CRANFIELD UNIVERSITY

Cristóbal Ruiz Cárcel

PREDICTIVE CONDITION MONITORING OF INDUSTRIAL
SYSTEMS FOR IMPROVED MAINTENANCE AND OPERATION

School of Engineering
Department of Power and Propulsion

Full Time Ph.D.
Academic Year: 2013 – 2014

Supervisors: Prof. David Mba
Dr. Yi Cao

July 2014

CRANFIELD UNIVERSITY

SCHOOL OF ENGINEERING
Department of Power and Propulsion

Full Time Ph.D.

Academic Year 2013 - 2014

CRISTÓBAL RUIZ CÁRCEL

PREDICTIVE CONDITION MONITORING OF INDUSTRIAL
SYSTEMS FOR IMPROVED MAINTENANCE AND OPERATION

Supervisors: Prof. David Mba
Dr. Yi Cao
July 2014

This thesis is submitted in partial fulfilment of the requirements for
the degree of PhD

ACKNOWLEDGEMENTS

ABSTRACT

Maintenance strategies based on condition monitoring of the different machines and devices in an industrial process can minimize downtime, increase the safety of plant operations and help in the process of decision-taking for control and maintenance actions in order to reduce maintenance and operating costs. Multivariate statistical methods are widely used for process condition monitoring in modern industrial sites due to the quantity of data available and the difficulties of building analytical models in complex facilities.

Nevertheless, the performance of these methodologies is still far away from being ideal, due to different issues such as process nonlinearities or varying operational conditions. In addition application of the latest approaches developed for process monitoring is not widely extended in real industry.

The aim of this investigation is to develop new and improve existing methodologies for predictive condition monitoring through the use of multivariate statistical methods. The research focuses on demonstrating the applicability of multivariate algorithms in real complex cases, the improvement of these methods in terms of fault detection and diagnosis by means of data fusion and the estimation of process performance degradation caused by faults.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

AE   Acoustic Emission

AIC   Akaike Information Criterion

ANC   Adaptive Noise Cancellation

CF   Crest Factor

CVA   Canonical Variate Analysis

DPCA   Dynamic Principal Component Analysis

DPLS   Dynamic Partial Least Squares

EU   European Union

FAPA   Fault and Prognostic Algorithms

GM   Gear Mesh Frequency

HHT   Hilbert-Huang Transform

IRD   Inner Race Defect Frequency

KDE   Kernel Density Estimators

LF   Line Frequency

LP   Linear Prediction

MVA   Multi Variate Analysis

ORD   Outer Race Defect Frequency

PCA   Principal Component Analysis

PLS   Partial Least Squares

SANC   Self-Adaptive Noise Cancellation

SCADA   Supervisory, control and data acquisition

SFFT   Short Fourier Transform

SK   Spectral Kurtosis

SNR   Signal to Noise Ratio

SS   Shaft Speed

SVD   Singular Value Decomposition

TSA   Time-domain Synchronous Averaging

UCL   Upper Control Limit

VSD   Variable Speed Drive

WT   Wavelet Transform

NOMENCLATURE

$\sum_{ff}$	Future covariance matrix
$\sum_{fp}$	Cross covariance matrix
$\sum_{pp}$	Past covariance matrix
A	State-space matrix
a	Reduced dimension in PCA
$a(k)$	Weight for sample k in LP
B	State-space matrix
c_α	Normal deviate of the (1-α) percentile
D	Virtual system damping
D	Matrix of canonical correlations
D	Bearing defect diameter
E	Residual matrix in PCA and CVA
e	Rotor unbalance eccentricity
$e(n)$	ANC and SANC filter output signal
f	Number of future samples considered in CVA
$F_0(t)$	Force acting on the system under normal operation
F_c	Central frequency in SK
$F_\alpha\,(a,\,n\text{-}a)$	F distribution with a and n-1 DOF
G	State-space matrix
H	Hankel matrix
H	State-space matrix
h	Bandwidth for the estimation of $p(y)$
H	Number of weighting coefficients in SANC
$h(t)$	Impulse response function
J	Transformation matrix in CVA
k	Arbitrary time point at CVA
K	Virtual system stiffness
$K(z)$	Kernel function
L	Transformation matrix in CVA
L	Length of bearing defect
L	Window length for extraction of features in vibration signal
M	Number of columns in past and future matrices in CVA
m	Number of sensors in PCA CVA
M	Number of samples in random variable y
m	Virtual system mass
m_r	Rotor mass
n	Number of samples in PCA and CVA
n_0	Additive background noise
n_1	Additional noise source uncorrelated with S in ANC
p	Number of past samples considered in CVA and LP
P	Loading matrix in PCA

$p(y)$	Probability density function of random variable y
Q	Q statistic (squared prediction error)
Q_{UCL}	Upper control limit for the Q statistic
r	Reduced dimension in CVA
R_τ	Autocorrelation function
S	Covariance Matrix in PCA
S	Signal of interest in SANC
T^2	Hotelling T^2 statistic
T^2_{UCL}	Upper control limit for the T^2 statistic
U	Rotation matrix in CVA
$u_0(t)$	Undamaged system motion
V	Rotation matrix in CVA
V	Loading vectors in PCA
V_r	Reduced rotation matrix in CVA
v_t	Background noise
W	Score matrix in PCA
w	SANC filter weights
W	Width of bearing defect
w_t	Background noise
X	Data matrix in PCA
x	Observation vector in PCA
$x(n)$	Vibration signal
$\hat{x}(n)$	Predictable part of $x(n)$
x_t	State vector
Y	Predicted matrix in FDA
y	Re-scaled principal components
Y_f	Future matrix in CVA
$y_{f,k}$	Future vector at CVA
$\hat{y}_{f,k}$	Normalized future vector
$\bar{y}_{f,k}$	Sample mean of $y_{f,k}$
Y_p	Past matrix in CVA
$y_{p,k}$	Past vector at CVA
$\hat{y}_{p,k}$	Normalized past vector
$\bar{y}_{p,k}$	Sample mean of $y_{p,k}$
y_t	Measurement vector
z	Canonical variates in CVA
α	Significance level
γ	Canonical correlations
Δ	SANC time delay
δ	Phase angle
Δf	Frequency band in SK
$\Delta F(t)$	Additional force caused by faults
$\Delta u(t)$	Residual motion induced by faults

ε	Residual
Λ	Diagonal matrix in PCA SVD
μ	SANC forgetting factor
μ	Average value of signal x in SK
$\rho_{fp}(a,b)$	Correlation between linear combinations $a^T(\hat{y}_{f,k})$
σ	Standard deviation of signal x in SK
Φ	State-space matrix
ω	Rotational speed

1 INTRODUCTION

Modern large scale industrial facilities such as oil separation plants and other chemical processes are complex systems composed of machines and devices of different nature. Such processes normally depend on a large number of variables that should be monitored and controlled to ensure the final quality of the product within safe, economical and environmentally friendly operating conditions. Condition monitoring of the different machines and devices in the process is important for deciding optimal production and maintenance strategies. Condition based maintenance can increase the safety of plant operations, minimize downtime and reduce operating and maintenance costs.

Breakdown maintenance and time-based maintenance strategies are still being used for plant maintenance, but condition based maintenance has been shown to be the most effective and economically profitable across most industries [1]. Breakdown maintenance consists basically of replacing a component when it finally fails. It is simple and does not require important initial investments but it has obvious disadvantages such as unplanned plant shutdowns or possible catastrophic damage in the facilities when the failures happen. Time based maintenance has the advantages of planned shutdowns and lower damage in the components. However, in some instances components are replaced unnecessarily, which means that component lifetime is not fully employed, causing an increase in the final operating costs.

Depending on the plant size, a large capital outlay is normally required to install an efficient and reliable condition monitoring system. This investment needs to be justified, and such justification can be drawn from the fact that a condition monitoring system will improve productivity and reliability, ultimately improving financial returns. The main advantages of condition based maintenance are:

- Plant shutdown planning: Pre order parts, reduced downtime grouping maintenance tasks.
- Reduction of labour costs: Minimize breakdowns and maintenance.
- Capacity to negotiate better insurance quotes for the plant.
- Unexpected shutdowns are reduced or eliminated
- Improved process performance: Gain extra production due to minimized downtime and operation according to process condition.
- Opportunity to improve and design a better plant
- Prevent secondary or extensive machine damage
- Safer plant operation

These advantages reiterate why condition based maintenance is a widely used strategy in industrial processes all over the world, especially when the use of electronics and information transmission systems are commonly extended [1].

Modern industrial facilities are heavily instrumented and automated; consequently there is a lot of process data available from the different sensors which can be used in detecting, diagnosing and predicting faults. Many methodologies have been developed to combine such data for analysis. One such approach is the use of multivariate analysis, which can take into account the relation between different variables measured, hence has an advantage in fault detection and diagnosis against the traditional univariate methodologies [2]. A large number of real industrial systems use multivariate algorithms to monitor operating conditions and performance. Nevertheless, there are still a lot of efforts concentrated on developing more effective techniques for fault detection and diagnosis, specially addressing challenges such as system dynamics and nonlinearity [3; 4].

The aim of this work is to develop and optimize new methodologies for predictive plant-wide condition monitoring. These methodologies will make use of statistical process monitoring of multi-variant data typically employed for monitoring process, electrical and mechanical machinery. The potential benefits of the application of such techniques include increased operational safety,

improved performance and reduced operating and maintenance costs of industrial processes.

1.1 The ENERGY-SMARTOPS project

The requirement of energy efficiency and reduction of CO_2 emissions is leading to new industrial processes and new ways of operating existing processes. In particular, the control and operation of the different interfaces of the system (process, mechanical and electrical) is becoming radically more integrated, giving new opportunities for energy saving through equipment management, automation, and optimization.

The 2006 Green Paper of the European Commission [5] identified the threats posed by security of energy supply and the need for sustainable, secure and affordable energy. The document highlighted the need for innovative technologies as well as diversification of energy supply and policies towards a single European strategy for energy and an internal energy market. The document also highlights that the EU needs to take measures to prevent energy supply crises developing, including better energy efficiency.

The ENERGY-SMARTOPS project starts from the premise that the savings in manufacturing industry must come from better operation of processes equipment and machinery that is already installed and running at the present time, which typically have an operative life of 30 to 50 years. In the light of these challenges, there is a need for new training and research action to address technology gaps at the interfaces between the process, mechanical and electrical domains, and to realize energy savings from integrated operation. The overall scientific and technical aim of the ENERGY-SMARTOPS project is to take a pivotal role in demonstrations of creative ideas for energy savings in large scale industrial sites making the best possible use of measurements from all plant subsystems. The project integrates in-depth understanding of the operational issues with analysis of measurements and first-principles physical knowledge to invent and develop tools that will be deployed in the field in case studies with the transmission operator partners. Three specific research areas are:

- Equipment and process monitoring; integrating multiple measurements from the process, mechanical and electrical sub-systems.
- Integrated automation; capturing information from all the subsystems and devising new algorithms that explicitly manage the interfaces and interactions between them.
- Optimization; to provide energy savings by better integration of operations across the process-mechanical-electrical interfaces.

This is a four years project financed by the Seventh Framework Program (FP7) of the European Commission which started in 2011. The Energy-Smartops consortium has detailed plans for cross-disciplinary training of a cohort of early stage engineering researchers through several academic and industrial participants. The project consortium includes universities like the London Imperial College, Cranfield University, ETH Zurich, Technical University of Krakow or Carnegie Mellon University and private partners like ABB, BASF, ThyssenKrup Acciai Speciali Terni, ESD Training Simulation Ltd and Statoil.

The present work covers only one of the 19 tasks that will be developed by the different project participants, and is included in the equipment and process monitoring research area. The aim of this work is to develop and improve condition monitoring techniques that can provide useful information about the process condition. This information can be used afterwards to optimize the operation of industrial processes from the maintenance point of view. Fig. 1 summarizes the concept of condition based maintenance for processes, where the data acquired from the different system interfaces (process, mechanical and electrical) is used to detect and diagnose faults, as well as to estimate the impact of these faults in the system performance. The information provided by the condition monitoring system is then used to produce optimal production and maintenance schedules. This optimization problem takes into account the different costs associated with maintenance actions (process stop and restart, wasted raw materials, downtime, energy consumption, spare parts, labour, etc.) to generate a schedule than satisfies the demand maximizing the economical revenue. All this information about process condition and optimal scheduling is
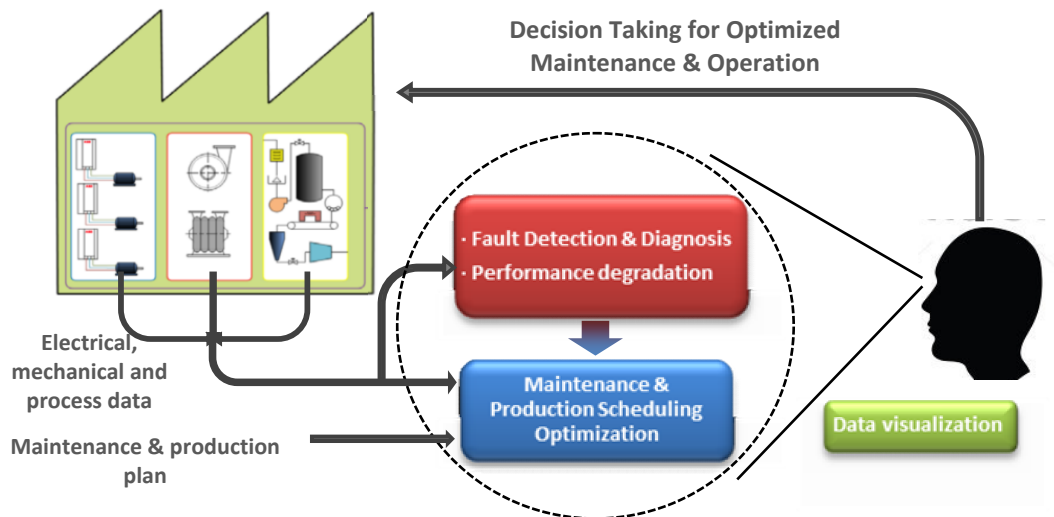
Fig. 1: Condition based maintenance scheme

visualized by the system operator, who takes the appropriate decisions to actuate on the system.

