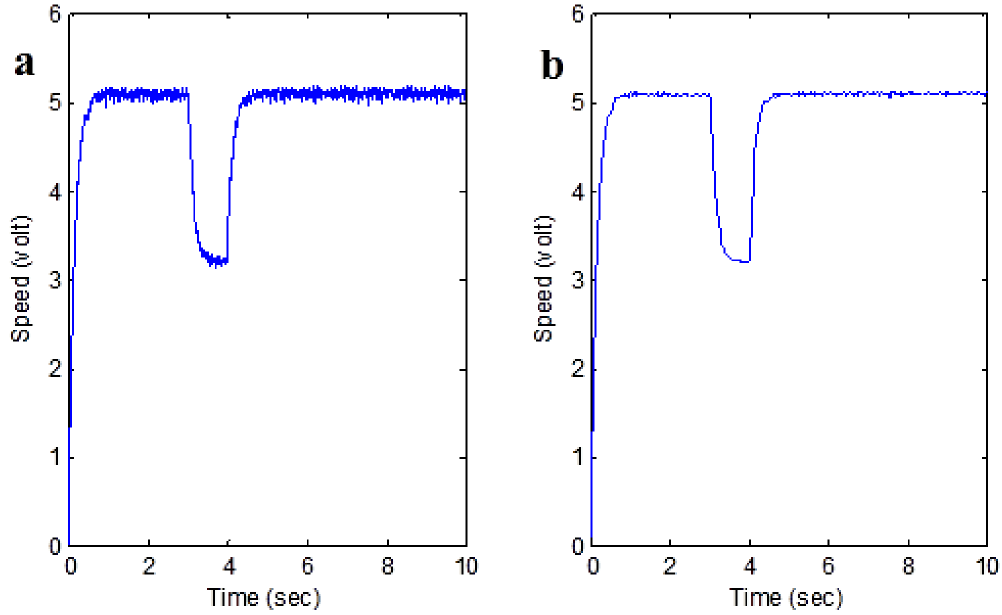


Figure 5.45. Measured Speed a) Noisy, b) De-noised by Wavelet.

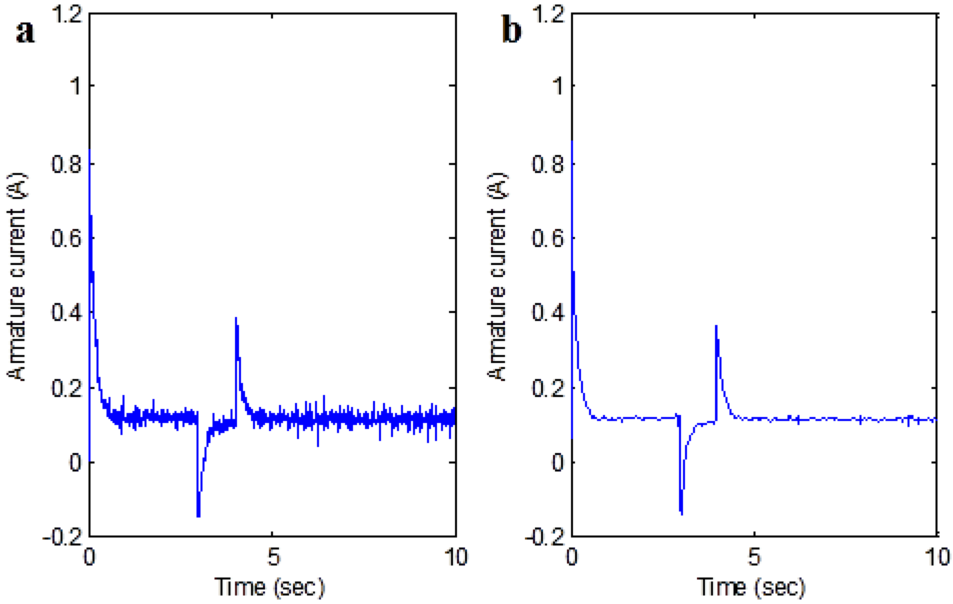


Figure 5.46. Measured Current a) Noisy, b) De-noised by Wavelet.

The results are plotted including the alarm signals in Figure 5.47 in which the left column plots show PCA method with their threshold values based on the noisy/de-noised signal and the right column plots show alarm signals produced.

T^2-statistic methods based on the fixed thresholds illustrated in Figure 5.47 (a) produces false alarm signal during the transient state of the fault signal applied. The missing fault signal components appear in the combination of fixed and adaptive threshold-based T^2-statistic method as shown in Figure 5.47(b). The alarm signal produced from the proposed method the combination of fixed and adaptive threshold-based T^2 with de-noised by wavelet, is illustrated in Figure 5.47(c) such that both the false alarm caused during the transient state and the missing data are eliminated.

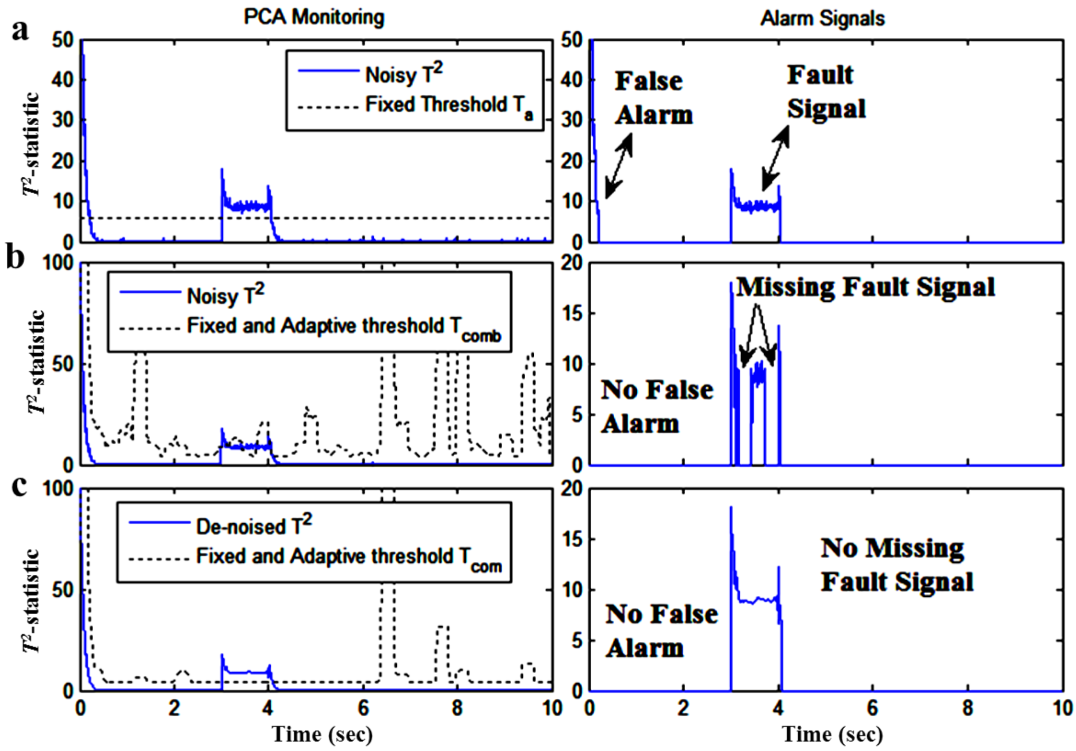


Figure 5.47. PCA Monitoring Charts with the Alarm Signals a) T_{α} , b) T_{comb} , and c) T_{comb} with Wavelet.

Experiment 2: Sensor Fault

When the system is under the steady-state operation condition a fault is injected by disconnecting and reconnecting the speed sensor for duration of 0.25 seconds at 1.1 sec. After the applied fault, noisy and de-noised measured speeds are shown in Figure 5.48. Applied sensor fault does not affect the input armatur current. Since the system runs in open-loop conditions.

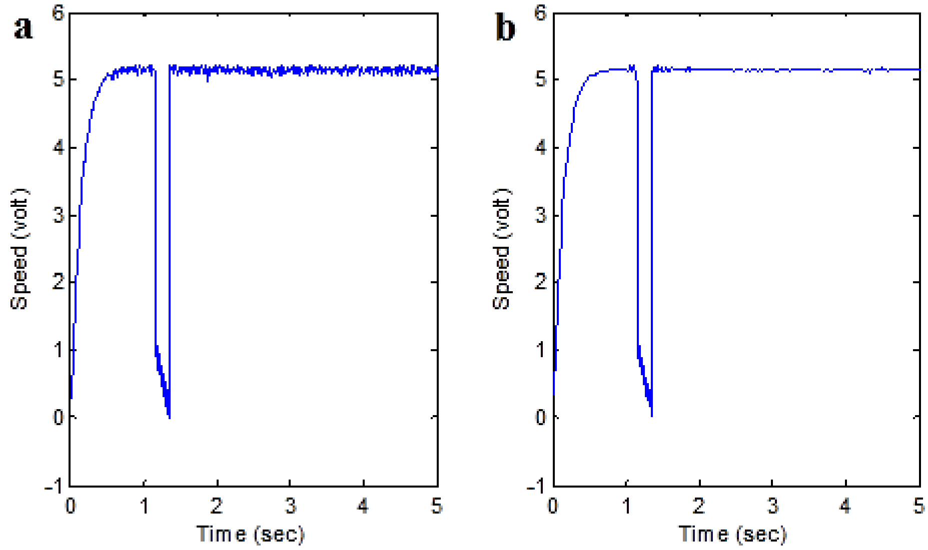


Figure 5.48. Measured Speed a) Noisy, b) De-noised by Wavelet.

The results are plotted including the alarm signals in Figure 5.49. The false alarm signal is produced in T^2-statistic methods as shown in Figure 5.49(a). Interrupted fault signal (missing fault signal component) is produced in the combination of fixed and adaptive threshold-based T^2-statistic method as illustrated in Figure 5.49(b). Proper fault signal without missing component is produced and the false alarm is eliminated in the proposed method as illustrated in Figure 5.49(c).

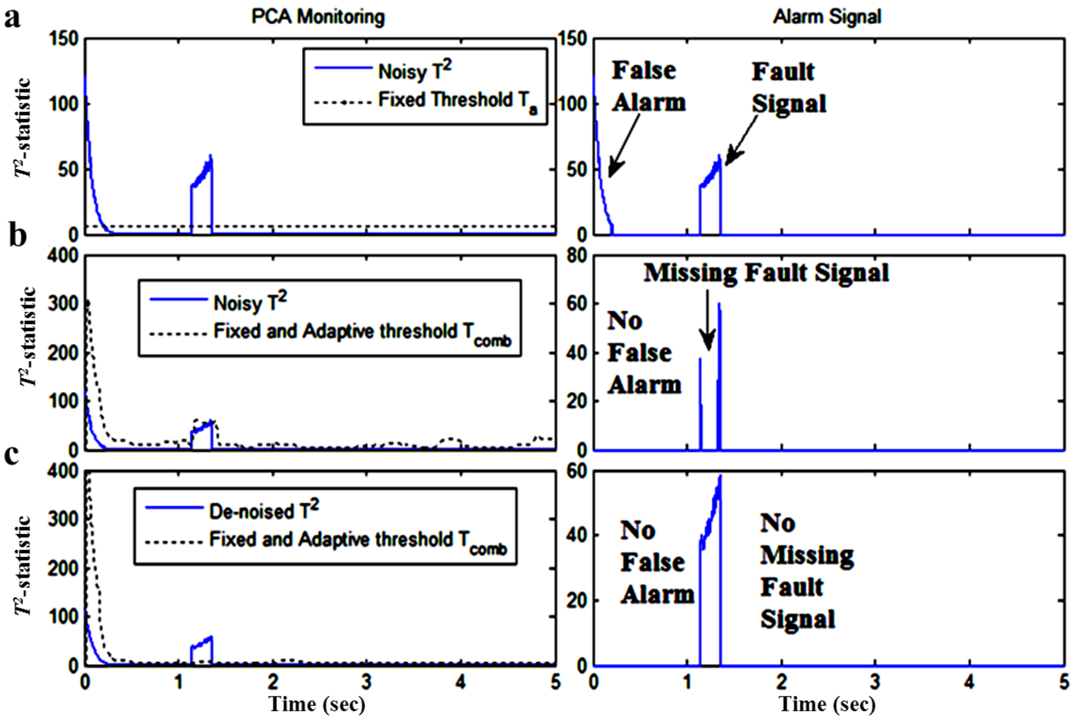


Figure 5.49. PCA Monitoring Charts with the Alarm Signals a) T_α , b) T_{comb} , and c) T_{comb} with Wavelet.

Experiment 3: Transient Test (No Fault)

A square wave input armature voltage is applied to test the transient condition such that the motor is permitted to run between 420 rpm – 1440 rpm. When 2 volts of the armature input voltage is applied to the plant the output speed is measured to be 420 rpm. The input armature voltage is increased from 2 volts to 6 volts as shown in Figure 5.50 and the output speed is measured to be 5.18 volts from the tachometer such that 5.18 volts of the output corresponds to 1440 rpm of speed. Both noisy and de-noised speed and armature current of the motor are plotted in Figure 5.51 and Figure 5.52, respectively.

False alarm signals are produced in the T^2-statistic method as shown in Figure 5.53(a). No alarm signal is produced in the combination of fixed and adaptive threshold-based T^2-statistic method and proposed method, the combination of fixed and adaptive threshold-based T^2-statistic with de-noised by wavelet method. Although no fault is applied to the system, the transient states are evaluated as a fault in the usual T^2-statistic method. This confirms the fact that usual T^2-statistic fault detection methods are not applicable to the systems with variable set-points. On the other hand, the proposed method prevents the false alarm as shown in Figure 5.53(b) and Figure 5.53(c).

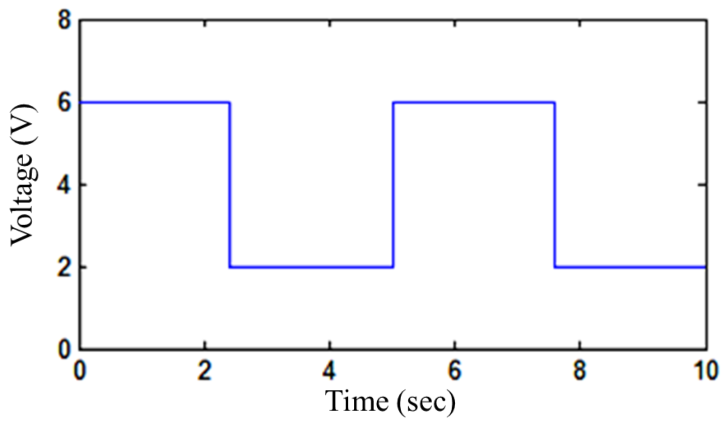


Figure 5.50. Applied Input Voltage.