The work presented in this thesis focuses exclusively on the condition monitoring part, providing tools for fault detection and diagnosis and estimation of performance degradation. Research involving scheduling optimization and complex data visualization was carried out by others and is not part of this thesis.

1.2 Objectives

The overall goal of the work presented in this thesis is to develop and optimize new methodologies for predictive plant-wide condition monitoring through multivariate statistical analysis. In order to improve the efficiency of the process it is necessary in the first instance to ensure that the process equipment is working smoothly and in good conditions. The implementation of global actions on the process like advanced control strategies or optimized workflow schemes to minimize the energy consumption only makes sense if the system condition is known and the performance of the process is taken into account. In the event of a fault, the early detection and diagnosis of the fault is crucial for the operation and maintenance of the system. The information provided by the condition monitoring system about the fault nature and its

25

impact on the system performance can be used to optimize the scheduling of maintenance and production accordingly.

The planed tasks undertaken to achieve the objectives of this study are:

1) Review current algorithms employed for multivariate analysis (MVA) and diagnosis.
2) Assess current multivariate fault diagnosis algorithms using simulated and real data.
3) Develop improved fault and prognostic algorithms based on literature.
4) Undertake experimental simulation studies on process, electrical and mechanical machines.
5) Validate and optimize fault and prognostic algorithms using the experimental data collected.

In order to fulfil the requirements of the ENERGY-SMARTOPS project this research will focus on faults that develop over time affecting the overall system performance rather than critical faults that require immediate maintenance action after the fault detection to avoid catastrophic failures. If the detected fault is not critical and allows the operation of the system under suboptimal conditions for a certain period, it is possible to optimally reschedule the maintenance and operation of the plant according to the actual system condition. The goal of this re-scheduling process is to exploit plant layout flexibility and group maintenance tasks to minimize maintenance and production costs while meeting the demand requirements, but the solution of this optimization problem is out of the scope of this work.

The state of the art in the field of process condition based maintenance and the potential improvements identified will be presented in chapter 0. Based on this review the main research targets were identified, which included:

1) Development of an experimental case study to assess currently used multivariate algorithms for process monitoring and prove the applicability of these techniques in real industrial environments. This case study must include multivariate data acquired from a large-scale complex facility, characterized by non-linear behaviour and dynamically changing operational conditions.

2) Undertake a comparative study of the capabilities of different monitoring algorithms in terms of fault detection and diagnosis using real process data.

3) Development of prognostic algorithms based on reviewed literature that can estimate the impact of faults over the system performance and predict the future behaviour of a faulty system.

4) Integration of different types of data to develop improved condition monitoring tools for earlier and more effective fault detection and diagnosis

The research work to address these four points is presented in this thesis in chapters 3 to 8.

1.3 Thesis structure

Following the introduction and the literature review, the thesis is structured in seven further chapters. Fig. 2 shows schematically the concepts investigated in each chapter, as well as the interconnections between the different chapters.

Each one of the chapters from 3 to 8 presents an independent research work, containing the common subsections that are normally found in scientific publications: introduction, methodology, results and discussion, and conclusion. Chapter 3 describes the experimental work carried out in a large scale experimental facility for investigation of multiphase flow. In this work multivariate data was acquired from the system working under varying operational conditions after seeding different faults. The objective is to generate a case

```
┌─────────────────────────────────────────────────────────────────────┐
│                       Chapter 1: Introduction                        │
└─────────────────────────────────────────────────────────────────────┘

┌─────────────────────────────────────────────────────────────────────┐
│                     Chapter 2: Literature review                     │
└─────────────────────────────────────────────────────────────────────┘
            ⇩                      ⇩                      ⇩
┌──────────────────────┐ ┌──────────────────────┐ ┌──────────────────────┐
│ Chapter 3: Development│ │ Chapter 6: Comparative│ │ Chapter 7: Application│
│ of experimental case  │ │ study of different    │ │ of Spectral Kurtosis  │
│ study                 │ │ signal processing     │→│ on acoustic emission  │
│                       │ │ techniques for        │ │ signals               │
│                       │ │ vibration-based       │ │                       │
│                       │ │ monitoring            │ │                       │
└──────────────────────┘ └──────────────────────┘ └──────────────────────┘
            ⇩                      ⇩
┌──────────────────────┐ ┌──────────────────────┐
│ Chapter 4: Assessment │ │ Chapter 8: Combination│
│ of CVA for fault      │ │ of process and        │
│ detection & diagnosis,│→│ vibration data for    │
│ comparative study     │ │ improved process      │
│                       │ │ condition monitoring  │
└──────────────────────┘ └──────────────────────┘
            ⇩
┌──────────────────────┐
│ Chapter 5: Estimation │
│ of process performance│
│ degradation under     │
│ faulty conditions     │
│ using CVA             │
└──────────────────────┘

┌─────────────────────────────────────────────────────────────────────┐
│                       Chapter 9: Conclusion                          │
└─────────────────────────────────────────────────────────────────────┘
```

Fig. 2: Schematic representation of the thesis structure

study to assess the capabilities of different multivariate algorithms for condition monitoring in terms of fault detection and diagnosis using real data. This case study is designed to be challenging for the algorithms containing data from different types of sensors, non-linear relations and varying operational conditions.

In chapter 4 the data sets described in chapter 3 were used to assess the capabilities of Canonical Variate Analysis (CVA) in terms of fault detection and diagnosis. This method has already been used for process monitoring using simulated data, reporting a better performance than other traditional algorithms such as the Principal Component Analysis (PCA) and the Partial Least Squares (PLS). Fault detection and diagnosis using CVA is analysed in detail in this chapter, and its performance is compared with PCA and PLS.

Following the application of CVA, chapter 5 presents an investigation on the use of this method for performance degradation analysis using real data acquired from the same test rig described in chapter 3. The goal of this research is to evaluate the capabilities of CVA to generate a model of a real complex process that can be used to estimate the degradation on the system performance when working under faulty conditions, as well as predicting the future behaviour of the faulty system under different operational conditions. After the event of fault detection and the isolation of the fault origin, this method can provide valuable information about the performance of the faulty system that can be used to optimize the production and maintenance plans according to the actual system condition.

Independently of the content in chapters 3 to 5, chapter 6 evaluates the performance of three different methods to enhance bearing fault features in the vibration signal spectrum. Vibration-based condition monitoring is probably the most common method for detection and diagnosis of mechanical faults in rotating machinery. The objective of this chapter is to explore vibration signal processing techniques in order to obtain a better understanding of vibration-based monitoring and signal processing, so as to assess the potential benefits of merging process and vibration data that will be investigated in chapter 8. Similarly chapter 7 presents the results of an investigation about the application of spectral Kurtosis on Acoustic Emission (AE) signals to detect seeded bearing faults of different size. The objective again is to explore the capabilities of condition monitoring based on AE, which is gaining popularity against vibration-based monitoring in the last years.

Chapter 8 is dedicated to the investigation of the benefits that can be obtained from merging different types of data for condition monitoring in a study carried out using real data obtained from a compressor test rig. The objective of this investigation is to improve the detection and diagnosis performance of currently used techniques for condition monitoring, trying to combine the benefits of process monitoring techniques and vibration-based monitoring. The research is centred on the extraction of specific features from the vibration signal that can

be combined with process data to improve the detection of mechanical faults in systems working under varying operational conditions.

Finally, chapter 9 summarizes how the objectives stated in section 1.2 were fulfilled and how the whole investigation presented in this thesis contributes to the objectives of the ENERGY-SMARTOPS project. This chapter also includes a list of the journal and conference papers generated as the result of the investigations carried out, which proves the novelty of the work developed in this thesis and the contributions to knowledge.

2 LITERATURE REVIEW

Once the main objectives of this research have been stated, this chapter attempts to summarize the state of the art and the main challenges to be faced in the different fields covered. The objective of this literature review is not to do an intensive research about the history and recent advances in each one of the areas treated but to give an overview of the actual development of these areas and the main research opportunities. Due to the variety of methodologies used for this study a detailed review of the origins and latest developments in each one of the research areas covered will be presented in the corresponding chapters of the thesis.

Condition based maintenance is a strategy where the maintenance actions are undertaken based on the information provided by condition monitoring systems. The aim of this strategy is to reduce operating and maintenance costs in industrial systems by minimizing the amount of maintenance carried out and avoiding major breakdowns by means of preventive maintenance.

The first maintenance strategy used was breakdown maintenance (also known as run-to-failure maintenance) which consists basically in undertaking maintenance actions only when a failure happens. It requires no planning effort or initial investments, but obviously it has several disadvantages such as unplanned shutdowns, extensive damage to other system parts, longer reparation times due to lack of planning, low operational safety, etc. Time based maintenance was introduced later to avoid these problems, undertaking preventive maintenance actions at certain time intervals in an attempt to avoid failures in the system. The main drawback of this maintenance strategy comes from the fact that maintenance actions are undertaken regardless of the actual condition of the assets, which increases the maintenance costs if parts are replaced or repaired before the end of their actual useful life. The complexity of modern industrial sites and the increased expectations in terms of reliability and product quality made these costs even higher. In condition based maintenance, the maintenance actions are planned when the condition monitoring system

shows evidences of damage, reducing unnecessary scheduled preventive maintenance.

The economic benefit of condition based maintenance against other strategies has been studied by several researchers. Recently Zhang [6] reported significant savings applying condition based maintenance in medicine dispensing stations. Rajan et al. [7] also reported condition-based maintenance as the most cost effective maintenance strategy in a case study using real data from a batch process in the pharmaceutical industry. Different studies have been developed recently to model the application costs of condition based maintenance strategies [8-13], but the cost analysis is out of the scope of this investigation. The procedures for the successful implementation of cost effective condition based maintenance were explored by Basim [14].

In the year 2000 Chiang et al. [2] published a book which compiles and analyses the most commonly used methodologies for process monitoring. Basically, the authors categorized these methods in three groups: data based methods, analytical methods and knowledge based methods. Analytical methods (or model-based methods) are the most traditional of all of them [4]. These approaches are based on the construction of mathematical models of the system using first principle equations. Faults are detected and diagnosed by looking at differences between process measurements and estimations produced by the model. The results obtained are more accurate than the other two methods as long as the model is reliable [4], however due to the complexity of modern industrial facilities it is complicated (sometimes impossible) to build reliable models. Knowledge based models rely on the knowledge about the process behaviour and the experience of the operators to apply techniques like causal analysis, machine learning, pattern recognition or Fuzzy logic, which is a time consuming and difficult procedure. Finally data-based methods are derived directly from process data, with no need of physical understanding of the process or expert knowledge. Due to the complexity and the high degree of instrumentation in modern industrial facilities this kind of methods became more popular in the recent years. The main drawback of data-based methods is the

amount of data required to build the model, but in most processes nowadays there is plenty of historical data stored that can be used for monitoring purposes.

Given the objective of this research is to develop plant-wide condition monitoring tools for large industrial sites where the impact of the process plant performance on revenue is high, and due to the benefits of data-based approaches, the investigation will focus on this type of methods. Multivariate algorithms such as the PCA, PLS and CVA can account for the correlation between the different variables measured in the process, and show advantages against the traditional univariate methods [2]. These techniques allow the conversion of the m-dimensional data acquired from the process into a single health indicator that provides information about the process condition.

PCA is probably the most widely used method for process monitoring. It is a linear dimensionality reduction technique which is optimal in terms of capturing the variability of the data. PLS is another dimensionality reduction technique that maximizes the covariance between a predictor (independent) data set X and the predicted (dependent) Y for each component in the reduced space. Finally CVA aims to find the linear combinations that maximize the correlation between two sets of variables. The main peculiarity of this method is that it is able to take into account time correlations due to the way in which the acquired data is structured before the analysis, making it more suitable for dynamic monitoring. Ku et al. [15] proposed the use of lagged variables to take into account time correlation to extend PCA to dynamic system monitoring (DPCA). Similarly, a dynamic version of PLS called dynamic PLS (DPLS) was proposed by Komulainen [16] . Despite of their success, DPCA and DPLS have been reported not to be as efficient as other state-space based methodologies such as CVA when applied to systems working under variable loading conditions, principally due to the representation of the system dynamics [17], [18; 19].

The literature gives examples of successful application of PCA, PLS and CVA for fault detection and diagnosis using computer simulated data [2; 17-30]. The Tennessee Eastman process simulator [31] has been widely used for the

assessment and comparison of the performance of different algorithms due to its realistic level of complexity and the challenges attached to the fact that it is a highly non-linear system. The popularity of this particular benchmark case is demonstrated by the high number of researchers that have used it in the last years to prove the validity of a large variety of approaches [32-50].

Despite the superior performance of CVA against other methods when applied to computer simulated data, examples of application using real industrial data are almost anecdotal. These examples are restricted to data acquired in small test rigs [51] or particular parts of a system [52] but there are no examples of application of CVA in real and complex systems working under varying operational conditions. That is why the first objective of this thesis is the development of an experimental case study that allows the assessment of the performance of different algorithms for fault detection and diagnosis.

A review about the research on data based process monitoring was presented recently by Ge et al. [4]. The authors stated that non-Gaussian and non-linear systems, time variance, dynamic monitoring and batch process monitoring are the most active fields of research in this area. The development of tools for monitoring complex dynamic processes, plant wide monitoring and multidata fusion are some of the most promising issues. The capabilities of CVA seem to match perfectly the type of solutions required by the industry, and consequently this method in particular will be explored in detail in this investigation.

Vibration-based condition monitoring is probably the most common method for detection and diagnosis of mechanical faults in rotating machinery and it has several advantages against other methods [1]. Typically, analysis of the vibration frequency spectrum can point directly to the source of the fault and there are plenty of signal processing techniques available to help the user to undertake diagnosis in conditions of high background noise. Assuming that the initial status of the machine was healthy, any changes observed in the measured vibration response are caused by the deterioration of the machine condition. However, this assumption is only valid if all the measurements are taken under the same loading conditions, as different levels of load will

generate different vibration levels [53]. One of the most active research areas in vibration based monitoring is the development of tools that can monitor the condition of machines working under varying loading conditions [54-60]. Due to the potential benefits of data fusion and the obvious advantages of vibration based monitoring, another key point of this investigation will be the combination of process data and key features obtained from vibration measurements to provide a more robust and reliable condition monitoring tool.

In addition to the detection and diagnosis of faults in the process, prognosis is a key aspect in the optimization of operation and maintenance schedules through condition based maintenance. The concept of prognosis consist in the forecast of remaining useful life, future condition or probability of reliable operation of the equipment based on the information provided by condition monitoring [61]. This forecast is used to minimize operation and maintenance costs by grouping maintenance tasks, pre-ordering parts, planning labour needed, modifying the operation strategy taking advantage of the plant flexibility, etc. The quantity of publications listed in recent reviews about the prognosis state of the art [61; 62] evidences how active is this field of research. Some of the issues enumerated in these reviews in the field of prognosis are the effect of varying operational conditions, non-linear relationships, practicability of the methods developed, prognosis performance evaluation and efficient on-line signal processing algorithms.

In the light of these challenges, this investigation will explore and test the capabilities of CVA to efficiently build a mathematical model of the system that allows the estimation of the process performance once it has been affected by a fault. In addition, the prediction of the future process performance working under faulty conditions will be investigated, with the objective of providing reliable estimations of the system behaviour for different operational conditions. This prediction can be used as an input for the calculation of remaining useful life of individual components. These approaches could then be used to take into account performance degradation and faulty system behaviour for the optimal

planning of maintenance and operation schedules, improving the overall efficiency and economic revenue of industrial processes.

3 STATISTICAL PROCESS MONITORING OF A MULTIPHASE FLOW FACILITY, PART I: THE BENCHMARK CASE

Abstract

The early detection and diagnosis of faults can reduce the operating costs of industrial processes by avoiding the inefficient operation of faulty equipment as well as minimizing unplanned shutdowns, downtime and extensive damage to other parts of the system.

The availability of process data in modern industrial facilities due to the high degree of instrumentation and automation has increased the popularity of data-driven methods for process monitoring. Today significant effort is placed on the development of improved methodologies for the detection and diagnosis of faults based on the manipulation of process data. These new techniques are usually tested using computer simulated data, typically using benchmark case studies such as the Tennessee Eastman Process Plant [31]. On the other hand, although more and more application case studies of statistical process monitoring have been reported in the literature, process data of these applications are generally not available in public domain due to commercial confidentiality.

In order to bridge the gap, this work aims to provide a benchmark case to demonstrate the ability of different monitoring techniques to detect and diagnose artificially seeded faults in a large scale test rig and measure the impact of those faults on the system performance. The objective of this case study is to provide a test bed for the development of new monitoring techniques and its implementation in real industrial facilities. This chapter describes in detail the experimental test rig and associated data for the different fault cases studied, including an example of analysis employing the PCA.

Keywords: fault detection; diagnosis; process monitoring; multivariate; dynamic; experimental

3.1 Introduction

Modern large scale industrial facilities are complex systems with various types of machines and devices. The process normally depends on a large number of variables that are monitored and controlled to ensure the final quality of the product and the process itself within safe, economical, and environmentally friendly operating conditions.

Condition monitoring of the different machines and devices in the process is important as the early and accurate detection and diagnosis of faults minimize downtime, increase the safety of plant operations and help with decision-making for optimal control and maintenance strategies. The latter reduces maintenance and operating costs through the improvement of the process efficiency, avoiding the operation of faulty and inefficient equipment. There is also a cost associated with unplanned shutdowns as a consequence of the energy consumed during the shutdown and restart process. These are clear examples of how fault detection and diagnosis can contribute to the reduction of maintenance costs and the improvement of plant safety and availability.

Industrial facilities are heavily instrumented and automated; consequently significant amount of process data is available from the different sensors which can be used for detecting and diagnosing faults. Data driven methods are widely used for fault detection and diagnostic applications in real industrial systems. Its success is dependent on the availability of measured data and the difficulties associated to the development of reliable models based on first principles equations for large and complex process plants [2].

Multivariate monitoring techniques take into account the relationship between the different variables measured in the process and they have demonstrated to be an advantage against the traditional univariate methodologies [2]. Despite the popularity of multivariate algorithms for condition monitoring of industrial processes, it is still necessary to develop more effective techniques that can deal with problems like changing operational conditions or nonlinear systems [3]. The literature gives examples of successful application of different improved algorithms based on well-known techniques such as PCA, PLS and CVA for

fault detection and diagnosis using computer simulated data [23-30]. It is especially interesting to use data obtained from a common benchmark case study, to compare the performance of different methodologies. The Tennessee Eastman process simulator [31] has been widely used for the assessment and comparison of the performance of different algorithms due to its realistic level of complexity and the challenges attached to the fact that it is a highly non-linear system. The popularity of this particular benchmark case is demonstrated by the high number of researchers that have used it in the last years to prove the validity of a large variety of approaches [32-50].

In addition to the use of computer simulated data, it is also possible to find numerous examples of process monitoring techniques that have been tested using real process data obtained from industrial facilities or experimental test rigs. These data sets have been acquired from different types of processes such as steel casting [63], ethylene compressors [64], blast furnaces [65], sequencing batch reactors [66] and many others [67-74]. The application of novel techniques using real data can prove the applicability of these methods in real industrial applications but in the other hand, the diversity in the cases studied makes it impossible for other researchers to compare the effectiveness of different techniques under the same test conditions. This underlines the authors' motivation for generating a common benchmark case study which can be used for the assessment of novel process monitoring techniques using real data available in the public domain.

This study shows the experimental work carried out in the Three-phase Flow Facility at Cranfield University, which is comparable in size and complexity to a real small-scale multiphase flow separation process. Six different faults were introduced deliberately in the system, simulating typical faults that may be experienced in real plants such as blockages, incorrect system operation or non-conventional operating conditions. The data sets were acquired under changing operational conditions to ensure that the fault detection can be undertaken not only in the steady-state regime but also under varying operational conditions, making it valid for the assessment of dynamic process

monitoring techniques. Additional data sets were acquired during steady conditions for comparative purposes. The changing operational conditions and the non-linear nature of the multiphase flow process, together with the size and complexity of the test rig makes this case study an ideal candidate for a benchmark case that can be used for the evaluation of novel multivariate process monitoring techniques performance using real experimental data.

The outcome of this work is organised in two parts. This first part is included in this chapter and describes in section 3.2 the test rig itself and how the data sets were acquired. Section 3.3 describes the software tool developed for the analysis of the data sets. In section 3.4 an example of application of the case study for the detection of faults by applying the PCA is presented, then the work is concluded in section 3.5. The second part of the investigation is presented in chapter 4 and will show the results obtained after the application of CVA and other algorithms to the data acquired in the laboratory.

3.2 Experimental set up

3.2.1 Description of the Three-phase Flow Facility

The Three-phase Flow Facility at Cranfield University is designed to provide a controlled and measured flow rate of water, oil and air to a pressurized system. Fig. 3 shows a simplified sketch of the facility. The test area consists of pipelines with different bore sizes and geometries, and a gas and liquid two-phase separator (0.5m diameter and 1.2m high) at the top of a 10.5 meters high platform. It can be supplied with single phase of air, water and oil, or a mixture of those fluids, at required rates. Finally the fluid mixtures are separated in a 11 m^3 horizontal three-phase separator at ground level (GS500). The air is returned to the atmosphere and the emulsions of oil and water are separated in their respective coalescers (CW500 and CO500), both having a capacity of 1.5 m^3 approximately, before returning to their respective storage tanks (T200 and T100). The capacity of each storage tank is approximately 12.5 m^3.

Air is supplied by a union of two compressors which are capable to deliver a flow rate up to 1410 m^3/h at 7 barg. The compressed air is received in a 8 m^3
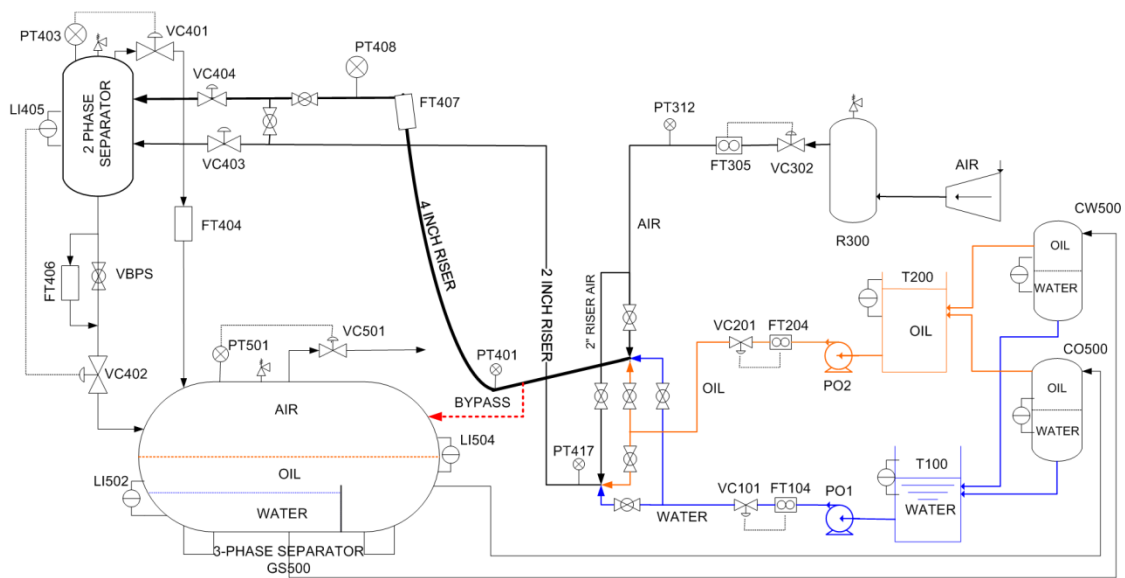
Fig. 3: Sketch of the three-phase flow facility

vessel (R300) to dampen pressure fluctuations. Then is filtered of droplets and particles and cooled before passing through the air flow meter FT305 and the pneumatic valve VC302 that controls the air flow. The water and oil are stored in tanks T100 and T200 and then supplied independently through multistage Grundfos CR90-5 pumps (PO1 and PO2), each one of which can provide up to 100 m^3/h at 10 barg with their rotational speed controlled by variable frequency inverters. The water flow rate is measured by FT104 and oil flow rate by FT204. The water and oil flow rates are controlled by pneumatic valves VC101 and VC201, respectively.

After the mixture, the fluids can flow either through a 4" diameter flow loop which has a 55m long and 2° downward inclined pipeline leading to a 10.5m high catenary riser, or via a 2" flow loop which is a 40m long horizontal pipeline, connecting to a 10.5m long vertical riser. Both flowlines are connected to the two phase separator but can be isolated each other by manual valve manifolds in both ends of the flowlines. During all the experiments adopted in this benchmark case the 4" line was used exclusively, except for the study of a particular fault (Case 6), where the 2" line was also involved. In the rig there is an alternative 4" line which can carry the flow directly from the mixing point to

41

the 3 phase separator, bypassing the riser and the top separator. This line is labelled as "BYPASS" and marked in dashed red in Fig. 3.

There are sensors measuring pressure at the air supply line before the mixing point (PT312), and along the 4" line at the bottom of the riser (PT401), at the top of the riser (PT408), inside of the top separator (PT403) and inside of the three-phase separator (PT501). The flow rate in the 4" line at the top of the riser can be measured by FT407, which also provides measurements of the density and the temperature. FT406 provides the measurements of the mass flow rate, density and temperature at the bottom outlet of the top separator. The liquid level inside of the top separator is measured by LI405 and controlled by VC402. The pressure of the 3-phase separator is controlled by VC501, while the levels between the different phases are measured by LI502 and LI504 respectively. In addition the pressure in the 2" line at the bottom of the riser is measured by PT417. The valves and sensors which control the level between phases in the coalescers have not been shown in Fig. 3 for simplicity. The whole system is managed using Delta V [75], a Fieldbus based supervisory, control and data acquisition (SCADA) software supplied by Emerson Process Management. Time stamped data of different variables can be retrieved, processed and visualized. The data can also be saved for post-processing. Fig. 4 and Fig. 5 show the test rig platform and the control room respectively.