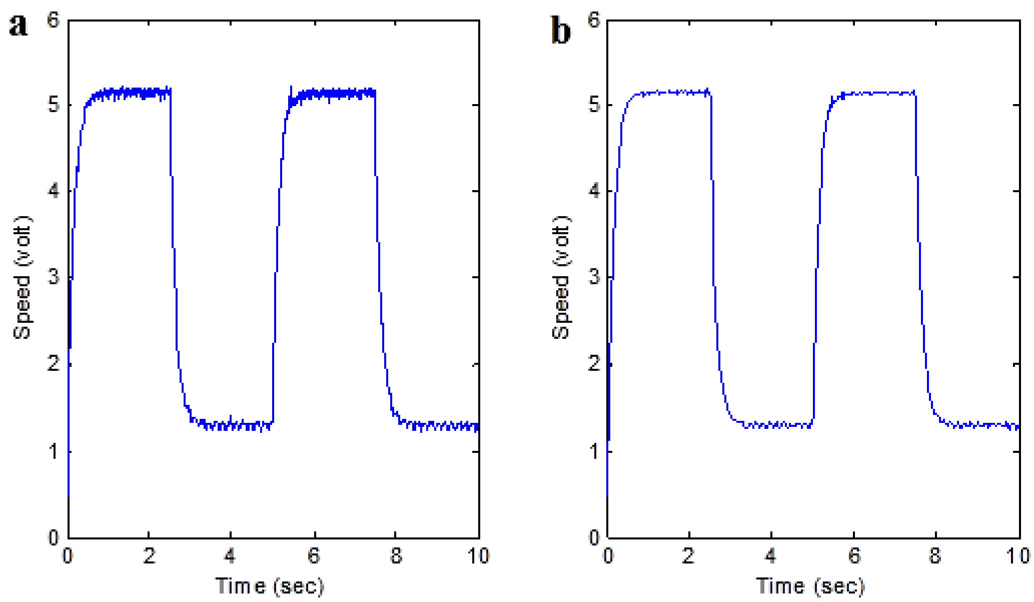


Figure 5.51. Measured Speed a) Noisy, b) De-noised by Wavelet.

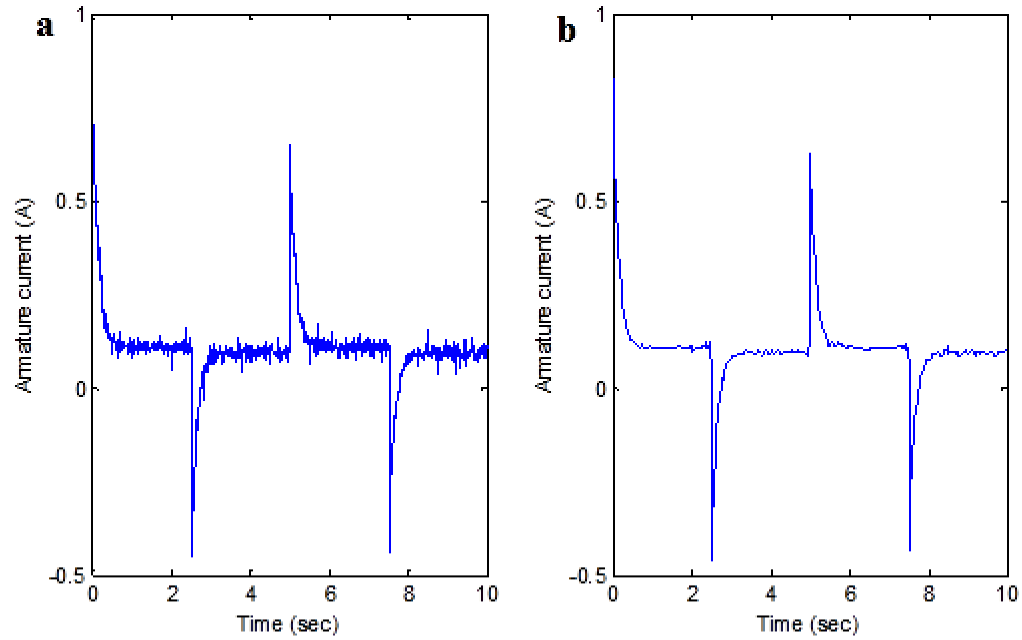


Figure 5.52. Measured Current a) Noisy, b) De-noised by Wavelet.

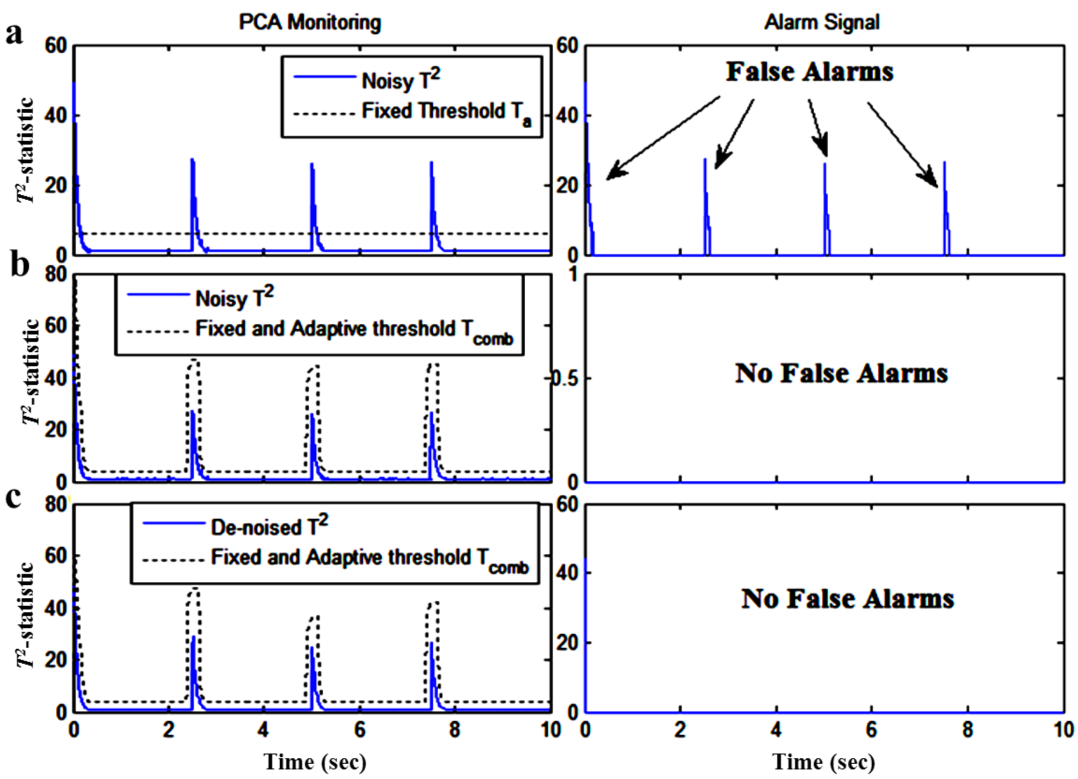


Figure 5.53. PCA Monitoring Charts with the Alarm Signals a) T_α , b) T_{comb} , and c) T_{comb} with Wavelet.

5.2.2.1. (b) Closed Loop Tests

Experiment 4: Actuator Fault

The experimental test is performed that 4 V of the set-point voltage (corresponding to 1130 rpm) is applied to the closed-loop system. Actuator fault is applied to the input (armature voltage) of the system while it is running in the steady-state operation such that the fault signal as shown in Figure 5.54 is applied to test response of the closed-loop system. Both noisy and de-noised signals of the measured speed and current are illustrated in Figure 5.55 and Figure 5.56, respectively.

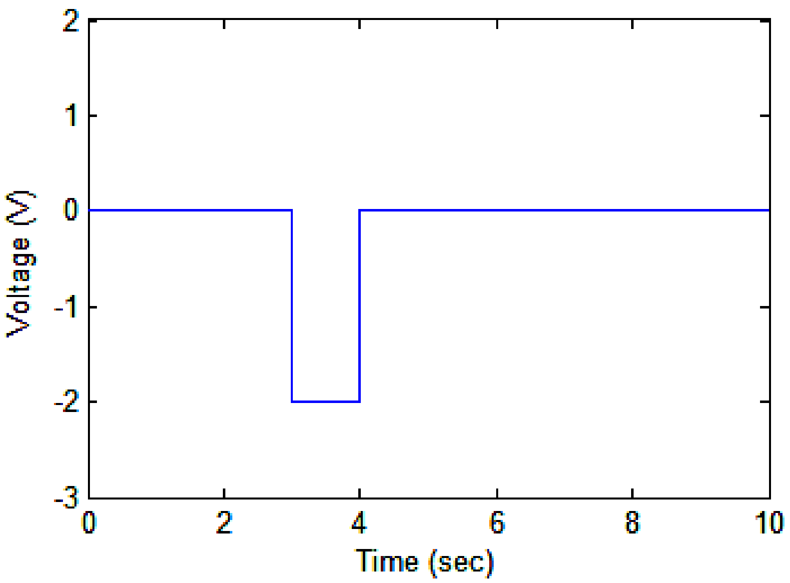


Figure 5.54. Applied Fault Signal to the Input (Armature).

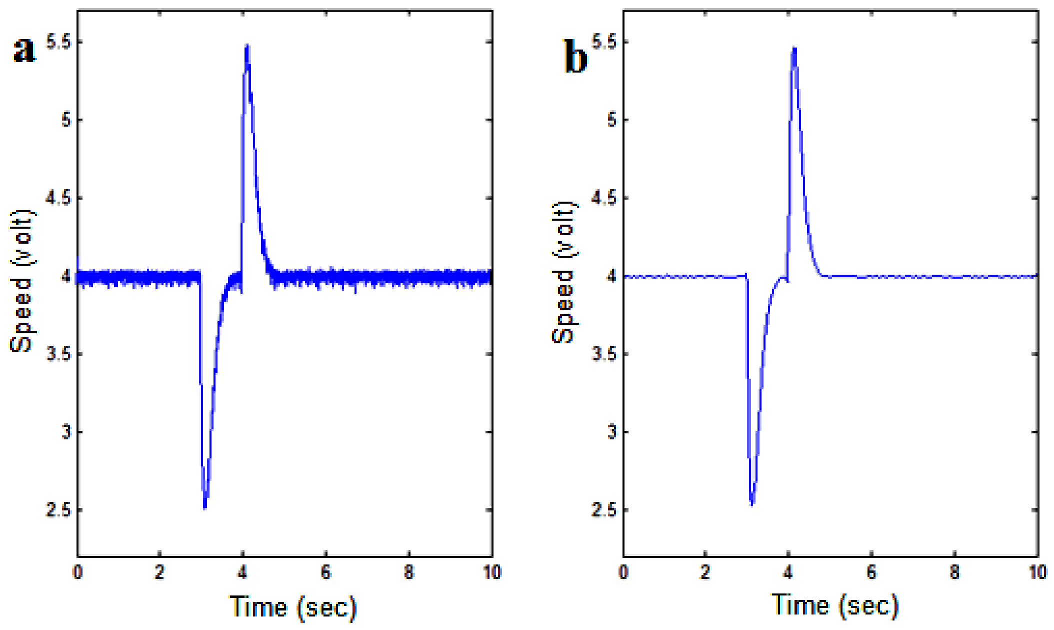


Figure 5.55. Faulty a) Measured Output Shaft Speed, and b) Wavelet Processed Shaft Speed Signals.

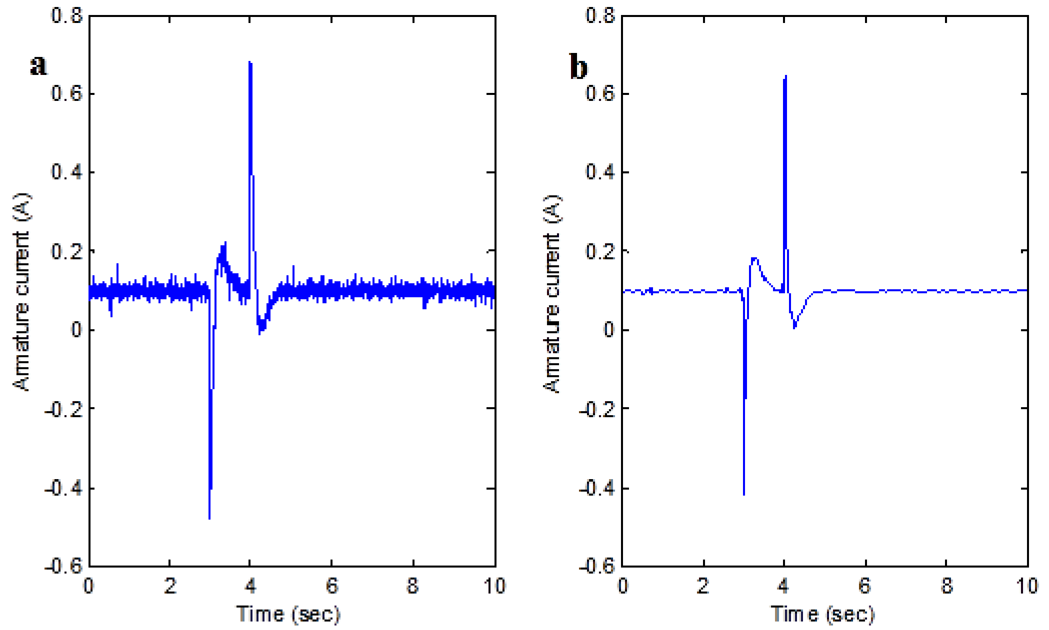


Figure 5.56. Faulty a) Measured Armature Current, and b) Wavelet Processed Armature Current Signals.

Fault alarm signals are indicated in all methods as shown in Figure 5.57 accounted without any false alarm since the system operates in the steady-state condition. However a missing fault signal is observed in the combined method, combination of fixed and adaptive threshold based T^2-statistic method that is illustrated in Figure 5.57(b).

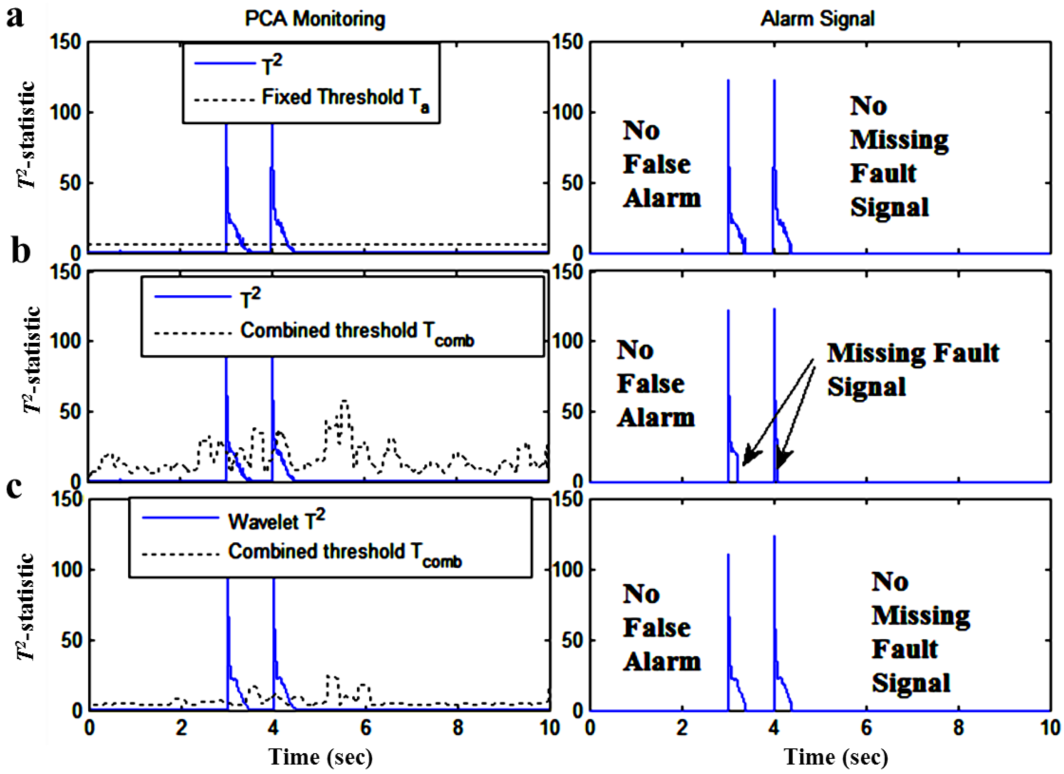


Figure 5.57. PCA Monitoring Charts with the Alarm Signals a) T_α , b) T_{comb} , and c) T_{comb} with Wavelet.

Experiment 5: Sensor Fault

In the current experiment, the fault is applied to the speed sensor. When the system is under the steady-state operation condition the fault is injected by disconnecting and reconnecting the speed sensor for duration of 0.5 seconds at 3.5 sec after the start-up. After the applied sensor fault, the measured noisy speed and denoised speed are shown in Figure 5.58 and the measured noisy current and de-noised current are illustrated in Figure 5.59.

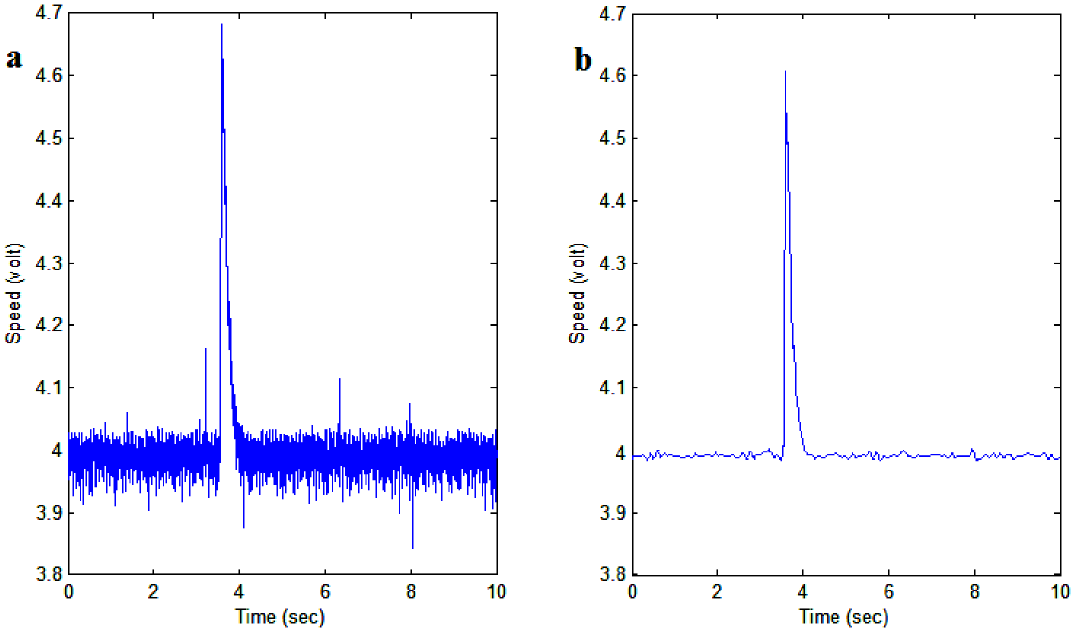


Figure 5.58. Faulty a) Measured Output Shaft Speed, and b) Wavelet Processed Shaft Speed Signals.

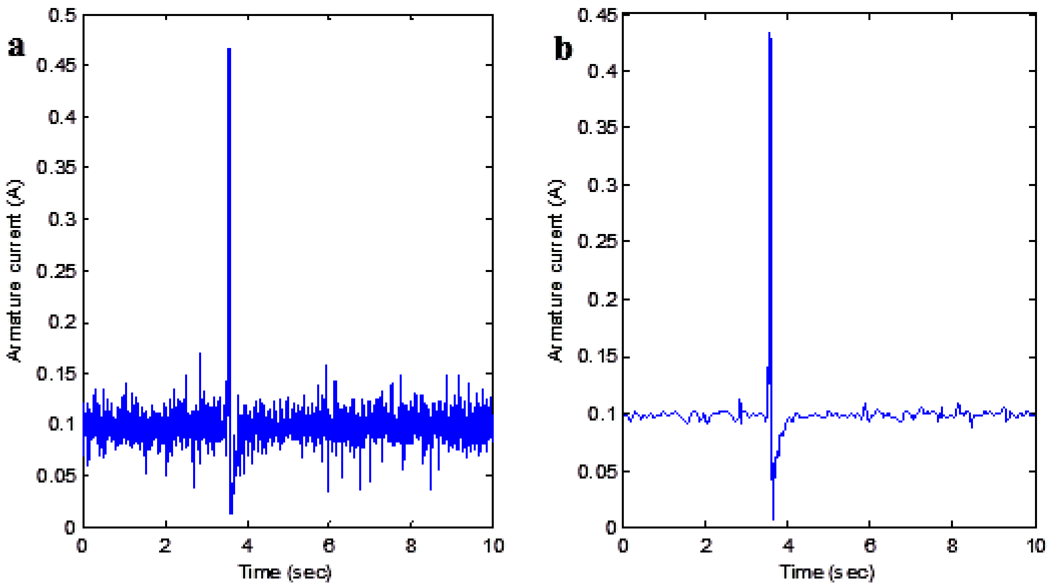


Figure 5.59. Faulty a) Measured Armature Current, and b) Wavelet Processed Armature Current Signals.

The results are plotted in Figure 5.60. The alarm signals are observed in all methods without any false alarm signal since the system operates in the steady-state condition. However a missing fault signal is observed in the combined method, the combination of fixed and adaptive threshold-based noisy T^2-statistic.

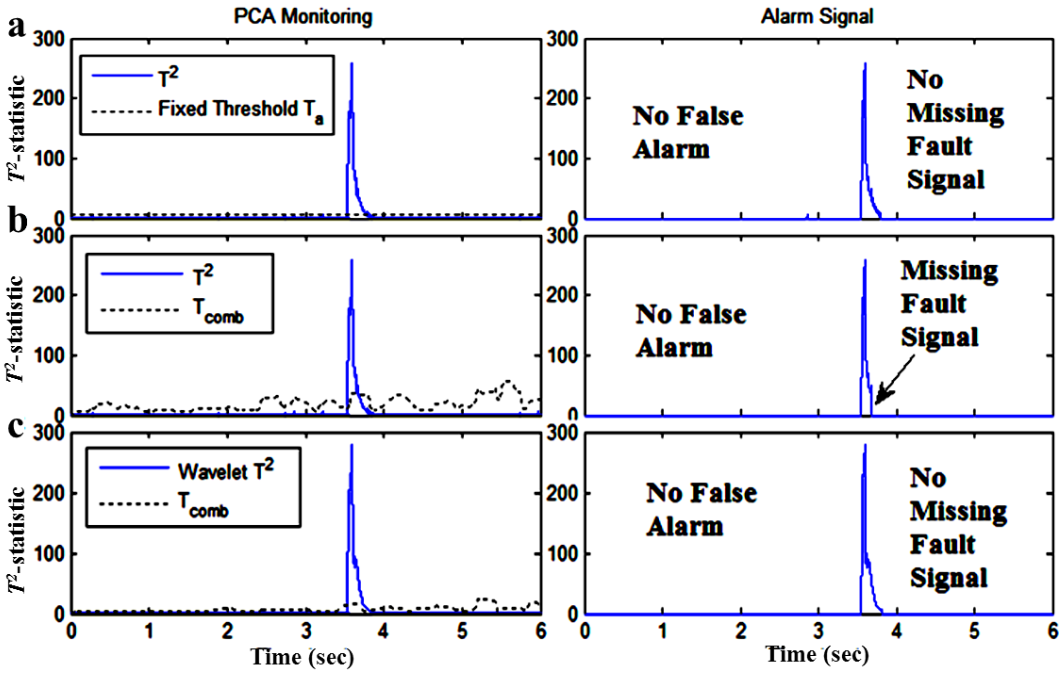


Figure 5.60. PCA Monitoring Charts with the Alarm Signals a) T_α , b) T_{comb} , and c) T_{comb} with Wavelet.

Experiment 6: Transient Test (Sensor Fault)

A variable set-point signal shown in Figure 5.61 is applied to the closed-loop system. Both the raw signal and wavelet processed speed and the armature current of the motor are shown in Figure 5.62 and Figure 5.63, respectively.

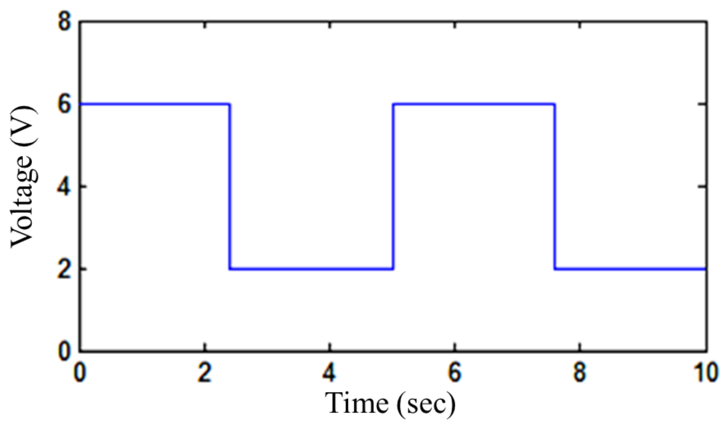


Figure 5.61. Applied Set-Point Speed Signal.

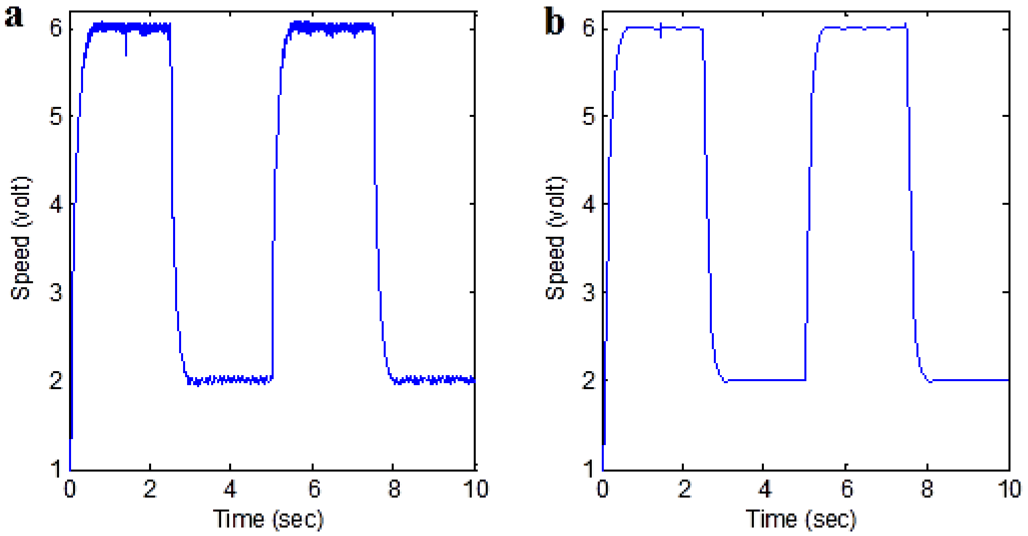


Figure 5.62. Fault-Free a) Measured Output Shaft Speed, and b) Wavelet Processed Shaft Speed Signals.

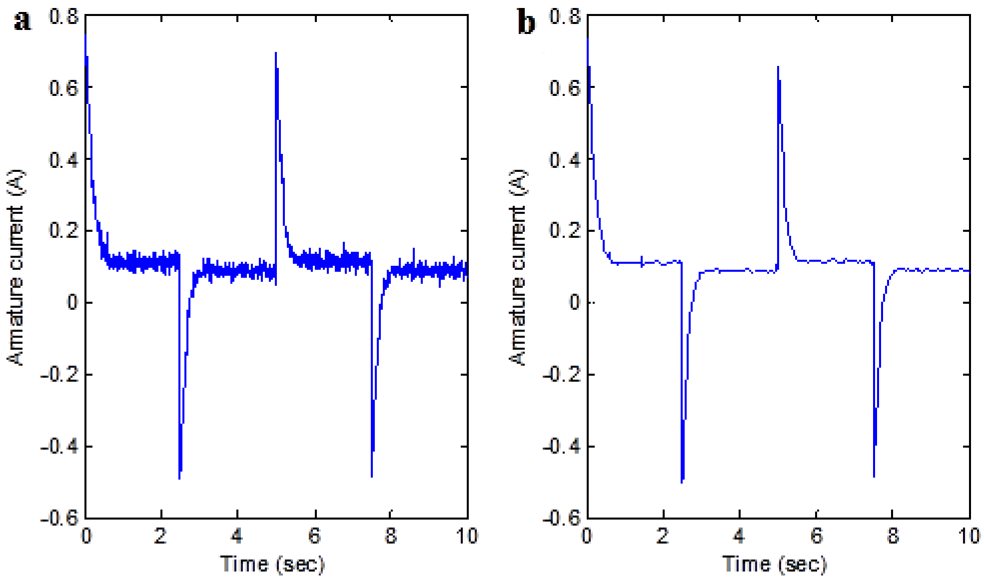


Figure 5.63. Fault-Free a) Measured Armature Current, and b) Wavelet Processed Armature Current Signals.

While the system operates under the servo (tracking) condition as shown in Figure 5. 61, a fault is injected by disconnecting and reconnecting the speed sensor at 1.75 s with duration of 0.7 seconds. After the sensor fault is applied, both the noisy and de-noised speed and current are plotted in Figure 5.64 and Figure 5.65.

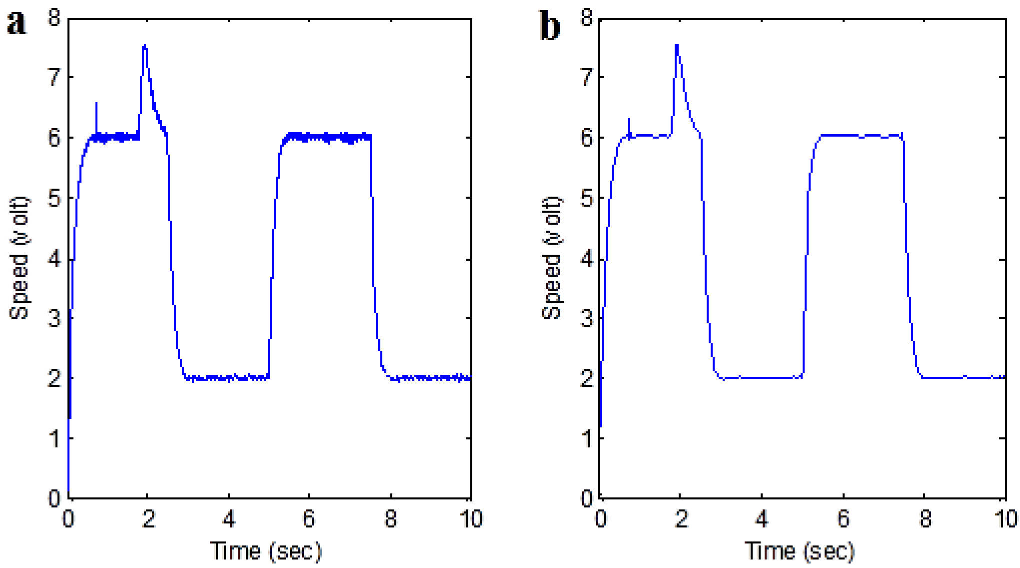


Figure 5.64. Faulty a) Measured Output Shaft Speed, and b) Wavelet Processed Shaft Speed Signals.