Fig. 4: Test rig platform

Fig. 5: Control room

In this study all the data was captured at a sampling rate of 1 Hz. The variables used include 24 different process variables (see Table 1) and two process inputs (air and water flow rate set point). Only air and water were used in all the experiments and the three phase separator was always pressurized to 1.0 barg.

Table 1: List of process variables used in this study

Variable nr	Location	Measured Magnitude	Unit
1	PT312	Air delivery pressure	barg
2	PT401	Pressure in the bottom of the riser	barg
3	PT408	Pressure in top of the riser	barg
4	PT403	Pressure in top separator	barg
5	PT501	Pressure in 3 phase separator	barg
6	PT408	Diff. pressure (PT401-PT408)	barg
7	PT403	Diff. pressure(PT408-PT403)	mbarg
8	FT305	Flow rate input air	Sm^3/h
9	FT104	Flow rate input water	kg/s
10	FT407	Flow rate top riser	kg/s
11	LI405	Level top separator	m
12	FT406	Flow rate top separator output	kg/s
13	FT407	Density top riser	kg/m^3
14	FT406	Density top separator output	kg/m^3
15	FT104	Density water input	kg/m^3
16	FT407	Temperature top riser	°C
17	FT406	Temperature top separator output	°C
18	FT104	Temperature water input	°C
19	LI504	Level gas-liquid 3 phase separator	%
20	VC501	Position of valve VC501	%
21	VC302	Position of valve VC302	%
22	VC101	Position of valve VC101	%
23	PO1	Water pump current	A
24	PT417	Pressure in mixture zone 2" line	barg

3.2.2 Normal Operation

In order to obtain data representing normal operating conditions three data sets (T1, T2 and T3) were acquired from the system. The set points of air and water flow rates were deliberately varied during the tests in order to obtain data from the process working under variable operating conditions. To ensure that the conditions during normal operation were representative 20 different combinations of air and water flow rates (see Table 2) were tested for each one of the three training data sets. In each one of them, the flow conditions were changed, though not identically, in order to obtain a good variety of large, small, long and short process changes happening in different directions (increment or decrement). The objective of this variety in the operational conditions is to ensure that the dynamics of the system are captured in all circumstances.

Fig. 6 represents the air and water flow rate for each data set.

Table 2: Typical set point values for air and water flow rates

Air flow rate (m³/h)	75	100	125	150	
Water flow rate (kg/s)	0.5	1	2	3.5	6
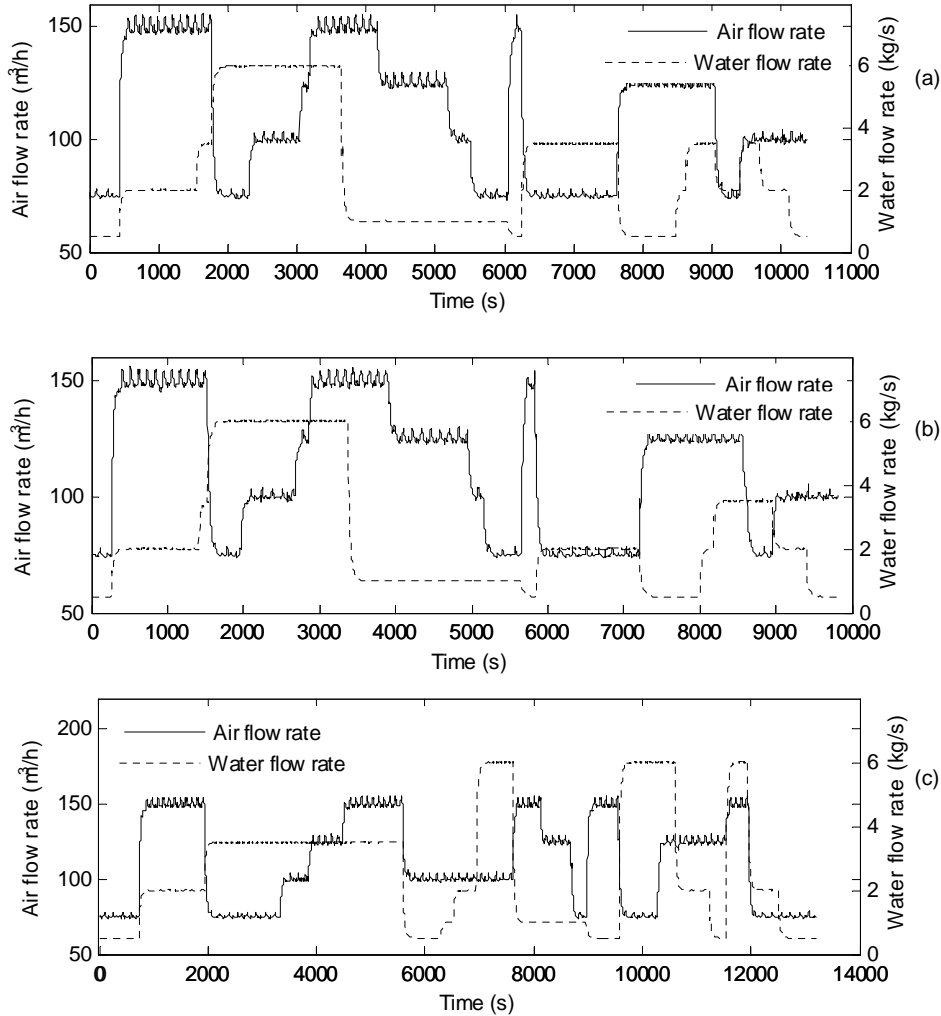
Fig. 6: Operational conditions for training data sets T1 (a), T2 (b) and T3 (c)

3.2.3 Operation with seeded faults

In addition to the training data sets, different sets of data were acquired from the system working in faulty conditions after seeding six different faults into the system. These faults were used in order to simulate typical malfunctions that could be experienced in a real system such as blockages in the pipelines, wrong system operation or abnormal operating conditions. The faults were

44

introduced after a certain time of normal operation in order to investigate changes in the health indicators generated by monitoring algorithms. These faults were introduced gradually when possible, in order to observe how the severity of the fault affects the indicator. After reaching a certain level of severity the fault condition was removed, returning the system to normal conditions. During these tests the flow rate conditions were changed in a similar way as was undertaken for normal operation. There are additional data sets acquired under constant flow rate conditions which allow the observation of the fault effects on the different measured variables without disturbances created by changes in the air or water flow rates. Table 3 summarizes the different faults tested and the next subsections explain in detail how each one of the faults were introduced.

Table 3: Summary of faults introduced

Case	Description	Type	Nr. of data sets (changing conditions)	Nr. of data sets (steady conditions)
1	Air line blockage	Gradual	1	2
2	Water line blockage	Gradual	1	2
3	Top separator input blockage	Gradual	1	3
4	Open direct bypass	Gradual	1	2
5	Slugging conditions	Random	2	0
6	Pressurization of the 2" line	Step	2	0

Case 1: Air line blockage

Just before the air, water and oil supply lines are joined at the mixing point, there are individual manual valves for each one of the lines. The fault introduced in this case was to gradually close the manual valve of the air line simulating an air blockage that develops over time. The valve angle was measured using a protractor attached to the valve. At the beginning of the experiments the valve was fully open (normal conditions) and the valve was gradually closed to introduce the fault condition simulating the blockage. Fig. 7 shows the manipulated manual valve:
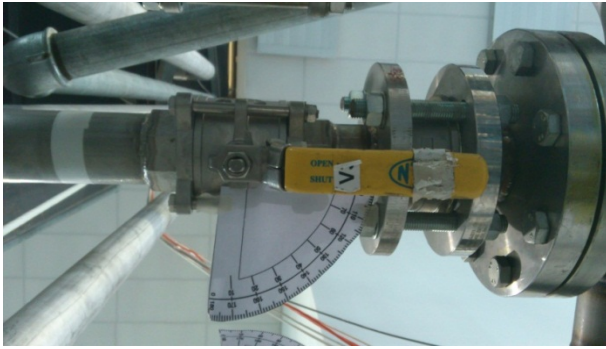
Fig. 7: Air line manual valve

For this fault case three data sets were acquired, one set under changing operational conditions and another two sets had constant air and water flow rates (see Table 4). All the details about the data sets acquired for this fault case can be seen in Fig. 8, Fig. 9, and Fig. 10, including the evolution of the fault, the profile of air and water flow rate set points, and actual measurements of air and water flow rate during the experiments.

Table 4: Operational conditions for data sets in Case 1

Data set	Operating conditions	Duration (s)	Air flow rate (m^3/s)	Water flow rate (kg/h)
1.1	Changing	5811	Varying	Varying
1.2	Steady-State	4467	2	150
1.3	Steady-State	4321	3.5	75

Data set 1.1
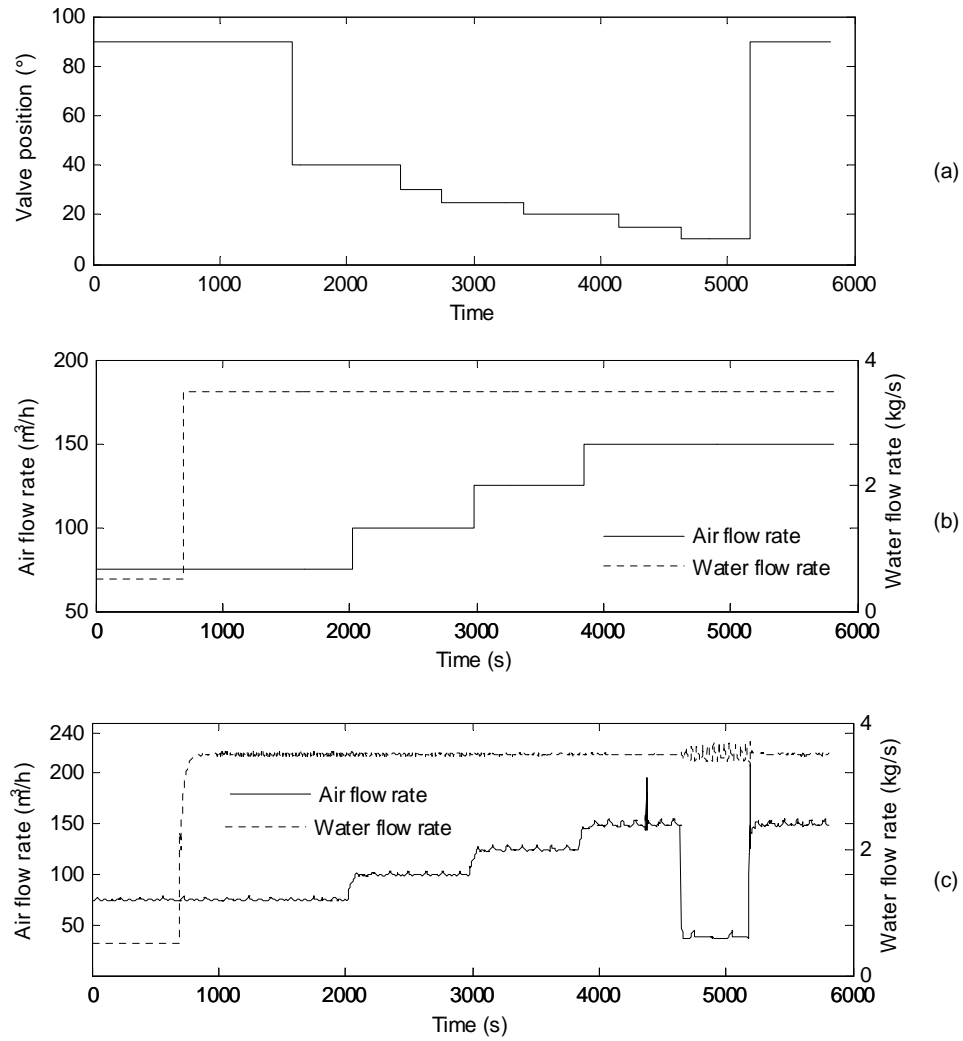
Fig. 8: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 1.1