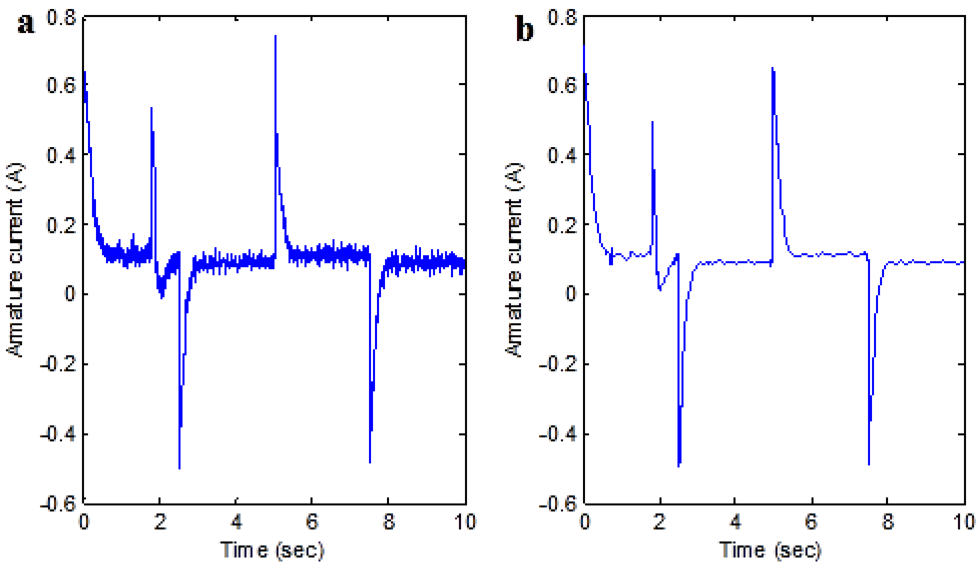


Figure 5.65. Faulty a) Measured Armature Current, and b) Wavelet Processed Armature Current Signals.

False alarms are produced in the T^2-statistic (Figure 5.66(a)) method. False alarm is not indicated in the combined method, combination of fixed and adaptive threshold-based T^2-statistic method but the fault signal appears with the missing component (Figure 5.66(b)). False alarm is not present in the proposed wavelet based

PCA method and the produced fault signal does not include any missing component (Figure 5.66(c)).

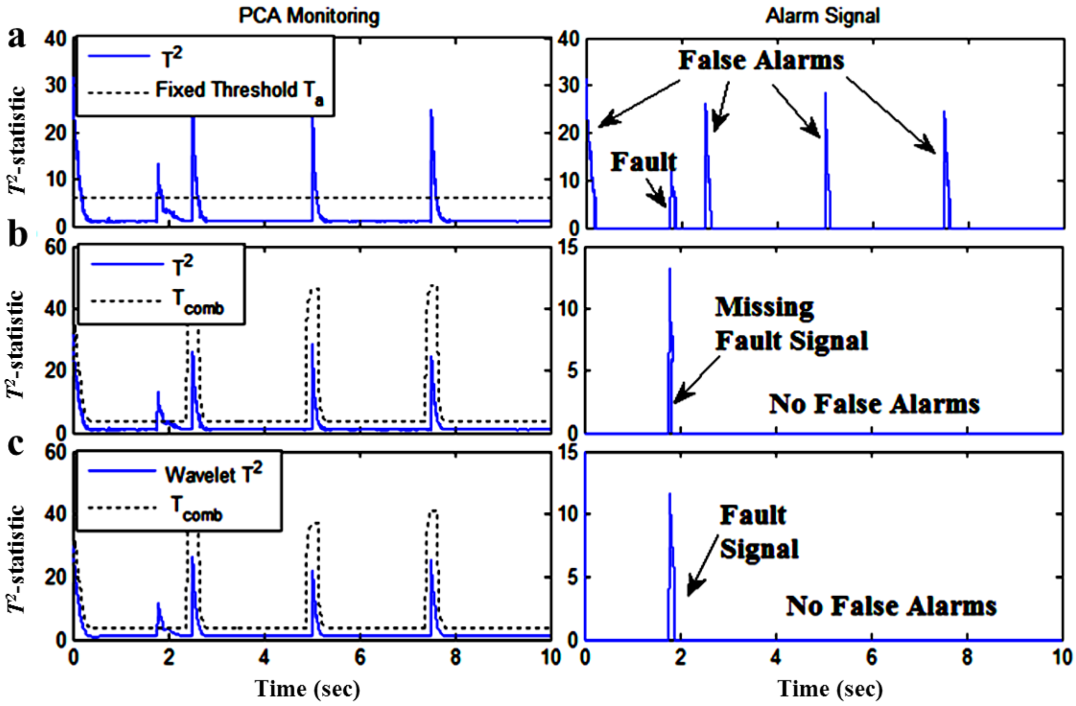


Figure 5.66. PCA Monitoring Charts with the Alarm Signals a) T_α , b) T_{comb} , and c) T_{comb} with Wavelet.

5.2.2.2. Experiments on Process Control System

Experiment 7: Actuator Fault

The experimental test is performed that 2.0 V of the step set-point voltage is applied to the closed-loop system. Original noisy and de-noised signals of the measured level and flow of the process are shown in Figure 5.67 and Figure 5.68, respectively.

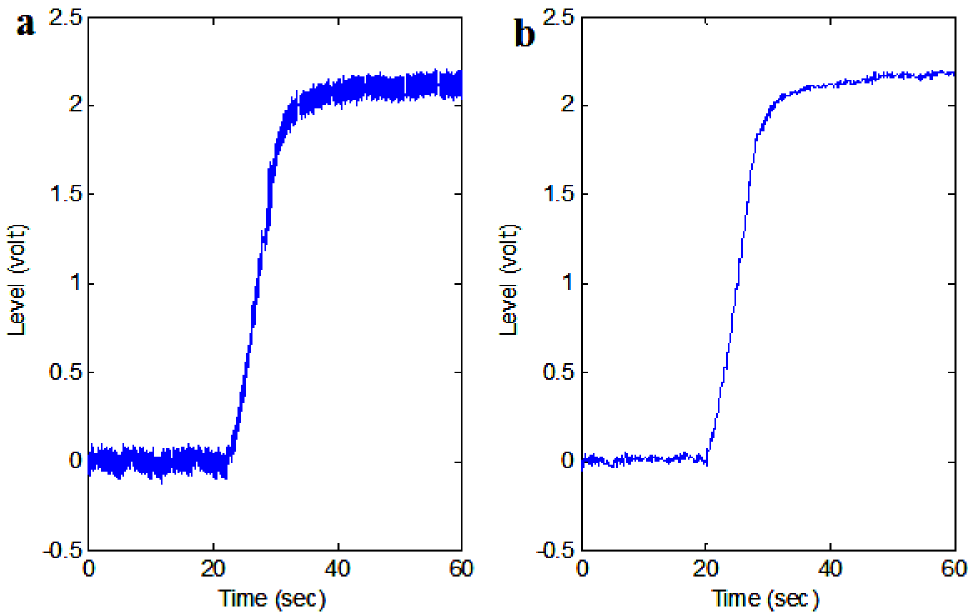


Figure 5.67. Fault-Free a) Measured Liquid Level, and b) Wavelet Processed Liquid Level Signals.

In the current experiment, the fault is applied to the pump input voltage. When the system is under the steady-state operation condition the fault is injected by disconnecting and reconnecting the pump input voltage signal for 2.0 seconds at 21.75 sec after the start-up. After the applied actuator fault, the measured liquid cleared level signal and level signal are shown in Figure 5.69 and the measured flow and wavelet processed flow are illustrated in Figure 5.70.

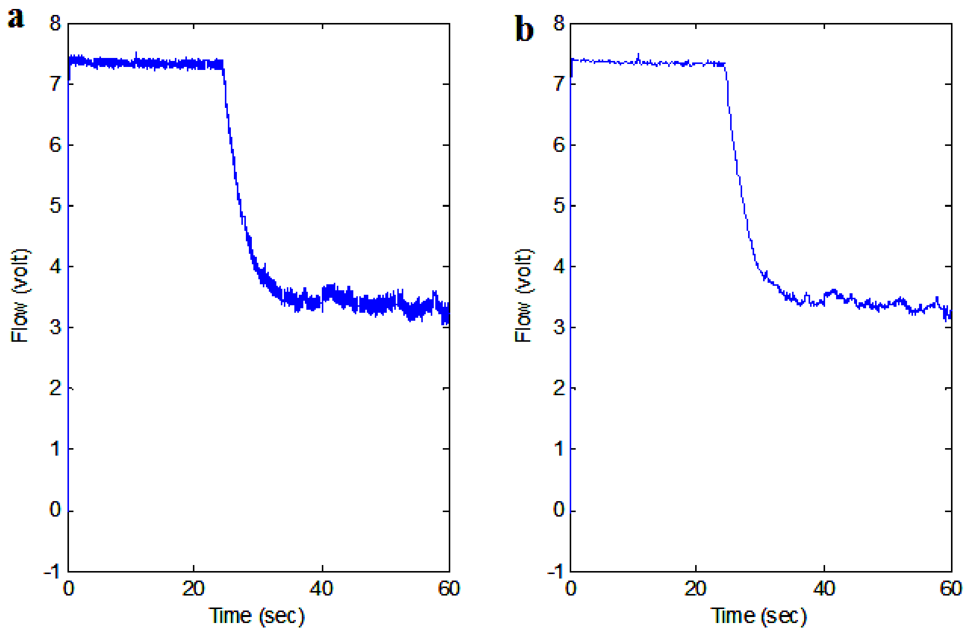


Figure 5.68. Fault-Free a) Measured Flow, and b) Wavelet Processed Flow Signals.

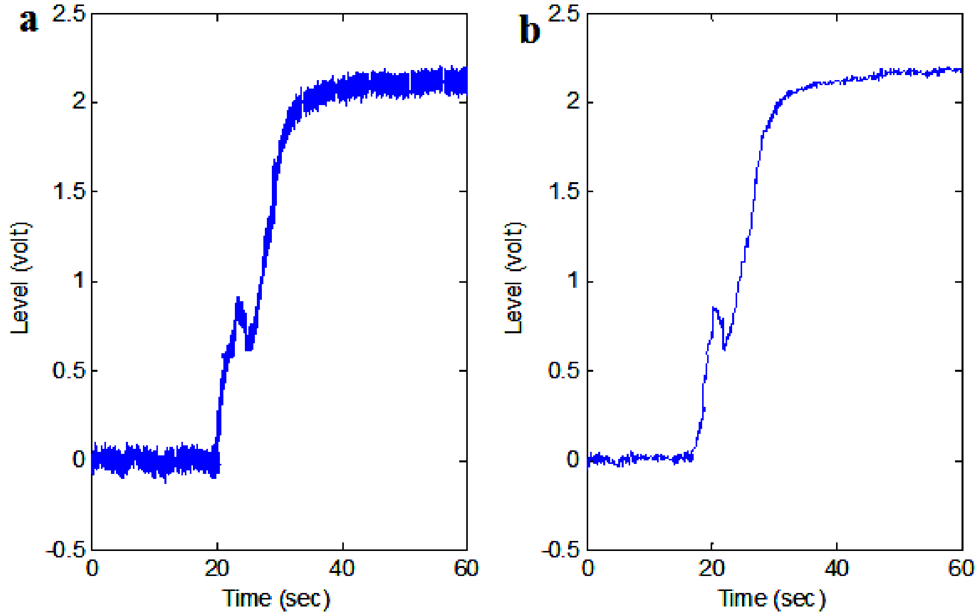


Figure 5.69. Faulty a) Measured Liquid Level, and b) Wavelet Processed Liquid Level Signals.

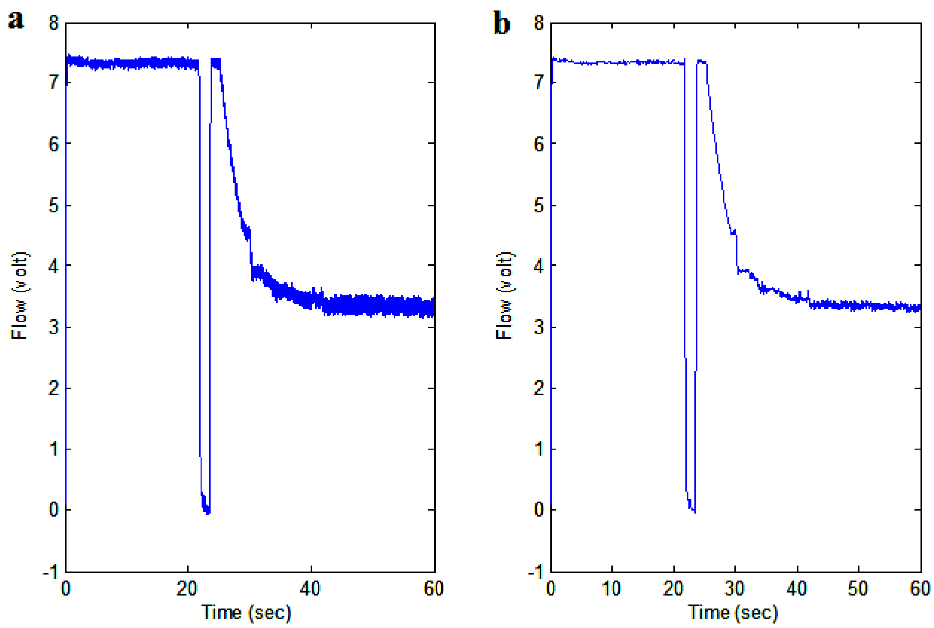


Figure 5.70. Faulty a) Measured Flow, and b) Wavelet Processed Flow Signal.

The results are plotted in Figure 5.71 in which the left column plots show PCA method with their threshold values based on the noisy/de-noised signal and the right column plots show alarm signals produced.

The T^2-statistic method based on the fixed thresholds illustrated in Figure 5.71(a) produces false alarm signals. The missing fault signal components appear in the combined method, combination of fixed and adaptive threshold-based T^2-statistic method as shown in Figure 5.71(b). The alarm signal produced from the proposed method, combined threshold-based T^2 with de-noised by wavelet method is illustrated in Figure 5.71(c) such that both the false alarm caused during the transient state and the missing data are eliminated.

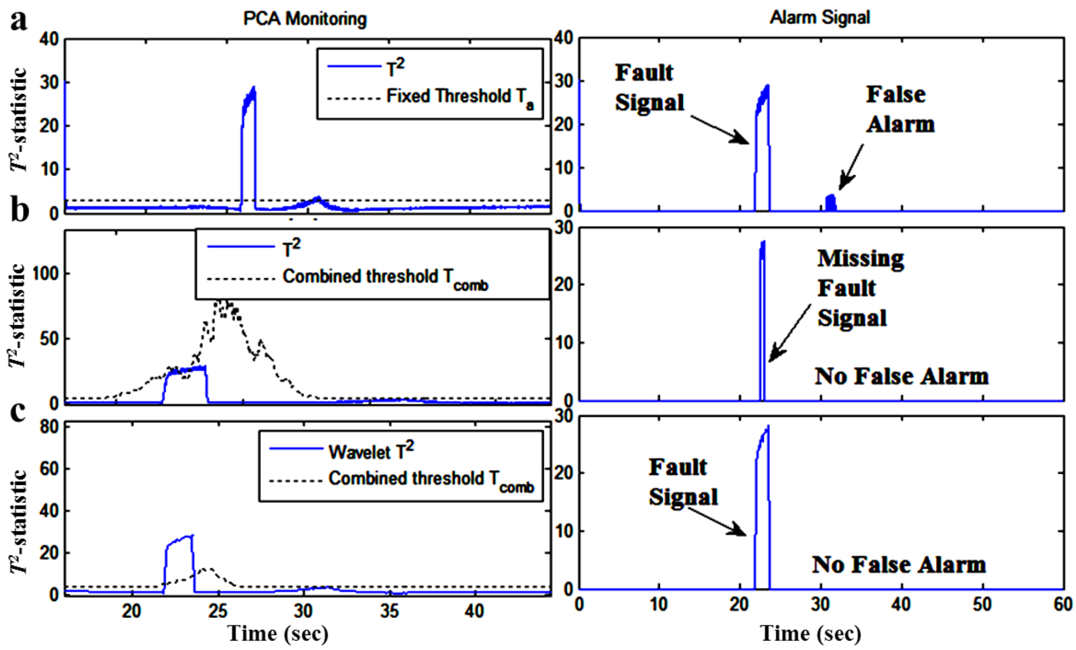


Figure 5.71. PCA Monitoring Charts with the Alarm Signals a) T_{α} , b) T_{comb} , and c) T_{comb} with Wavelet.

Experiment 8: Sensor Fault

When the system is under the steady-state operation condition a fault is injected by disconnecting and reconnecting the speed sensor for duration of 2.5 seconds at 20.0 sec. After the applied fault, measured and wavelet processed output liquid level signals are shown in Figure 5.72. Applied level sensor fault does not affect the flow sensor signal.

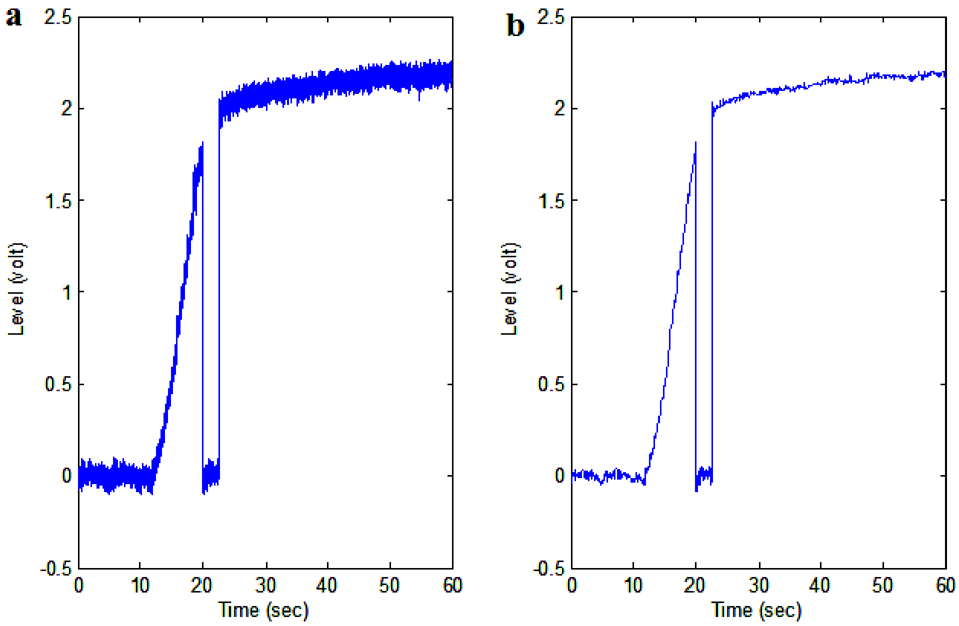


Figure 5.72. Faulty a) Measured Liquid Level, and b) Wavelet Processed Liquid Level Signals.

The results are plotted in Figure 5.73. The false alarm signal is produced in T^2-statistic method as shown in Figure 5.73(a). Interrupted fault signal (missing fault signal component) is produced in the combined method in Figure 5.73(b). Proper fault signal without missing component is produced and the false alarm is eliminated in the proposed method as illustrated in Figure 5.73(c).

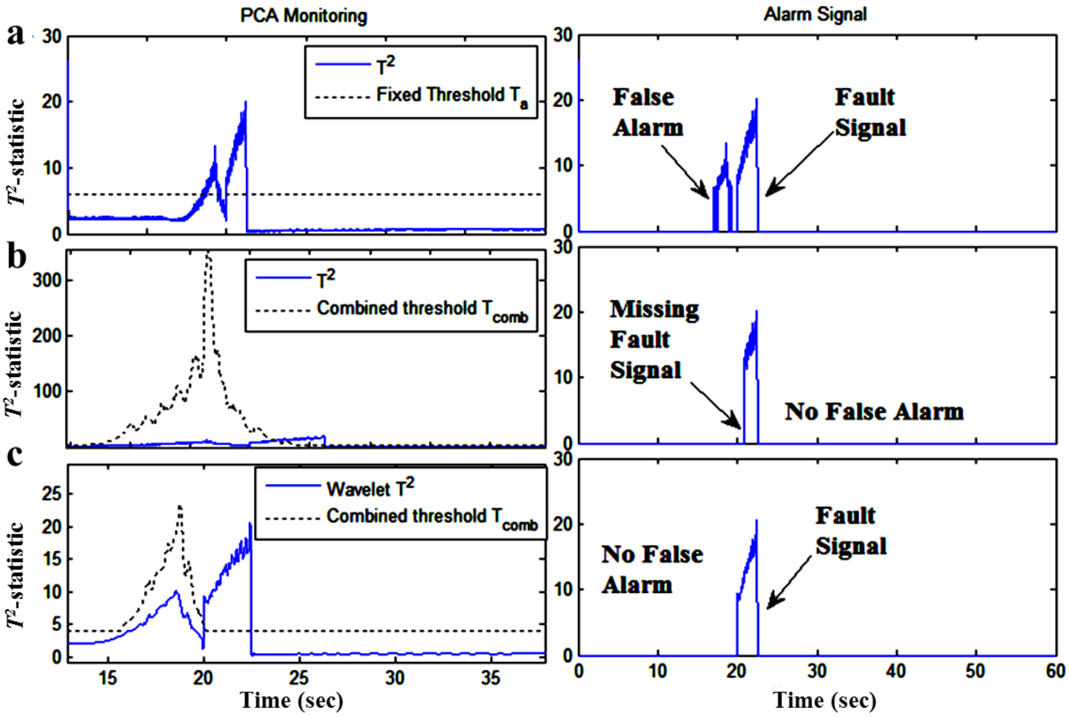


Figure 5.73. PCA Monitoring Charts with the Alarm Signals a) T_{α} , b) T_{comb} , and c) T_{comb} with Wavelet.

Experiment 9: Transient Test (Actuator Fault)

A variable set-point signal is applied to the closed-loop system to test the transient behavior. Level set-point is increased from 2.0 V to 3.0 V at 60 sec after the start-up. Both raw and wavelet processed liquid level and flow signals of the process are plotted in Figure 5.74 and Figure 5.75, respectively.

When the system is under the steady-state operation the fault is injected by disconnecting and reconnecting the pump voltage for duration of 7.5 s at 64.0 seconds. After the applied actuator fault, the measured noisy liquid level and de-noised level are shown in Figure 5.76 and the measured noisy flow and de-noised flow are illustrated in Figure 5.77.

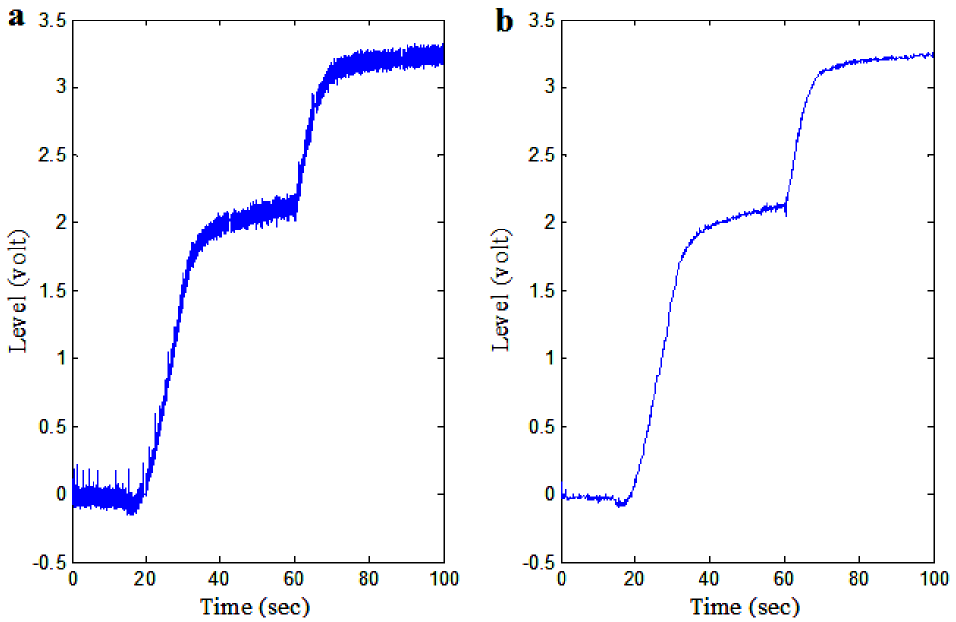


Figure 5.74. Fault-Free a) Measured Liquid Level, and b) Wavelet Processed Liquid Level Signals.

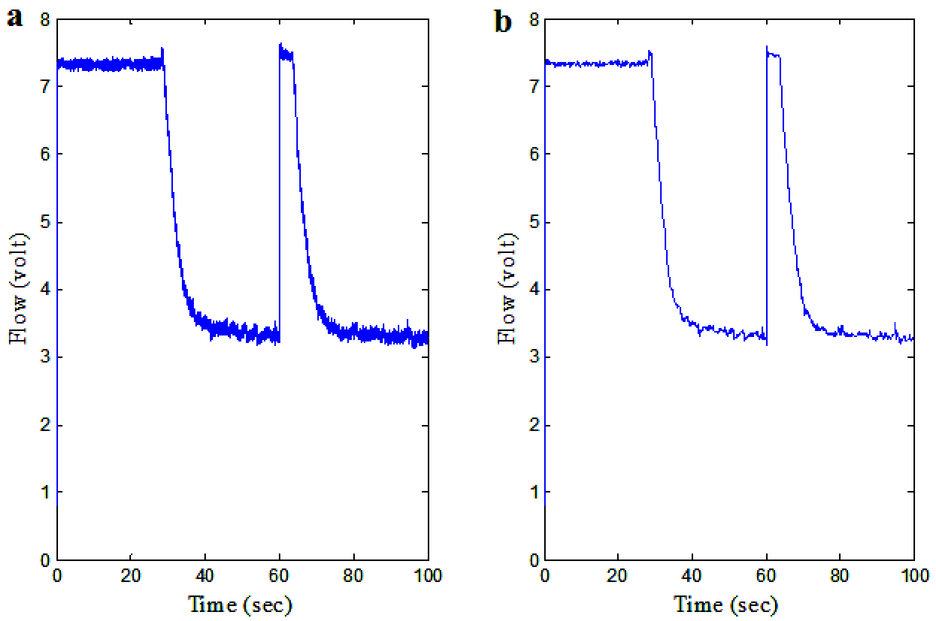


Figure 5.75. Fault-Free a) Measured Flow, and b) Wavelet Processed Flow Signal.

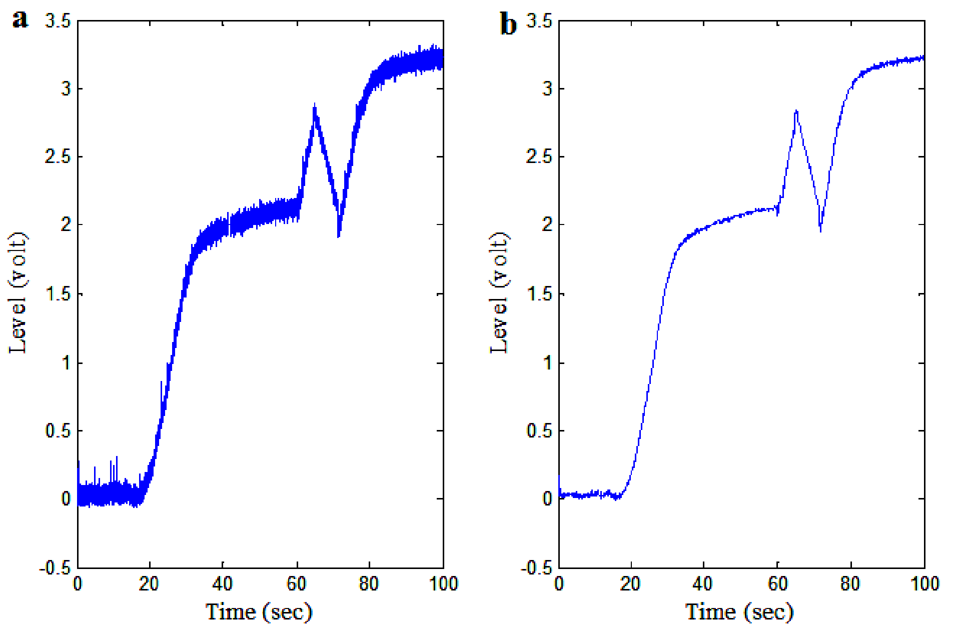


Figure 5.76. Faulty a) Measured Liquid Level, and b) Wavelet Processed Liquid Level Signals.

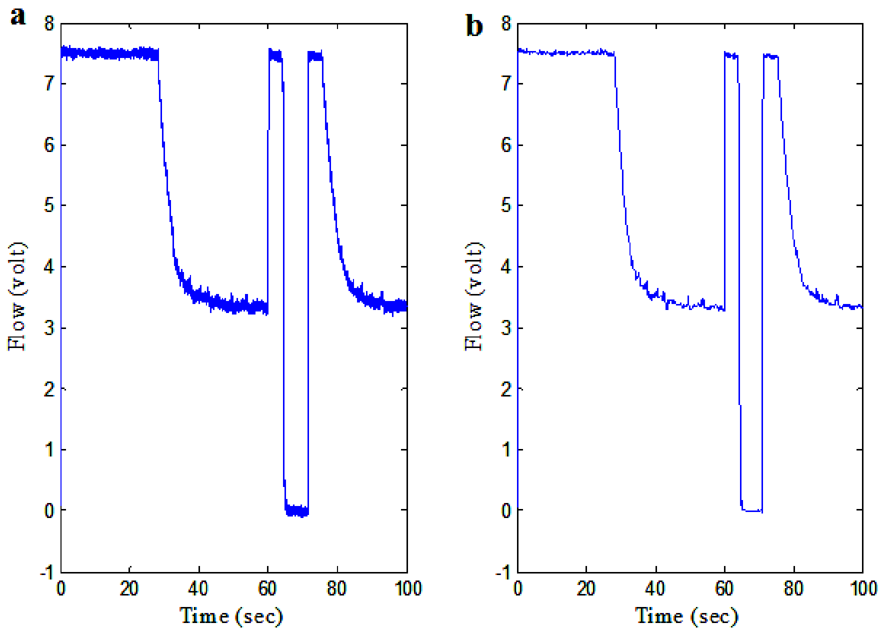


Figure 5.77. Faulty a) Measured Flow, and b) Wavelet Processed Flow Signal.