Data set 1.2
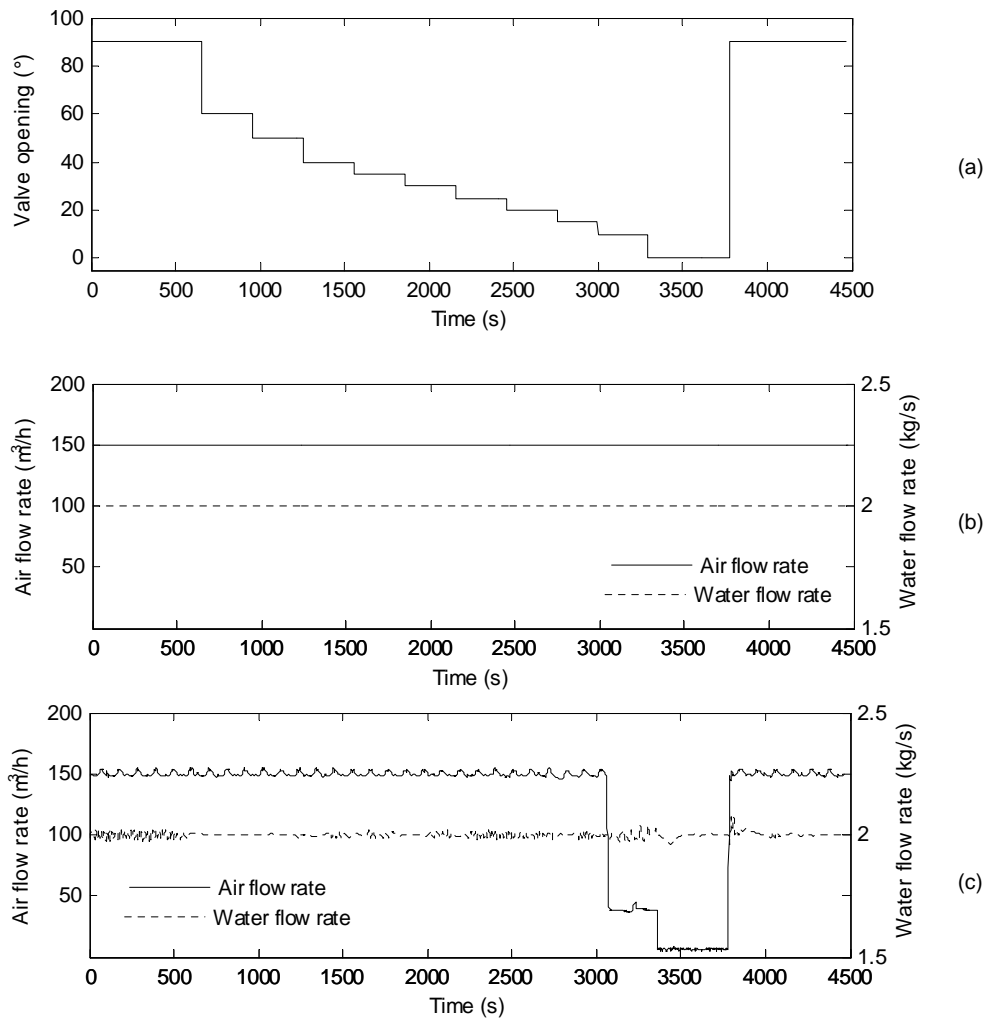
Fig. 9: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 1.2

48
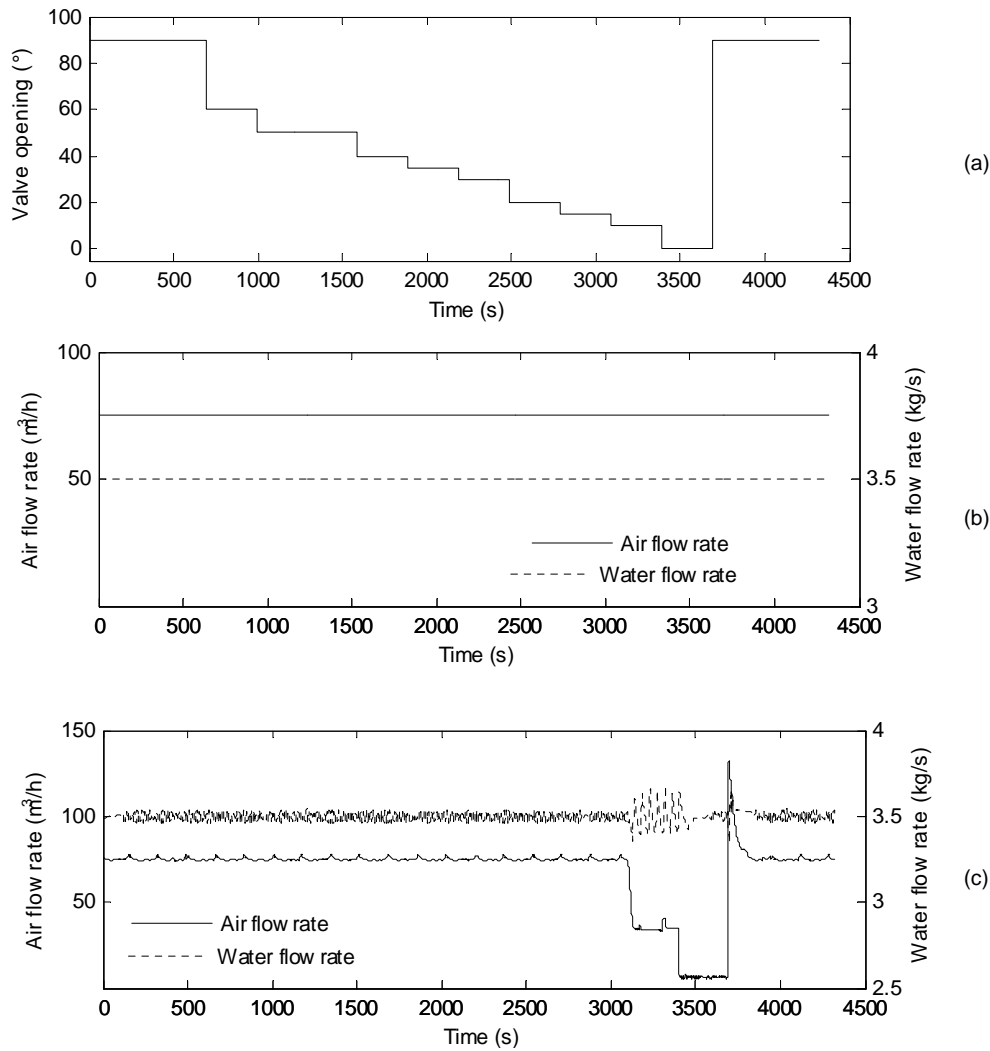
Data set 1.3

Fig. 10: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 1.3

Case 2: Water line blockage

This fault is exactly the same as the fault introduced in Case 1, with the difference that in this case the water line valve (see Fig. 11) was closed gradually to simulate the blockage instead of the air line valve. Despite the fault is basically the same, the results are expected to be different, as the physical properties of air and water (especially density and viscosity) are considerably different. For this case three data sets were acquired, one under changing operational conditions and the other two sets with constant air and water flow rates (see Table 5). All the details about the data sets acquired for this fault

49
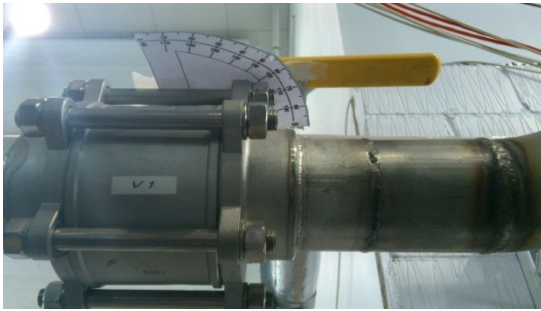
Fig. 11: Water line manual valve

case can be seen in Fig. 12, Fig. 13 and Fig. 14 including the evolution of the fault, the profile of air and water flow rate set points and actual measurements of air and water flow rate during the experiments.

Table 5: Operational conditions for data sets in Case 2

Data set	Operating conditions	Duration (s)	Air flow rate (m^3/s)	Water flow rate (kg/h)
2.1	Changing	9192	Varying	Varying
2.2	Steady-State	3496	2	100
2.3	Steady-State	3421	3.5	150

Data set 2.1
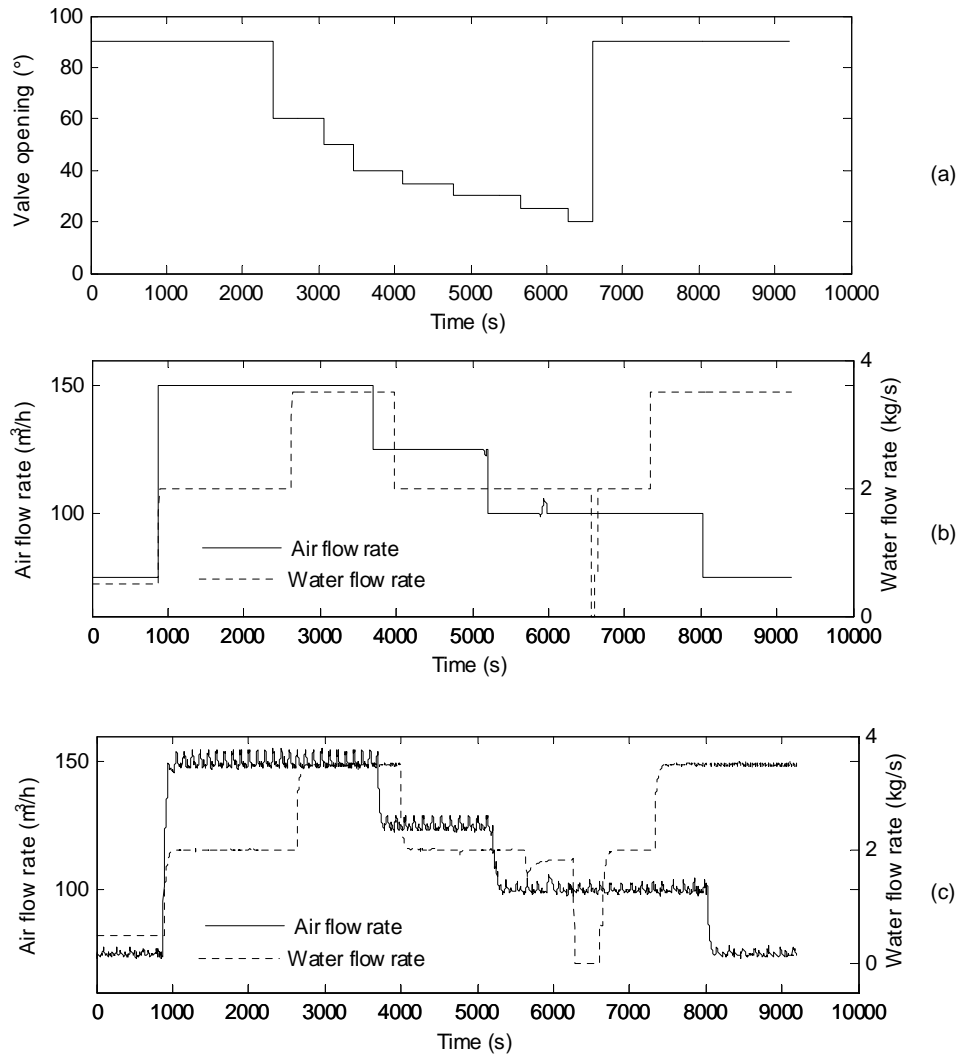
Fig. 12: Fault evolution (a) flow rate set points (b) and measured flow rate conditions (c) for data set 2.1

Data set 2.2

(a)

(b)

(c)
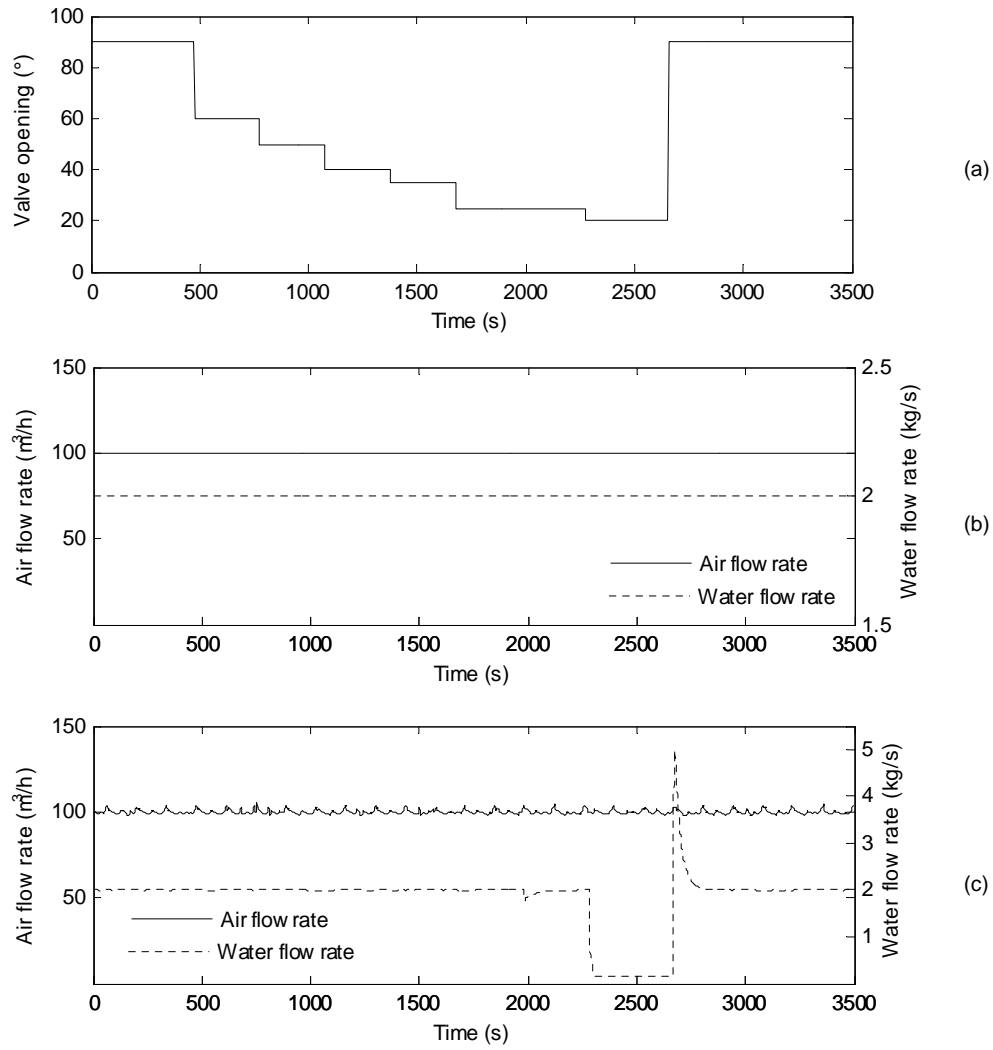
Fig. 13: Fault evolution (a) flow rate set points (b) and measured flow rate conditions (c) for data set 2.2