The results are plotted in Figure 5.78**.** False alarm is produced in T^2-statistic method as shown in Figure 5.78(a). A false alarm is not indicated in the combined threshold-based T^2-statistic method but the fault signal appears with a missing component as illustrated in Figure 5.78(b). False alarm is not present in the proposed method, the combined threshold-based T^2-statistic with wavelet processed, and the produced fault signal does not include any missing (interrupted) component as illustrated in Figure 5.78(c).

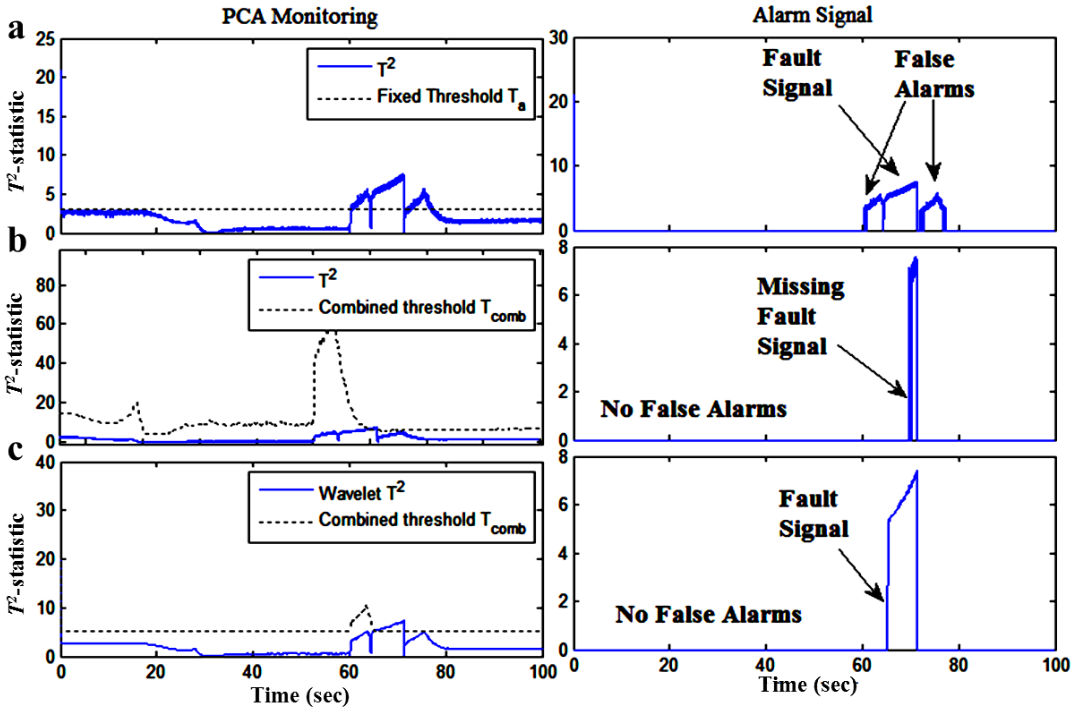


Figure 5.78. PCA Monitoring Charts with the Alarm Signals a) T_{α} , b) T_{comb} , and c) T_{comb} with Wavelet.

Although no fault is applied to the system, the transient states are evaluated as a fault in the usual T^2-statistic method. This confirms the fact that usual T^2-statistic fault detection method is not applicable to the systems with variable set-points. On the other hand, the proposed method prevents the false alarm as shown in Figure 5.78(b) and Figure 5.78(c).

Integral Squared Alarm Signal (ISAS) index is calculated to compare signal energy produced:

$$ISAS = \sum_{n=0}^{N} z^2(n) \tag{5.1}$$

where $z(n)$ denotes the alarm signal with N samples, n is a positive integer.

Calculated ISAS values are presented in Table 5.2 to provide information about the produced alarm signal for fault detection. The numerical values are regulated by 10^3. In Table 5.2, if the calculated numerical values and the alarm signal energy are compared, the results obtained from the proposed method (T_{comb} with wavelet) are numerically larger than the others. It confirms the fact that the alarm signal produced from the proposed method is more powerful than the others for the both electromechanical and process control systems.

Table 5.2. Integral Squared Alarm Signal Comparisons

		Electromechanical Experiments			**Process Control Experiments**		
		EXP1	**EXP2**	**EXP3**	**EXP4**	**EXP5**	**EXP6**
PCA algorithms	**T_α Fixed Threshold**	169.95	419.59	NA	NA	NA	NA
	T_{comb}	147.89	411.12	1.7423	60.705	75.132	20.142
	T_{comb} With Wavelet	**171.96**	**574.63**	**2.5077**	**213.332**	**92.865**	**46.736**

NA: Not Applicable-Due to the false alarms

5.3. Conclusion

The work of this chapter can be summarized in three main experimental applications:

1. ***Application of observer-based fault detection method:*** A scheme for the detection and isolation of sensor faults for electromechanical system is presented. The single observer scheme, based on one Luenberger observer driven by the motor speed sensor, is utilized for fault detection. The capability of the residuals to detect and isolate different kind of faults in the sensors such as abrupt, incipient, intermittent and disconnection faults

are demonstrated. Four different experiments are performed to validate the observer based method.

2. ***Variance sensitive adaptive threshold-based PCA method:*** In order to demonstrate the efficacy of the proposed PCA method, some experimental tests are performed. The experiments are carried out in open-loop and closed –loop conditions to test both the steady-state and transient operating conditions for the actuator and sensor faults under the constant load conditions. The results of the experimental tests (the fixed threshold-based *Q*-statistic (Q_α), T^2-statistic (T_α), combination of fixed and adaptive threshold-based T^2-statistic (T_{comb}) and proposed variance sensitive adaptive threshold-based (T_{vsa}) methods) are presented for PCA monitoring. The alarm signals are computed for each method. The overall experimental results are summarized in a table for a complete comparison. Seven different experiments are performed to validate the variance sensitive adaptive threshold-based PCA method.
3. ***Wavelet based combined PCA method:*** In order to validate the feasibility and reliability of the proposed combined PCA method using wavelet, some experimental tests are performed on an electromechanical system and process control system. Closed-loop experimental tests are performed to test the system for the actuator and sensor faults. The results of the experimental tests obtained from fixed threshold-based T^2-statistic $\left(T_\alpha\right)$, combination of fixed and adaptive threshold-based T^2-statistic $\left(T_{comb}\right)$, and combination of fixed and adaptive threshold-based T^2-statistic $\left(T_{comb}\right)$ with wavelet based de-noised signal are presented for PCA monitoring. Integral Squared Alarm Signal (ISAS) index is calculated to compare signal energy produced to provide information about the produced alarm signal for fault detection.

6. CONCLUSIONS AND FUTURE WORK

Fault detection and diagnosis has been becoming more and more important for process monitoring because of the increasing demand for higher performance as well as for increased safety and reliability of dynamic systems. Fault detection and diagnosis deals with the timely detection, diagnosis and correction of abnormal conditions of faults in a process. The early detection of the occurrence of faults is critical in avoiding product deterioration, performance degradation, major damage to the machinery itself and damage to human health or even loss of lives.

The development and application of an observer-based scheme for FD has been demonstrated by application to an electromechanical system (DC motor). A scheme for the detection and isolation of sensor faults for the electromechanical system is presented. Single observer scheme, based on one Luenberger observer driven by the motor speed sensor, is utilized for fault detection. The capability of the residuals to detect and isolate different kind of faults in the sensors such as abrupt, incipient, intermittent and disconnection faults are demonstrated.

Date-driven methods for FD use a data collected from a steady-state process to monitor T^2-statistic with fixed threshold. For the systems where the transient values of the processes must also be taken into account, the usage of the fixed threshold in a PCA method causes the false alarms which significantly compromise the reliability of the monitoring systems. A combined algorithm (Combination of Fixed and Adaptive Threshold (T_{comb})) is proposed to overcome the problems raised from the fixed threshold and provide required confidence limit. However, the data collected from industrial processes often contain measurement noise that causes to produce the missing fault signal components when the combined threshold method is used.

In order to overcome drawbacks, variance sensitive adaptive threshold method is proposed that is sensitive to the high variance which occurs due to noise. Wavelet method is used remove the noise before combined PCA analysis.

The proposed methods are implemented and validated experimentally both on a process control system and an electromechanical control system operating in open-loop and closed-loop conditions. Actuator fault, sensor fault and servo tracking conditions are experienced to demonstrate advantages and applicability of the proposed methods. The experimental results are compared with the results obtained from the conventional PCA monitoring methods. Tabulated data and experimental test results confirm the fact that the fixed threshold-based T^2-statistic methods produce false alarms during the variable set-point signals. Fault signal with the missing (interrupted) components is obtained with the combined threshold-based T^2-statistic method. The false alarm and missing alarm signal problems are eliminated using the proposed variance sensitive adaptive threshold-based and combined threshold-based T^2-statistic with wavelet processed methods.

The main differences and important contributions of this research can be stated as:

- Application of the observer based fault detection method to an electromechanical system.
- A combined threshold PCA method is proposed and implemented in the present research to prevent false alarm signal when transient states taken into account.
- A variance sensitive adaptive threshold PCA method is proposed and implemented to prevent missing fault alarm signal.
- Wavelet-based combined PCA method is implemented to overcome the false alarms and to produce uninterrupted fault alarm signal.

Some research topics that can be studied in the future are discussed as follows:

- An alternative design approach of residual generators can also be derived (Parameter estimation, Parity space).
- Another challenging research direction is to investigate the FD problem of nonlinear systems.
- Adaptive threshold can be developed for residual signal.

- Generated residual can be evaluated by wavelet.
- Partial Least Square (PLS) or Independent Component Analysis (ICA) can be used as statistical tools for monitoring processes.
- Proposed threshold algorithm can be developed for Q-statistics.
- All of the developed methods can be used for on-line applications to check the real time applicability.

REFERENCES

ALKAYA, A., and EKER, İ., 2011. Variance sensitive adaptive threshold-based PCA method for fault detection with experimental application. ISA Transactions, 50:287-302.

ANGELI, C., 2004. On-Line Fault Detection Techniques for Technical Systems: A Survey. International Journal of Computer Science and Applications, 1(1):12 – 30.

ANTORY, D., 2007. Application of a data-driven monitoring technique to diagnose air leaks in an automotive diesel engine: A case study. Mechanical Systems and Signal Processing, 21:795–808.

BAKSHI, B. R., and STEPHANOPOULOS, G., 1993. Wave-net: A multi resolution, hierarchical neural network with localized learning. AICHE J. 39(1):57 – 81.

BAKSHI, B. R., 1998. Multi scale PCA with applications to multivariate statistical process monitoring. AIChE Journal 44(7):1596 – 1610.

BARALDI, P., CAMMI, A., MANGILI, F., and ZIO, E., 2010. An ensemble approach to sensor fault detection and signal reconstruction for nuclear system control. Annals of Nuclear Energy, 37:778–790.

BASILA, M., JR., STEFANEK, G., and CINAR, A., 1990. A model-object based supervisory expert system for fault tolerant chemical reactor control. Computers and Chemical Engineering 14(4 – 5):551 – 560.

BEARD, R., 1971. Failure accommodation in linear systems through self – reorganization, Technical Report MVT-71-1. Cambridge (Man Vehicle Laboratory): Mass.

BECRAFT, W., and LEE, P., 1993. An integrated neural network/expert system approach for fault diagnosis. Computers and Chemical Engineering 17 (10):1001 – 1014.

BENDAPUDI, S., and BRAUN J. E., 2002. A review of literature on dynamic models for vapor compression equipment. HL 2002-9, Report #4036-5, Ray Herrick Laboratories, Purdue University.

BHATTACHARYA, R., and WAYMIRE, E. C., 1990. Stochastic processes with applications. New York: Wiley.

BIN, H. X., and YANG, Y. P., 2008. Variable MWPCA for adaptive process monitoring. Industrial & Engineering Chemistry Research, 47(2):419 – 427.

BLANKE, M., KINNAERT, M., LUNZE, J., and STAROSWIECKI, M., 2003. Diagnosis and Fault-Tolerant Control, Springer-Verlag.

BURRUS, C. S., COPINATH, R. A., and GUO, H., 1998. Introduction to wavelets and wavelet transforms, Prentice Hall.

CARPENTER, C. A. and GROSSBERG, S., 1988. The art of adaptive pattern recognition by a self-organizing neural network. Computer 21 21(3):77 – 88.

CHEN W., KHAN A.Q., ABID M., and DING S.X., 2011. Integrated Design of Observer Based Fault Detection for a Class of Uncertain Nonlinear Systems. Int. J. Appl. Math. Computer, Sci., 21(3):423–430.

CHEN, B. H., WANG, X., YANG, Z. S. H., and MCGREAVY, C., 1999. Application of wavelets and neural networks to diagnostic system development, 1. Feature extraction. Computers and Chem. Eng. 23(7):899 – 906.

CHEN, J., and PATTON, R., 1999. Robust model-based fault diagnosis for dynamic systems. Kluwer, Boston.

CHEN, W., and CHOWDHURY, F. N., 2010. A synthesized design of sliding-mode and Luenberger observers for early detection of incipient faults. International Journal of Adaptive Control and Signal Processing, 24(12):1021-1035.

CHESTER, D., LAMB, D., and DHURJATI, P., 1984. Rule-based computer alarm analysis in chemical process plants. In Proceedings of 7th Micro-Delcon. 22 – 29.

CHEUNG, J. T., and STEPHANOPOULOS, G., 1990. Representation of process trends part I. A formal representation framework. Computers and Chemical Engineering 14(4 – 5):495 – 510.

CHO, J. H., LEE, J. M., CHOI, S. W., LEE, D., and LEE, B., 2005. Fault identification for process monitoring using kernel principal component analysis. Chemical Engineering Science, 60(1):279–288.

CHOW, E. Y., and WILLSKY, A. S., 1984. Analytical redundancy and the design of robust failure detection systems. IEEE Transactions on Automatic Control, vol. 29(7):603–614.

CHUI, C. K., 1992. An introduction to wavelets, volume 1, Academic Press.

DEPERSIS, C., and ISIDORI, A., 2000. A Geometric Approach to Nonlinear Fault Detection and Isolation. In proc.: IFAC Safeprocess2000. pp. 209–214.

DING X., and JEINSCH T., 1999. An approach to analysis and design of observer and parity relation based FDI systems. 14th IFAC World Congress. July, 5-9, Beijing, China.

DONG, D., and MCAVOY, T. J., 1996. Nonlinear Principal Component Analysis - Based on principal curves and neural networks. Computers and Chemical Engineering 20(1):65-78.

DONOHO, D. L., 1994. De-noising by soft-thresholding. IEEE Trans. Inf. Theory, 41:613–627.

DONOHO, D. L., and JOHNSTONE, I. M., 1995. Adapting to unknown smoothness via wavelet shrinkage. Journal of American Statistical Assoiciation, 90:1200–1224.

EDWARDS, C., SPURGEON S. K., and PATTON R. J., 2000. Sliding mode observers for fault detection and isolation. Automatica, 36(4):541–553.

EDWARDS C., and YAN X. G., 2007. Nonlinear robust fault reconstruction and estimation using a sliding mode observer. Automatica, 43:1605-1614.

EKER, İ., 2010. Second-order sliding mode control with experimental application. ISA Transactions, 49:394–405.

ELSHENAWY, L. M., YİN, S., NAİK, A. S., and DİNG, S. X., 2010. Efficient recursive principal component analysis algorithms for process monitoring. Ind. Eng. Chem. Res., 49:252–259.

EVA, P. E., 1996. The MATLAB handbook. Harlow: Addison-Wesley.

FALKENHAINER, B., and FORBUS, K., 1991. Compositional modeling: Finding the right model for the job. Artificial Intelligence, 51:95 – 143.

FAN, J. Y., NIKOLAOU, M., and WHITE R. E., 1993. An approach to fault diagnosis of chemical processes via neural networks, AICHE J. 39(1):82 – 88.

FANGLAI, Z., and FENG, C., 2010. Full-order observer-based actuator fault detection and reduced-order observer-based fault reconstruction for a class of uncertain nonlinear systems. Journal of Process Control, 20(10):1141-1149.

FISCHER, D., BÖRNER, M., SCHMITT, J., and ISERMANN, R., 2007. Fault detection for lateral and vertical vehicle dynamics. Control Engineering Practice, 15(3):315–324.

FOURIE, S. H., and DEVAAL, P., 2000. Advanced process monitoring using an on-line non-linear multi scale principal component analysis methodology. Computers and Chemical Engineering 24(2-7):755-760.

FOURIE, S., 2000. Dissertation: Advanced Process Monitoring using Wavelets and Non-Linear Principal Component Analysis. Department of Chemical Engineering, University of Pretoria.

FRANK P.M., and DING S.X., 1997. Survey of robust residual generation and evaluation methods in observer-based fault detection systems. Journal of Process Control, 7(6):403–424.

FRANK, P. M., 1996. Analytical and Qualitative Model-based Fault Diagnosis - A Survey and Some New Results. European Journal of Control 2(1):6–28.

GABOR, D., 1964. Theory of communication, J. IEE, 93:429-457.

GANESAN, R., DAS, T. K., and VENKATASUBRAMANIAN, V., 2004. Wavelet-based multiscale statistical process monitoring: A literature review. IIE Transactions, 36(9):787-806.

GARCÍA, E. A., and FRANK, P. M., 1997. Deterministic Nonlinear Observer-Based Approaches to Fault Diagnosis: A Survey. Control Engineering Practice 5(5):663–760.

GENOVESI, A., HARMAND, J., and STEYER, J. P., 1999. A fuzzy logic based diagnosis system for the on-line supervision of an anaerobic digestor pilot-plant. Biochemical Eng. J., 3(3):171 – 183.

GERTLER, J., 2000. All linear methods are equal and extendible to nonlinearities. IFAC Sefeprocess. Jun, 14-16, Pages 52 – 63, Budapest, Hungary.

GERTLER, J., and MONAJEMY, R., 1995. Generating directional residuals with dynamic parity relations. Automatica, 31(4):627 – 635.

GERTLER, J., 1998. Fault detection and diagnosis in engineering systems. Marcel Dekker, New York.

HAMMOURI, H., and TMAR, Z., 2010. Unknown input observer for state affine systems: A necessary and sufficient condition. Automatica, 46(2):271–278.

HAN, C., SHIH, R., and LEE, L., 1994. Quantifying signed directed graphs with the fuzzy set for fault diagnosis resolution improvement. Indust. and Eng., Chemistry Research, 33(8):1943 – 1954.

HE, T., XIE, W. R., WU, Q. H., and SHI, T. L., 2006. Process fault detection and diagnosis based on principal component analysis. IEEE Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian.

HENLEY, E. J., 1984. Application of expert systems to fault diagnosis. In AIChE annual meeting, San Francisco, CA.

HIMMELBLAU, D., 1978. Fault detection and diagnosis in chemical and petrochemical processes, New York: Elsevier.

HOSKINS, J. C., KALIYUR, K. M., and HIMMELBLAU, D. M., 1991. Fault diagnosis in complex chemical plants using artificial neural networks. AICHE J. 37(1):137 – 141.

HOTELLING, H., 1947. Multivariate quality control illustrated by the testing of sample bombsights. In: Eisenhart C, Hastay MW, and Wallis WA (Eds.), Selected techniques of statistical analysis, New York, McGraw-Hill.

HUI, L., YUPING, Z., and HAIQI, Z., 2011. Application of Hermitian wavelet to crack fault detection in gearbox. Mechanical Systems and Signal Processing, 25:1353–1363.

IRI, M., AOKI, K., OSHIMA, E., MATSUYAMA, H., 1979. An algorithm for diagnosis of system failures in the chemical process. Computers and Chem. Eng., 3(1-4):489 – 493.

ISERMANN, R., 1984. Process fault detection based on modeling and estimation methods–A survey. Automatica, 20(4):387 – 404.

ISERMANN, R., 2005. Model-based fault-detection-status and applications. Annual Reviews in Control, 29:71 – 85.

ISERMANN, R., 2006. Fault-diagnosis systems–An introduction from fault detection to fault tolerance, Springer, Heidelberg, 475 p.

ISERMANN, R., 2011. Fault–Diagnosis Applications, Springer Heidelberg Dordrecht, London New York, 372 p.

ISERMANN, R., and BALLÉ P., 1997. Trends in the application of model – based fault detection and diagnosis of technical processes. Control Eng. Practice, 5(5):709 – 719.

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

JACKSON, J. E, and MUDHOLKAR, G. S., 1979. Control procedures for residuals associated with principal component analysis. Technometrics, 21(3):341–349.

JACKSON, J. E., 1991. A User's Guide to Principal Components, Wiley, New York.

JANG, J. O., JEON, and G. J., 2000. A parallel neuro-controller for DC motors containing nonlinear friction. Neuro Computing, 30:233–48.

JANUSZ, M., and VENKATASUBRAMANIAN, V., 1991. Automatic generation of qualitative description of process trends for fault detection and diagnosis. Engineering Applications of Artificial Intelligence 4(5):329 – 339.

JIA, M., CHU, F., WANG, F., and WANG, W., 2010. On-line batch process monitoring using batch dynamic kernel principal component analysis. Chemometrics and Intelligent Laboratory Systems, 101:110–122.

JOHNSON, R., A., and WICHERN, D.W., 1992. Applied Multivariate Statistical Analysis, Prentice-Hall, London.

JOLLIFE, I. T., 1986. Principal component Analysis. Springer, New York.

KANEV S., 2004. Robust fault – tolerant control, PhD thesis, University of Twente, Netherlands.

KANO, M., NAGAO, K., OHNO, H., et al., 1999. Process monitoring using moving principal component analysis. Kagaku Kogaku Ronbunshu, 25(6):998-1003.

KANO, M., HASEBE, S., HASHIMOTO, I., and OHNO, H., 2001 A new multivariate statistical process monitoring method using principal component analysis. Computers and Chemical Engineering, vol. 25, pp. 1103–1113.

KARAMI, F., POSHTAN, J., and POSHTAN, M., 2010. Detection of broken rotor bars in induction motors using nonlinear Kalman filters. ISA Transactions, 49(2):189–95.

KATIPAMULA, S., and BRAMBLEY, M. R., 2005. Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I. HVAC&R Research, 11(1):3 – 25.

KAY, H., and KUIPERS, B., 1993. Numerical behavior envelopes for qualitative models, In "Proceedings of AAAI-93", pp. 606—613, Menlo Park, CA.

KELLY, B. E., and LEES, F. P., 1986. The propagation of faults in process plants: 2. fault tree synthesis. Reliability Eng., 16(1):39 – 62.

KOKAWA, M., SATOSHI, M., and SHIGAI, S., 1983. Fault location using digraph and inverse direction search with application. Automatica, 19(6):729 – 735.

KOURTI, T., 2005, Application of latent variable methods to process control and multivariate statistical process control in industry. Journal of Adaptive Control and Signal Processing, 19:213–246.

KRAMER, M. A., and PALOWITCH, B. L., 1987. A rule based approach to fault diagnosis using the signed directed graph. AIChE J., 33(7):1067 -1078.

KRAMER, M. A., 1991. Nonlinear Principal Component Analysis using autoassociative neural networks. AIChE Journal 37(2):233-243.

KRESTA, J. V., MACGREGOR, J. F, and MARLIN, T. E., 1991. Multivariate Statistical Monitoring of Process Operating Performance. The Canadian Journal of Chemical Engineering 69(1):35-47.

KU, W., STORER, R. H., and GEOGAKIS, C., 1995. Disturbance Detection and Isolation by Dynamic Principal Component Analysis. Chemometrics and Intelligent Laboratory Systems, 30(1):179 – 196.

KUIPERS, B., 1986. Qualitative simulation, Artificial Intelligence, 29(3):289 – 338.

KUNPENG, Z., SAN, W. Y., and SOON, H. G., 2009. Wavelet analysis of sensor signals for tool condition monitoring: A review and some new results. Int. Journal of Machine Tools & Manufacture, 49:537–553.

LEE, J. M., YOO, C. K., CHOI, S. W., VANROLLEGHEM, P. A., and LEE, I. B., 2004. Non-linear Process Monitoring using Kernel Principal Component Analysis. Chemical Engineering Science, 59:223–224.

LENNOX, J., and ROSEN, C., 2002. Adaptive multi scale principal components analysis for online monitoring of wastewater treatment. Water Science and Technology, 45(4–5):227–235.

LEUNG, D., and ROMAGNOLI, J., 2000. Dynamic probabilistic model-based expert system for fault diagnosis. Computers and Chemical Engineering, 24 (11):2473 – 2492.

LI, R., and WANG X., 2001. Qualitative/Quantitative simulation of process temporal behavior using clustered fuzzy digraphs. AIChE J., 47(4):906 – 919.

LI, W., YUE, H. H., VALLE-CERVANTES, S., and QIN, S. J., 2000. Recursive PCA for Adaptive Process Monitoring. J. Process Control, 10:471-486.

LIN, W., QIAN, Y., and LI, X., 2000. Non-linear dynamic principal component analysis for on-line process monitoring and diagnosis. Computers and Chemical Engineering, 24(2-7):423-429.

Li, X., 2002. A brief review: acoustic emission method for tool wear monitoring during turning. International Journal of Machine Tools & Manufacture, 42:157–165.

LUO, R., MISRA, M., and HIMMELBLAN, D. M., 1999. Sensor fault detection via multiscale analysis and dynamic PCA. Industrial and Engineering Chemistry Research, 38(4):1489-1495.

MACGREGOR, J. F., JAEKLE, C, KIPARISSIDES, C., and KOUTOUDI, C., 1994. Process monitoring and diagnosis by multiblock PLS methods. AIChE Journal, 40(5):826-838.

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

MAKI, Y., and LOPARO, K. A., 1997. A neural-network approach to fault detection and diagnosis in industrial processes. IEEE Trans. on Contr. Syst. Technology 5(6):529 – 541.

MALLAT, S. G., 1989. Multiresolution approximations and wavelet orthonormal bases, Transactions of the American Mathematical Society, 315(1):69–87.

MALLAT, S. G., and HWANG, W. L., 1992. Singularity detection and processing with wavelets. IEEE Transactions: On Information Theory, 38(2):617–643.

MARTINEZ, L., and MARTINEZ, A., 2004. Exploratory Data Analysis with Matlab, CRC Press , USA.

MAULUD, A. W. D., and ROMAGNOLI, J. A., 2006. A multi-scale orthogonal on linear strategy for multivariate statistical process monitoring. Journal of Process Control, 16(7):671-683.

MISRA, M, Y. H. H., QIN S J, and LING ,C., 2002. Multivariate process monitoring and fault diagnosis by multiscale PCA. Computers and Chemical Engineering 26(9):1281-1293.

MONTGOMERY, D. C., 1996. Introduction to Statistical Control, John Wiley and Sons, Inc.

NIIDA, K., 1985. Expert system experiments in processing engineering. In Institution of Chemical Engineering Symposium Series, pp. 529 – 583.

NOBUYUKI, K., TOMOYUKI, K., and HIROSHI, W., 1996. An estimating method of instantaneous load torque by Wavelet Transform for the induction motor drive. IEEE AMC-MIE, pp. 321-325.

NOMIKOS, P., and MACGREGOR, J. F., 1995. Multivariate SPC charts for Monitoring Batch Processes. Technometrics, 37(1): 41 – 59.

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

ODGAARD, P. F., and MATAJI, B., 2008. Observer-based fault detection and moisture estimating in coal mills. Control Engineering Practice, 16(8):909–921.

OYELEYE, O. O., and KRAMER, M. A., 1988. Qualitative simulation of chemical process systems: Steady state analysis. AICHE J., 34(9):1441 – 1454.

PAN, M. C., and SAS, P., 1996. Transient analysis on machinery condition monitoring. International Conference on Signal Processing Proceedings, 2:1723–1726.

PASTI, L., WALCZAK, B., MASSART, D. L., and RESCHIGLIAN P., 1999. Optimization of signal de-noising in discrete wavelet transform. Chemometrics and Intelligent Laboratory Systems, 48:21–34.

PATTON, R. J., and CHEN, J. 1997. Observer-based fault detection and isolation: robustness and applications. Control Engineering Practice 5(5):671–682.

PAUL, D., 1994. Detection of change in processes using Wavelets, Proceeding of IEEE-SP international symposium on time-frequency and time-scale analysis. pp. 174-177.

PEARSON, K., 1901. On lines and planes of closest fit to systems of points in space. Philosophical Magazine Series B, 2:559–572.

PENG, C., YUE, D., TIAN, E. G., et al., 2010. Observer-based fault detection for networked control systems with network Quality of Services. Applied mathematical modeling, 34(6):1653–1661.

PENG, C., YUE, D., TIAN, E. G., et al., 2010. Observer-based fault detection for networked control systems with network Quality of Services. Applied mathematical modeling, 34(6):1653–1661.

PENG, Z. K., and CHU, F. L., 2004. Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. Mechanical Systems and signal Processing, 18:199 – 221.

PERSIS D.C., and ISIDORI A., 2001. A geometric approach to nonlinear fault detection and isolation. IEEE Transactions on Automatic Control, 46(6):853–865.

PÖLLÄNEN, K., HÄKKINEN, A., REINIKAINEN, S., RANTANEN, J., and MINKKINEN, P., 2006. Dynamic PCA-based MSPC charts for nucleation prediction in batch cooling crystallization processes. Chemometrics and Intelligent Laboratory Systems, 84:126–133.

QIN, S. J., VALLE, S., and PIOVOSO, M. J., 2001. On unifying multiblock analysis with application to decentralized process monitoring. Journal of Chemometrics, 15(9):715-742.

QIU, H., LEE, J., LIN, J., and YU, G., 2006. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. Journal of Sound and vibration, 289:1066 – 1090.

RAfiEE, J., RAfiEE, M. A., PRAUSE, N., and SCHOEN, M. P., 2011. Wavelet basis functions in biomedical signal processing. Expert Systems with Applications, 38:6190–6201.