Data set 2.3

(a)

(b)

(c)
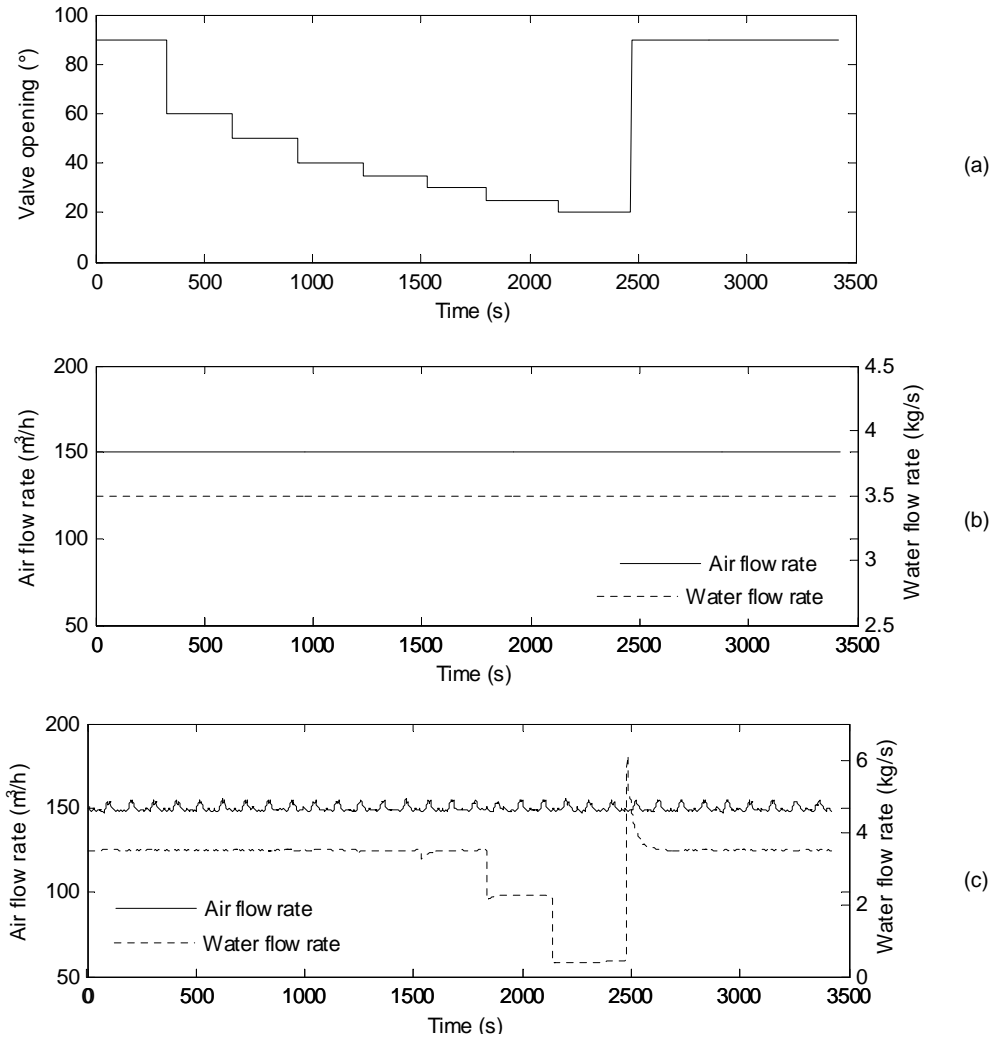
Fig. 14: Fault evolution (a) flow rate set points (b) and measured flow rate conditions (c) for data set 2.3

Case 3: Top separator input blockage

In this case the type of fault introduced is again similar to the air and water line blockage. The main difference is that the valve manipulated in this case is VC404 (top separator input, see Fig. 15) which is operated pneumatically. The valve can be controlled remotely from the control room and the angular position can be accurately measured. This allows the precise observation of how the fault severity affects the results provided by condition monitoring algorithms. The fault was introduced in most cases with a fast evolution in a first instance due to the inherent behaviour of ball valves: in this type of valves the change in

the pressure drop is almost insignificant for valve positions near the fully open position, but it is much more sensitive to changes in the valve position as the valve is gradually closed. For this case four data sets were acquired, one of them under changing operational conditions and another three with constant air and water flow rates (see Table 6). All the details about the data sets acquired for this fault case can be seen in Fig. 16, Fig. 17, Fig. 18 and Fig. 19 including the evolution of the fault, the profile of air and water flow rate set points and actual measurements of air and water flow rate during the experiments.

Fig. 15: Top separator input valve VC404

Table 6: Operational conditions for data sets in Case 3

Data set	Operating conditions	Duration (s)	Air flow rate (m^3/s)	Water flow rate (kg/h)
3.1	Changing	9090	Varying	Varying
3.2	Steady-State	6272	2	100
3.3	Steady-State	10764	3.5	75
3.4	Steady-State	8731	3.5	75
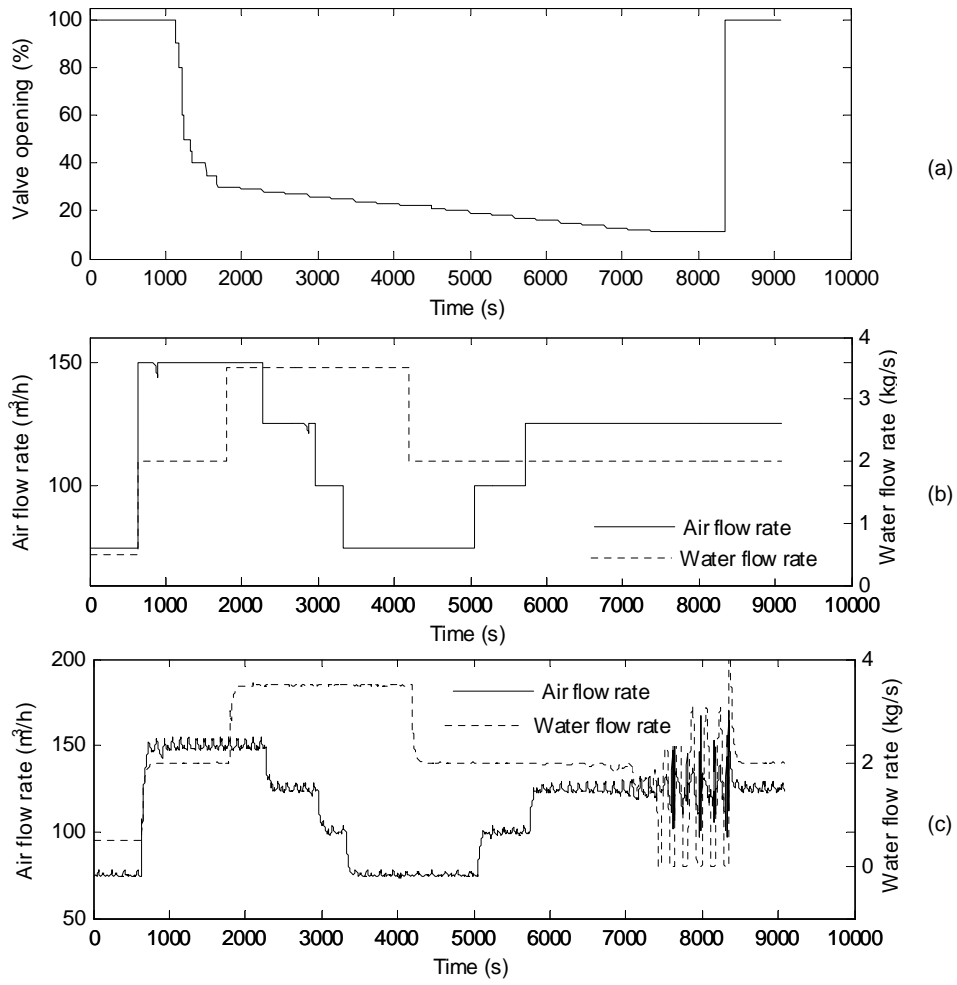
Fig. 16: Fault evolution (a) flow rate set points (b) and measured flow rate s (c) for data set 3.1

Data set 3.2
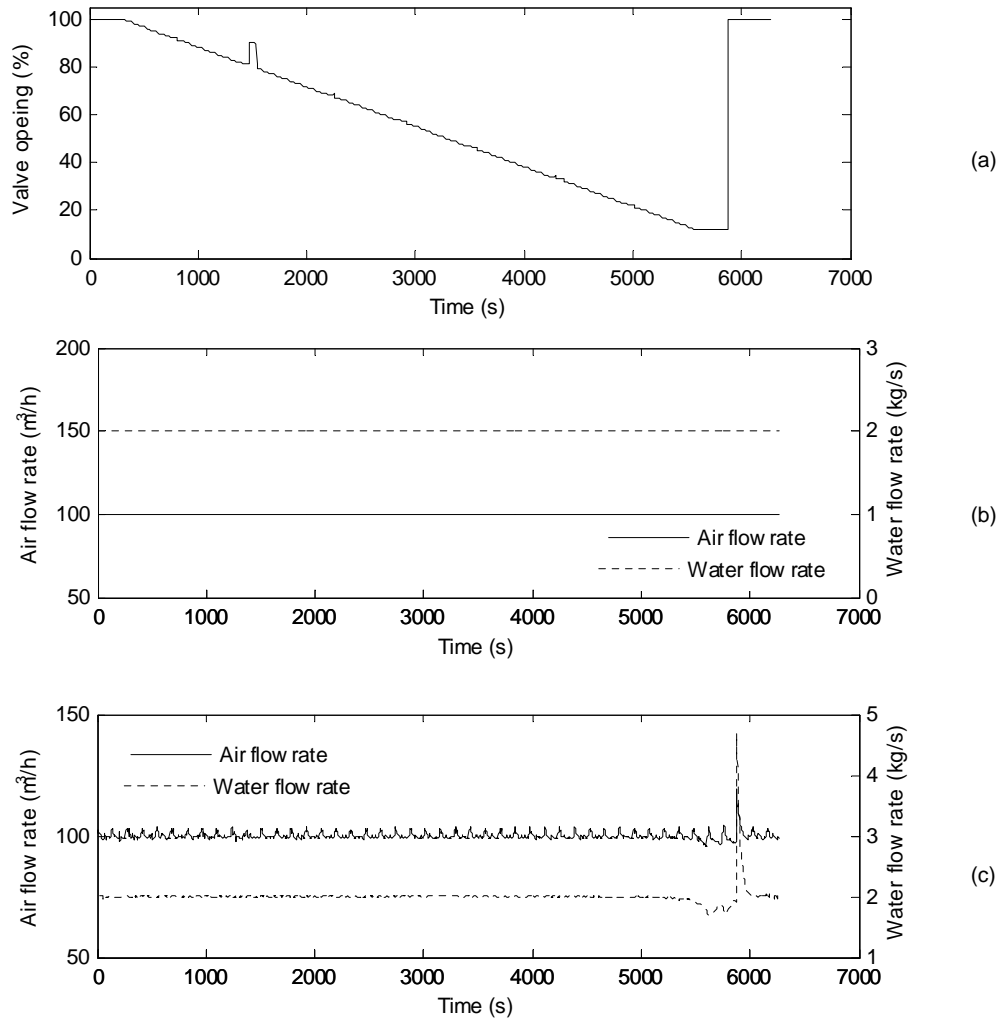
Fig. 17: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 3.2

Data set 3.3
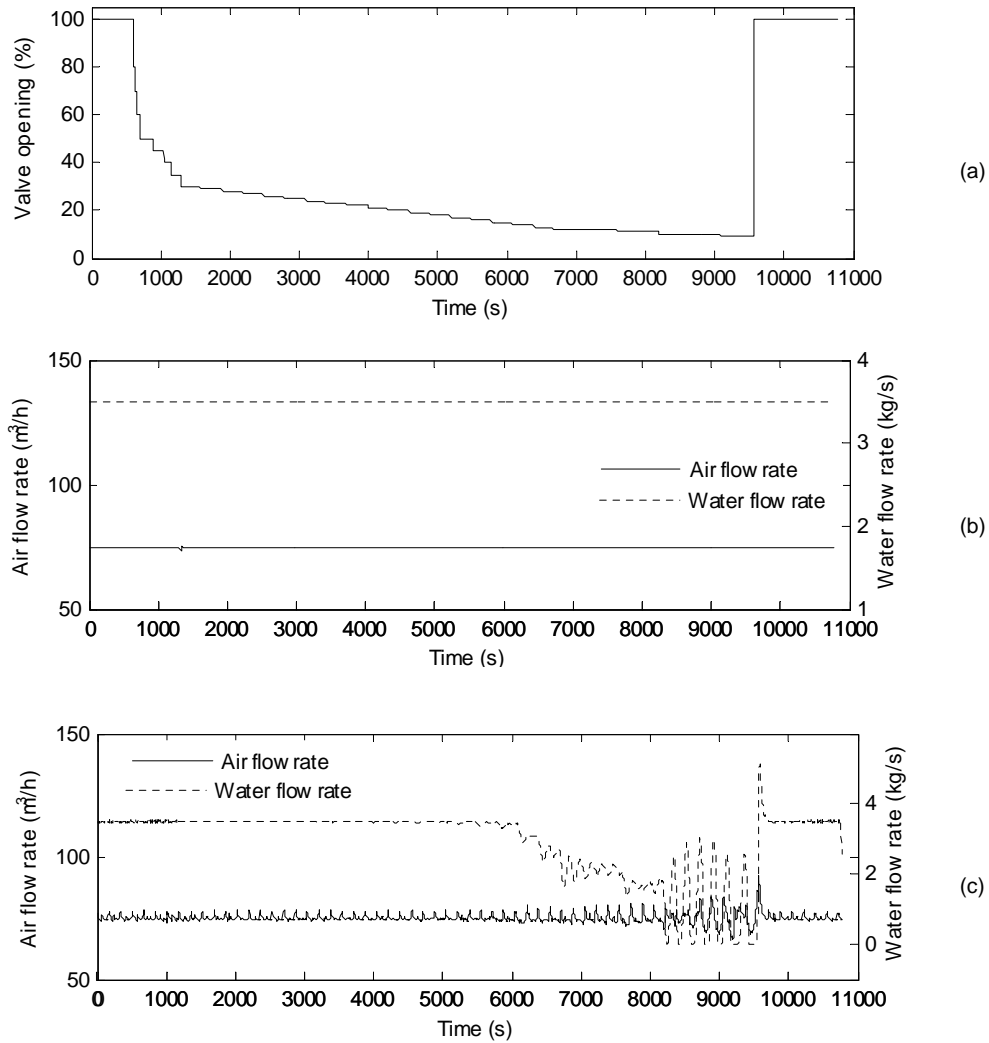
Fig. 18: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 3.3