RENGANATHAN, K., and BHASKAR, V., 2010. Observer based on-line fault diagnosis of continuous systems modeled as Petri nets. ISA Transactions, doi:10.1016/j.isatra.2010.04.003.

RENGASWAMY, R., HAGGLUND, T., and VENKATASUBRAMANIAN, V., 2001. A qualitative shape analysis formalism for monitoring control loop performance. Engineering Applications of Artificial Intelligence, 14(1):23–33.

RENGASWAMY, R., and VENKATASUBRAMANIAN, V., 1995. A syntactic pattern-recognition approach for process monitoring and fault diagnosis. Engineering Applications of Artificial Intelligence, 8(1):35–51.

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

RICH, S. H., VENKATASUBRAMANIAN, V., NASRALLAH, M., and MATTEO, C., 1989. Development of a diagnostic expert system for a whipped toppings process. Journal of Loss Prevention in the Process Industries, 2(3):145–154.

ROSEN, C., and LENNOX, J. A., 2001. Multivariate and multiscale monitoring of waste water treatment operation. Water Res., 35(14):3402–3410.

ROSSI, T. M., and BRAUN, J. E., 1997. A statistical, rule-based fault detection and diagnostic method for vapor compression air conditioners. International Journal of Heating, Ventilating, Air Conditioning and Refrigerating Research, 3(1):19 – 37.

RUSSELL, E., CHIANG, L. H., and RICHARD, D. B., 2000. Data-driven techniques for fault detection and diagnosis in chemical process. London, Springer. 192 p.

SACKS, E., 1988. Qualitative analysis of piecewise linear approximation. J. of Artificial Intelligence in Eng., 3(3):151 – 155.

SHAO, R., JIA, F., MARTIN, E. B., and MORRIS, A. J., 1999. Wavelets and nonlinear principal component analysis for process monitoring. Control Engineering Practice, 7(7):865-879.

SHIELDS, D. N., 2005. Observer-based residual generation for fault diagnosis for non-affine non-linear polynomial systems. International Journal of Control, 78(5):363–384.

SHIOZAKI, J., MATSUYAMA, H., TANO, K., and OSHIMA, E., 1985. Fault diagnosis of chemical processes by the use of signed, directed graphs. Extension to five-range patterns of abnormality. Int. Chem. Eng., 25(4):651–659.

SHUN, L., 2009. A Model-Based Fault Detection and Diagnostic Methodology for Secondary HVAC Systems. Drexel University, Ph.D. Thesis.

TAFRESHI, R., 2005. Feature Extraction Using Wavelet Analysis with Application to Machine Fault Diagnosis. The University Of British Columbia, Ph.D Thesis.

TAN, S., and MAVROVOUNIOTIS, M. L., 1995. Reducing data dimensionality through optimizing neural network inputs. AIChE Journal, 41(6):1471–1480.

TARIFA, E., and SCENNA, N., 1997. Fault diagnosis, directed graphs, and fuzzy logic. Computers and Chemical Engineering, 21:649 – 654.

TEPPOLA, P., MUJUNEN, S., MINKKINRN, P., PUIJOLA, T., and PURSIHEIMO, P., 1998. Principal component analysis, contribution plots and feature weights in the monitoring of sequential data from a paper machine's wet end. Chemometrics and Intelligent Laboratory Systems, 44(1-2):307-317.

TREETRONG, J., SINHA, J. K., GUB, F., and BALL, A., 2009. Bi-spectrum of stator phase current for fault detection of induction motor. ISA Transactions, 48:378–382.

TSAI, C. S., and CHANG. C. T., 1995. Dynamic process diagnosis via integrated neural networks. Computers and Chem. Eng., 19:747 – 752.

TSUNG, F., 2000. Statistical monitoring and diagnosis of automatic controlled processes using dynamic PCA. International Journal of Production Research, 38(3):625–637.

ULERICH, N. H., and POWERS, G. A., 1988. Online hazard aversion and fault diagnosis in chemical processes: The digraph /fault tree method. IEEE Trans. on Reliability, 37(2):171–177.

UMEDA, T., KURIYAMA T., OSHIMA E., and MATSUYAMA H. 1980. A graphical approach to cause and effect analysis of chemical processing systems. Chem. Eng. Science, 35(12):2379 – 2388.

UNGAR, L. H., POWELL, B. A., and KAMENS, S. N., 1990. Adaptive networks for fault diagnosis and process control. Computers and Chem. Eng., 14(4-5):561 – 572.

VAIDHYANATHAN, R., VENKATASUBRAMANIAN, V., 1995. Digraph-based models for automated HAZOP analysis. Reliability Eng. and Syst. Safety, 50(1):33 – 49.

VALLE, S. L., W., and QIN. S. J., 1999 Selection of the number of principal components: The variance of the reconstruction error criterion with a comparison to other methods. Ind. Eng. Chem. Res., 38:4389 – 4401.

VEDAM, H., and VENKATASUBRAMANIAN, V., 1997a. Signed digraph based multiple fault diagnosis. Computers and Chem. Eng., 21:655 – 660.

VEDAM, H., and VENKATASUBRAMANIAN, V., 1997b. A wavelet theory-based adaptive trend analysis system for process monitoring and diagnosis. In American Control Conference, pp. 309 – 313.

VENKATASUBRAMANIAN, V., RENGASWAMY, R., YIN, K., and KAVURI, S.N., 2003a. A review of process faults detection and diagnosis. Part I: Quantitative model-based methods. Computers & Chemical Engineering, 27:293–311.

VENKATASUBRAMANIAN, V., RENGASWAMY R., YIN K., and KAVURI S.N., 2003b. A review of process faults detection and diagnosis. Part II: Qualitative models and search strategies. Computers and Chemical Engineering, 27:313–326.

VENKATASUBRAMANIAN V., RENGASWAMY R., YIN K., and KAVURI S.N., 2003c. A review of process faults detection and diagnosis. Part III: Process history based methods. Computers and Chemical Engineering, 27:327–346.

VILLEZ, K., SRINIVASAN, B., RENGASWAMY, R., NARASIMHAN, S., and VENKATASUBRAMANIAN, V., 2011. Kalman-based strategies for Fault Detection and Identification (FDI): Extensions and critical evaluation for a buffer tank system. Computers and Chemical Engineering, 35(5):806 – 816.

WANG, R., 2003. Statistical theory. China: Xian Jiaotong University Press.

WANG, S. W., and CHEN, Y. M., 2004, Sensor validation and reconstruction for building central chilling systems based on principal component analysis. Energy Conservation and Management, 45:673–695.

WANG, S., ZHOU, Q., and XIAO, F., 2010. A system-level fault detection and diagnosis strategy for HVAC systems involving sensor faults. Energy and Buildings, 42:477–490.

WANG, X. Z., CHEN, B. H., YANG, S. H., and MCGREAVY, C., 1999. Application of wavelets and neural networks to diagnostic system development. 2. An integrated framework and its application. Computers and Chem. Eng., 23(7):945 – 954.

WATANABE, K., MATSURA, I., ABE, M. KUBOTA, M., and HIMMELBLAU, D. M., 1989. Incipient fault diagnosis of chemical processes via artificial neural net-works. AICHE J., 35(11):1803 – 1812.

WESTERHUIS, J. A., KOURTI, T., and MACGREGOR, J. F., 1998. Analysis of multi-block and hierarchical PCA and PLS models. Journal of Chemometrics, 12:301–321.

WHITELEY, J. R., and DAVIS, J. F., 1994. A similarity-based approach to interpretation of sensor data using adaptive resonance theory. Computers and Chem. Eng., 18(7):637 – 661.

WILCOX, N. A., and HIMMELBLAU, D. M., 1994a. Possible cause and effect graphs (PCEG) model for fault diagnosis I: Methodology. Computers and Chem. Eng., 18(2):103 – 116.

WILCOX, N. A., and HIMMELBLAU, D. M., 1994b. Possible cause and effect graphs (PCEG) model for fault diagnosis II: Applications. Computers and Chem. Eng., 18(2):117 – 127.

WILLSKY, A. S., 1976. A Survey of Design Methods for Failure Detection in Dynamic Systems. Automatica, 12(6), 601–611.

WO, M., GUI, W, SHEN, D, and WANG, Y., 2000. Export fault diagnosis using role models with certainty factors for the leaching process. In Proceedings of the third world congress on intelligent control and automation. Hefei, China, 1:238 – 241.

WOLD, S., 1994. Exponentially weighted moving principal component analysis and projection to latent structures. Chemometrics Intelligent Laboratory Systems 23(1): 149-161.

WOLD, S., GELADI, P., ESBENSEN, K., and OHMAN, J., 1987. Multi-way principal components and PLS-analysis. Journal of Chemometrics, 1:41–56.

WOLD, S., KETTANE, H. N., and TJESSEM, K., 1996. Hierarchical multiblock PLS and PC model for easier model interpretation and as an alternative to variable selection. Journal of Chemometrics, 10(5-6):463 – 482.

WU, N., and WANG, Q., 2011. Experimental studies on damage detection of beam structures with wavelet transform. International Journal of Engineering Science, 49:253–261.

XIAO, F., WANG, S. W., and ZHANG, J. P., 2006 A diagnostic tool for online sensor health monitoring in air-conditioning systems. Automation in Construction, 15:489–503.

XIAO, F., WANG, S., XU, X., and GE, G., 2009. An isolation enhanced PCA method with expert-based multivariate decoupling for sensor FDD in air-conditioning systems. Applied Thermal Engineering, 29:712 – 722.

XINHUA, X., XIAO, F., and SHENGWEI, W., 2008. Enhanced chiller sensor fault detection, diagnosis and estimation using wavelet analysis and principal component analysis methods. Applied Thermal Engineering, 28:226–237.

YANG, Q., 2004. 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.

YOO, C., and KYOO, L., 2006. Nonlinear multivariate filtering and bioprocess monitoring for supervising nonlinear biological processes. Process Biochemistry, 41(8):1854 – 1863.

YOON, S., and MACGREGOR, J. F., 2000 Statistical and Causal Model-Based Approaches to Fault Detection and Isolation. AIChE Journal, 46(9):1813–1824.

YOUMING, C., and LILI, L., 2009. A fault detection technique for air-source heat pump water chiller/heaters. Energy and Buildings, 41:881–887.

YOUMING, C., and LILI, L., 2010. Fault detection, diagnosis and data recovery for a real building heating/cooling billing system. Energy Conversion and Management, 51:1015–1024.

YÜKSEL, T., and SEZGIN, A., 2010. Two fault detection and isolation schemes for robot manipulators using soft computing techniques. Applied Soft Computing, 10:125–134.

ZHANG, P., YE, H., DING, S. X., WANG, G. Z., and ZHOU, D. H., 2006. On the relationship between parity space and H-2 approaches to fault detection. Systems and Control Letters, 55(2):94 – 100.

ZHANG, X., POLYCARPOU, M. M., and PARISINI, T., 2002. A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems. IEEE Transactions on Automatic Control, 47(4):576–593.

ZHIQIANG, G., FURONG, G., and ZHIHUAN, S., 2011 Batch process monitoring based on support vector data description method. Journal of Process Control, 21(6):949 – 959.

ZHINQIANG, G., and QUNXIONG, Z., 2005. Multi scale Nonlinear Principal Component Analysis (NLPCA) and its Application for Chemical Process Monitoring. Industrial and Engineering Chemistry Research, 44(10):3585-3593.

ZHIQIANG, G., and ZHIHUAN, S., 2010. Semiconductor Manufacturing Process Monitoring Based on Adaptive Sub-statistical PCA. IEEE Transactions on Semiconductor Manufacturing, 23(1):99–108.

ZHONG, M., DING, Q., and SHI, P., 2009. Parity space-based fault detection for Markovian jump systems. International Journal of Systems Science, 40(4):421-428.

ZVOKELJ, M., ZUPAN, S., and PREBIL, I., 2010. Multivariate and multiscale monitoring of large-size low-speed bearings using Ensemble Empirical Mode Decomposition method combined with Principal Component Analysis. Mechanical Systems and Signal Processing, 24:1049–1067.

CURRICULUM VITAE

Alkan ALKAYA was born in Hatay on September 25, 1979. He received his B.Sc. degrees in Electrical and Electronics Engineering from Gaziantep University in 2001 and M.Sc. degrees in Electrical and Electronics Engineering from Mersin University in 2005. He has been working as a research assistant in Electrical and Electronics Engineering Department, Mersin University, since 2001. He began to a Ph.D. degree in Electrical and Electronics Engineering Department of the Çukurova University in 2006.

He has worked in several research projects supported by Scientific and Technological Council of Turkey (TBAG-2212 (102T091)) and Scientific Research Project Unit of Çukurova University (BAP-MMF2010D2). His research areas are System quality, Fault Detection and Diagnosis, System monitoring, Renewable energy and solar cell. He is a member of Turkish Chamber of Electrical Engineers.

APPENDIX

APPENDIX

GLOSSARY

The terminology in this field is not unique. This makes it difficult to understand the goals of contributions and to compare different approaches. For example, what, the differences between fault or failure detection, isolation, identification and diagnosis are, is not very clear. Hence, the SAFEPROCESS Technical Committee of IFAC (International Federation of Automatic Control) discussed this matter and tried to find commonly accepted definitions (Isermann and Ball, 1997).

About the states and the signals;

Fault: An unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable / usual / standard condition.

Failure: A permanent interruption of a system's ability to perform a required function under specified operating conditions.

False alarm: is an indication of a fault, when in actuality a fault has not occurred.

Malfunction: An intermittent irregularity in the fulfillment of a system's desired function.

Error: A deviation between a measured or computed value (of an output variable) and the true, specified or theoretically correct value.

Disturbance: An unknown and uncontrolled input acting on a system.

Missed detection: when there is no indication of a fault, though a fault has occurred.

Residual: A fault indicator, based on deviation between measurements and model-equation-based computations.

Symptom: A change of an observable quantity from normal behavior.

About the functions;

Fault detection: Determination of faults present in a system and the time of detection.

Fault isolation: Determination of the kind, location and time of detection of a fault. Follows fault detection.

Fault identification: Determination of the size and time-variant behavior of a fault. Follows fault isolation.

Monitoring: A continuous real time tasks of determining the conditions of a physical system, by recording information, recognizing and indicating anomalies in the behavior.

Residual computation: residual value is computed from the known variable.

Residual evaluation: the residual is evaluated in order to detect, isolate and identify faults.

Supervision: Monitoring a physical system and taking appropriate action to maintain the operation in the case of faults.

About the system properties and its measurements;

Reliability: Ability of a system to perform a required function under stated conditions, with in a given scope, during a given period of time.

Safety: Ability of a system not to cause danger to persons or equipment or the environment.

Availability: Probability that a system or equipment will operate satisfactorily and effectively at any point of time.

About the models;

Quantative model: Use of static and dynamic relations among system variables and parameters in order to describe a system's behavior in quantative mathematical terms.

Qualitative model: Use of static and dynamic relations among system variables and parameters in order to describe a system's behavior in qualitative terms such as causalities and if-then rules.

Diagnostic model: A set of static or dynamic relations which link specific input variables-the symptoms-to specific output variables-the faults.

Analytical redundancy: Use of two or more (but not necessary identical) ways to determine a variable, where one way uses a mathematical process model in analytical form.