Data set 3.4
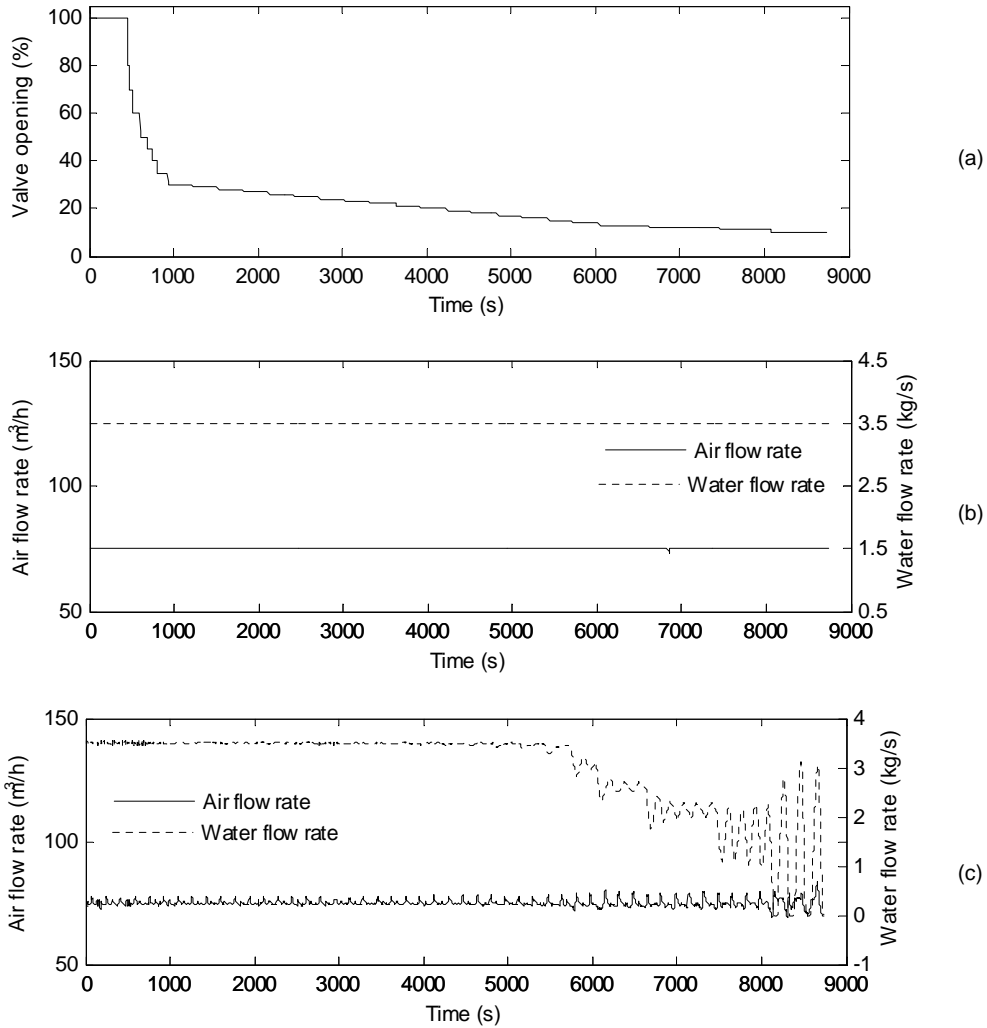
Fig. 19: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 3.4

Case 4: Open direct bypass

The bypass line can direct the multiphase flow after the mixing point to the 3-phase separator, bypassing the riser. There are valves at the beginning and at the end of this alternative line to isolate it from the rest of the system, and in normal conditions they are always closed. One of these valves is represented in Fig. 20. The objective of this fault is to simulate a leakage at the bottom of the riser, causing a lack of flow in the top of the riser. For this case three data sets were acquired, one set under changing operational conditions and another two with constant air and water flow rates (see Table 7). All the details about the
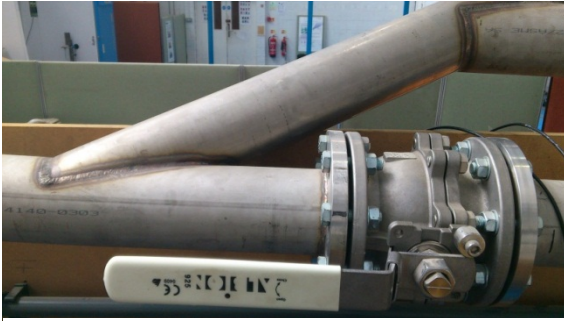
Fig. 20: Bypass line derivation and valve

data sets acquired for this fault case can be seen in Fig. 21, Fig. 22 and Fig. 23 including the evolution of the fault, the profile of air and water flow rate set points and actual measurements of air and water flow rate during the experiments.

Table 7: Operational conditions for data sets in Case 2

Data set	Operating conditions	Duration (s)	Air flow rate (m^3/s)	Water flow rate (kg/h)
4.1	Changing	7208	Varying	Varying
4.2	Steady-State	4451	2	150
4.3	Steady-State	3661	3.5	75
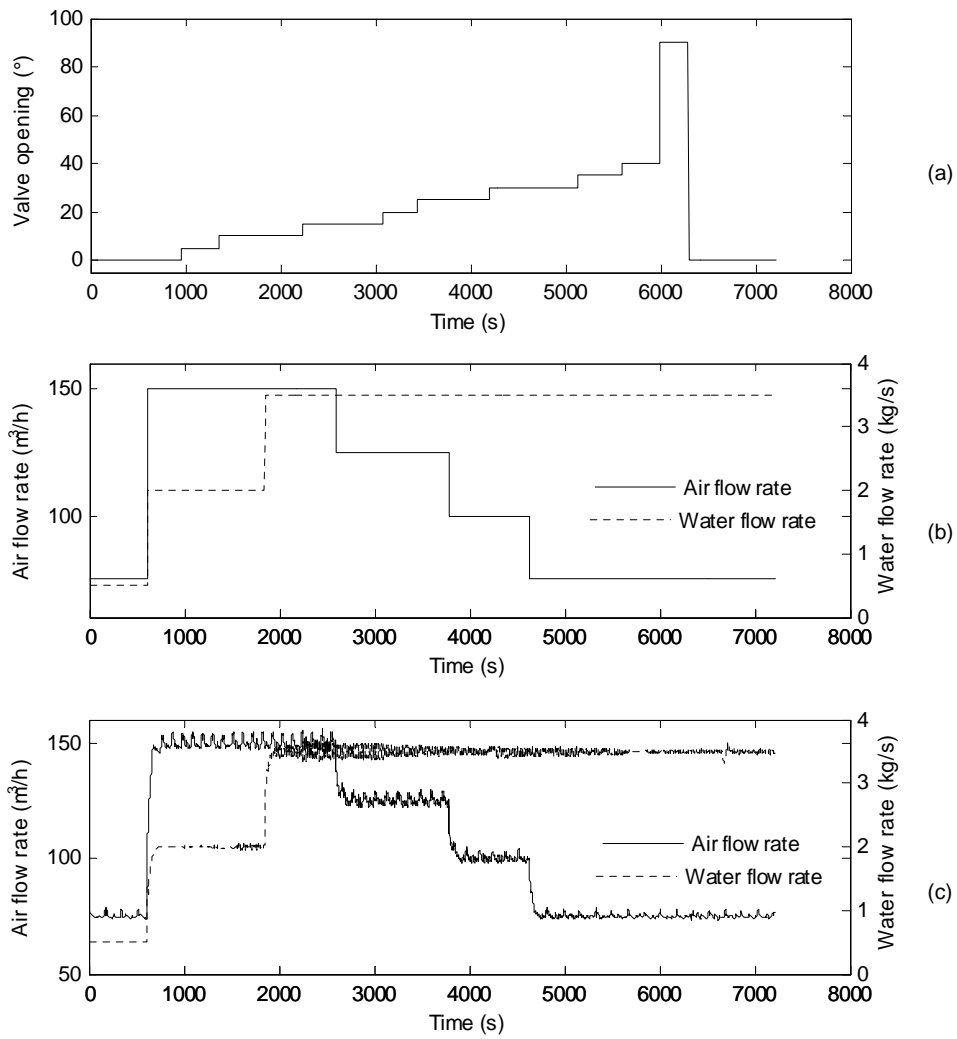
Fig. 21: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 4.1

Data set 4.2
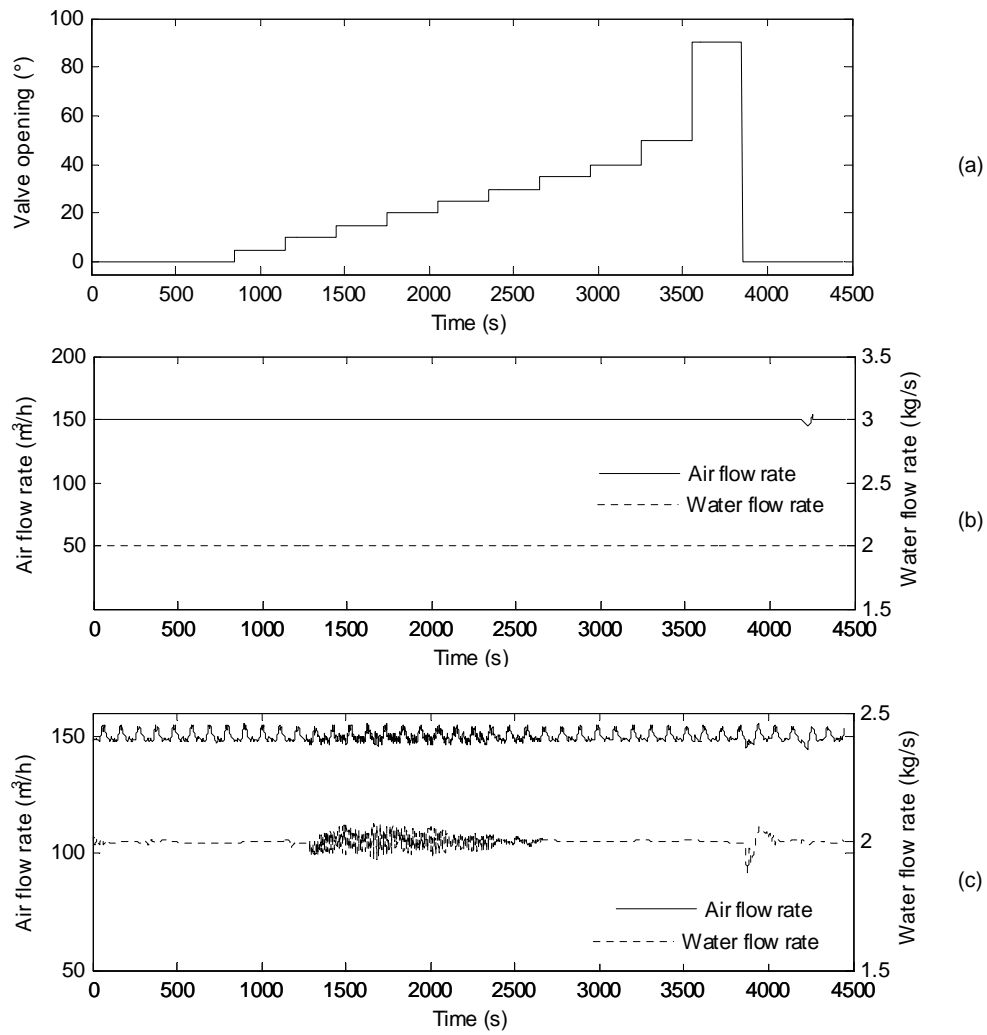
Fig. 22: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 4.2

61
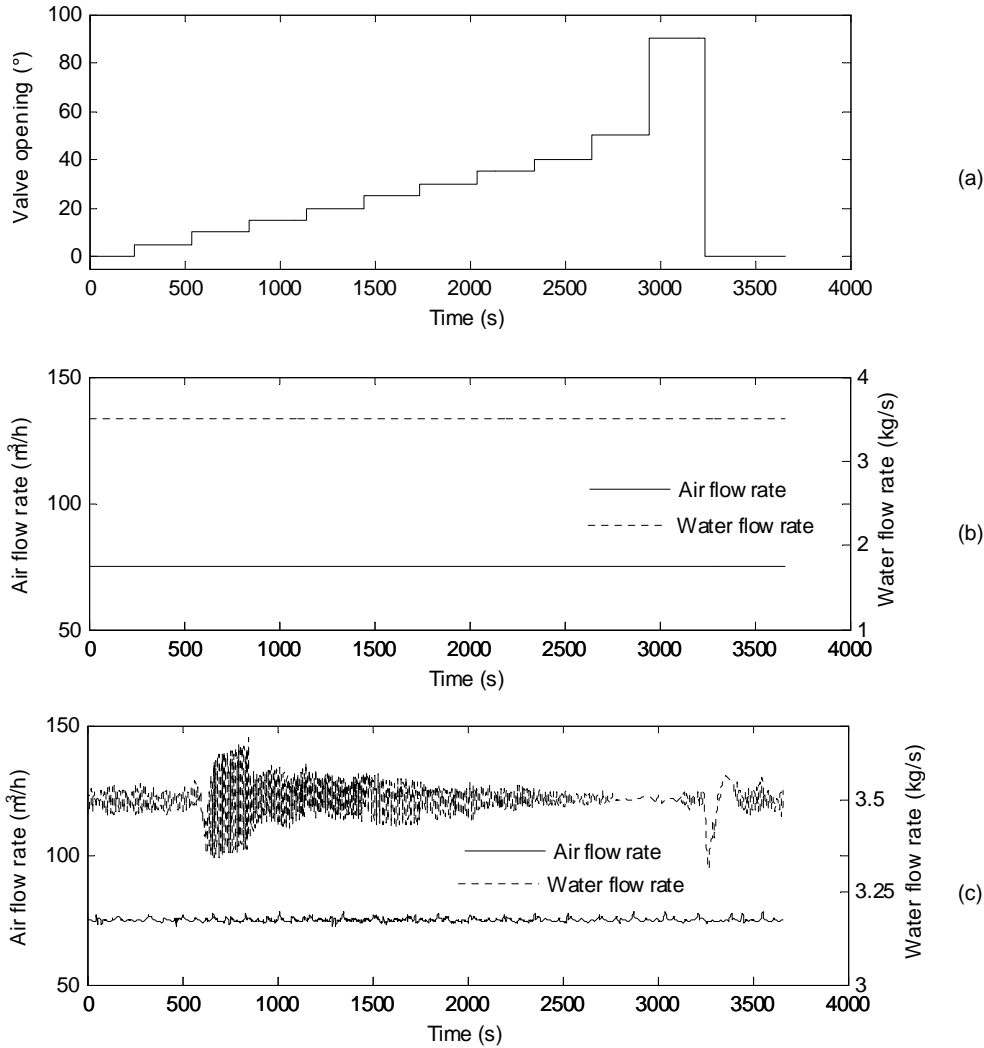
Data set 4.3

Fig. 23: Fault evolution (a) flow rate set points (b) and measured flow rates (c) for data set 4.3

Case 5: Slugging conditions

Slugging [76] is a transient phenomenon that can occur in risers with multiphase flow when the velocities of the gas and the liquid are relatively low. The liquid tends to accumulate in the base of the riser blocking the flow. Due to this blockage, the pressure builds upon the blockage until it is sufficient to flush the liquid (and the gas) out of the riser. After this surge, the liquid remaining in the riser falls down, creating a new blockage and starting the cycle again. This phenomenon is typical for offshore oil production systems, where multiphase hydrocarbon fluids travel from an oil filed along a long pipeline on the sea bed to

a riser connecting to a separation process on an offshore oil rig. It can produce large amplitude fluctuations in the pressure and flow rates, which can affect and damage the equipment [77]. The fault was introduced by reducing the air and water flow rates to regimes where slugging is produced. During the test, the flow rate was varying at different points from normal to slugging conditions. For this case 2 data sets were acquired, both of them under changing operational conditions (see Table 8).

Table 8: Operational conditions for data sets in Case 5

Data set	Operating conditions	Duration (s)	Air flow rate (m^3/s)	Water flow rate (kg/h)
5.1	Changing	2541	Varying	Varying
5.2	Changing	10608	Varying	Varying

Fig. 24 and Fig. 25 describe in detail each one of the data sets, including the profile of air and water flow rate set points, actual measurements of air and water flow rate and the evolution of the pressure in the bottom of the riser during the experiments, where it is possible to observe the fluctuations caused by the slugging. For each data set the flow rate combinations where slugging is expected are shaded in grey in the chart corresponding to the flow rate set points.

Data set 5.1
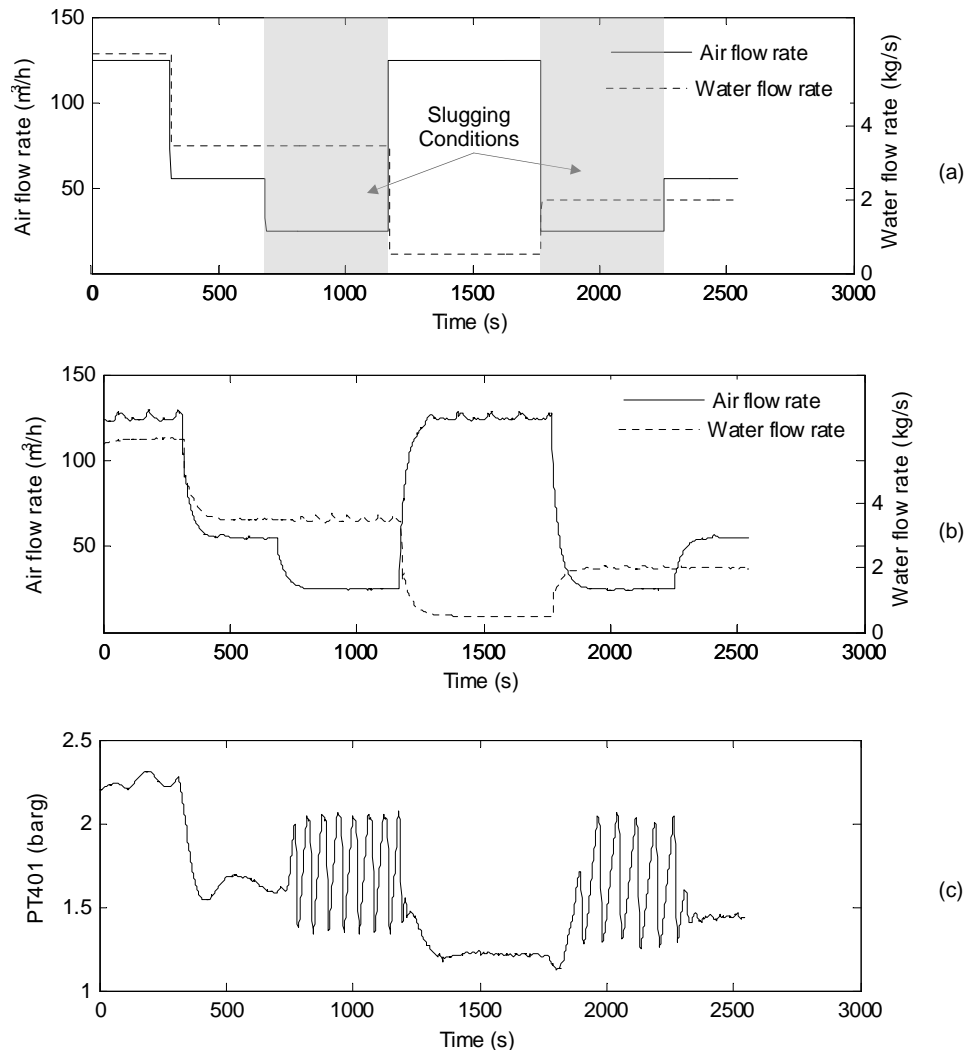
Fig. 24: Flow rate set points (a), measured flow rates (b) and bottom riser pressure PT401(c) for data set 5.1

Data set 5.2
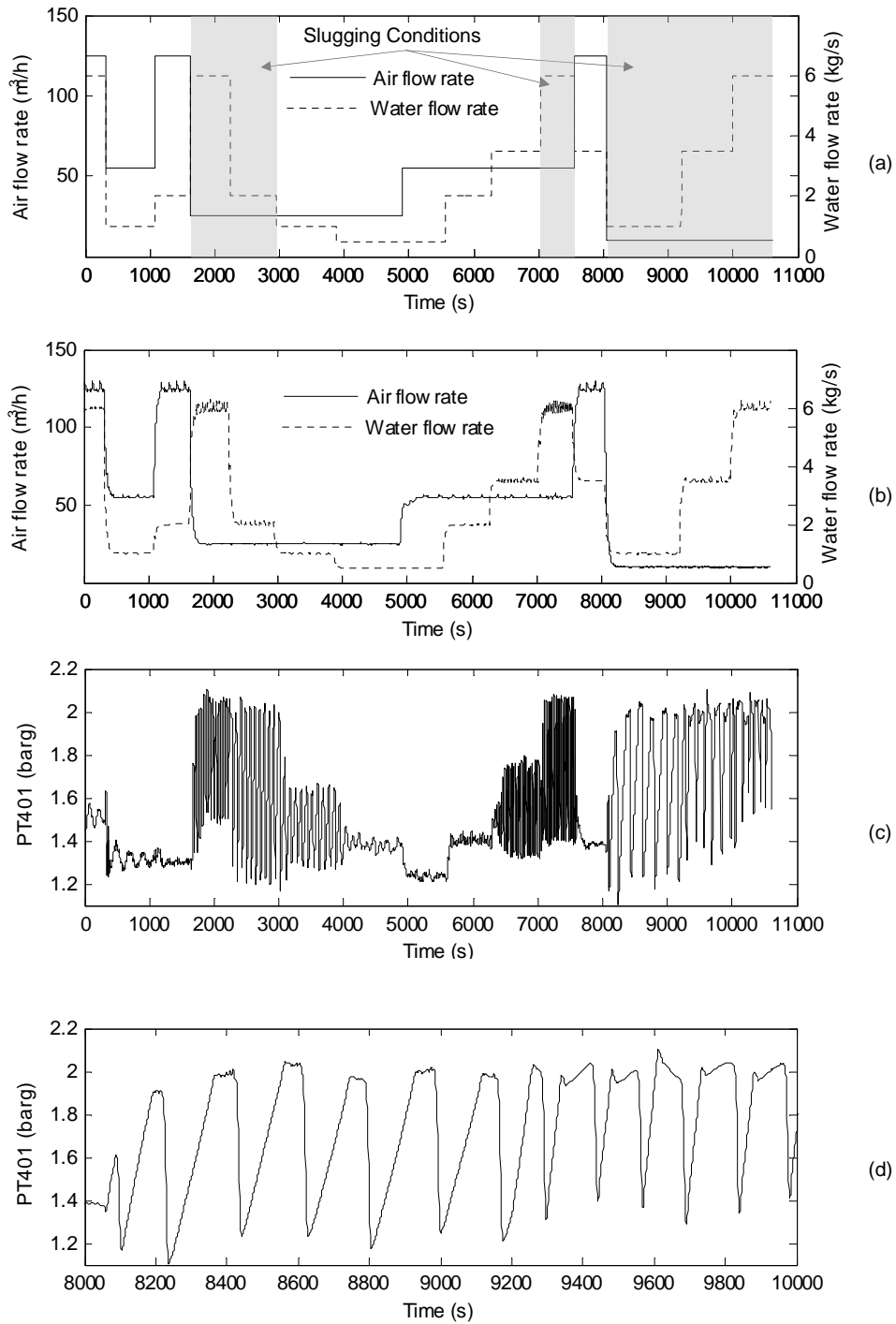
Fig. 25: Flow rate set points (a), measured flow rates (b) and bottom riser pressure PT401(c and d) for data set 5.2

Case 6: Pressurization of the 2" line

In the cases presented previously the 4" line was used to carry the flow to the top of the riser. In these conditions, the 2" line is totally isolated from the rest of the system and thus it should not be pressurized. In the top of the riser, just before the input of the top separator there is a bridge with a valve that connects the 4" and the 2" line. Fig. 26 shows a detail of this valve:

Fig. 26: Bridge valve between the 4" line and the 2" line

The fault introduced consisted basically in opening this bridge valve keeping both sides of the 2" isolated from the rest of the installation. The objective of this fault condition is to simulate an unusual operation of the system that in a real process can cause degradation in the performance or in the output product quality or even safety issues. This fault should not affect the flow conditions in the 4" line or other parts of the rig, but it will pressurize the 2" line. In this particular case, an additional variable was measured and included in the analysis. This variable is the pressure measured in the bottom of the riser in the 2" line by PT417. For this case 2 data sets were acquired, both of them under changing operational conditions (see Table 9).

Table 9: Operational conditions for data sets in Case 6

Data set	Operating conditions	Duration (s)	Air flow rate (m^3/s)	Water flow rate (kg/h)
6.1	Changing	2800	Varying	Varying
6.2	Changing	4830	Varying	Varying

Fig. 27 and Fig. 28 describe in detail each one of the data sets including the fault evolution, the pressure measurements observed in PT417, the profile of air

and water flow rate set points and actual measurements of air and water flow rates during the experiments.

Data set 6.1
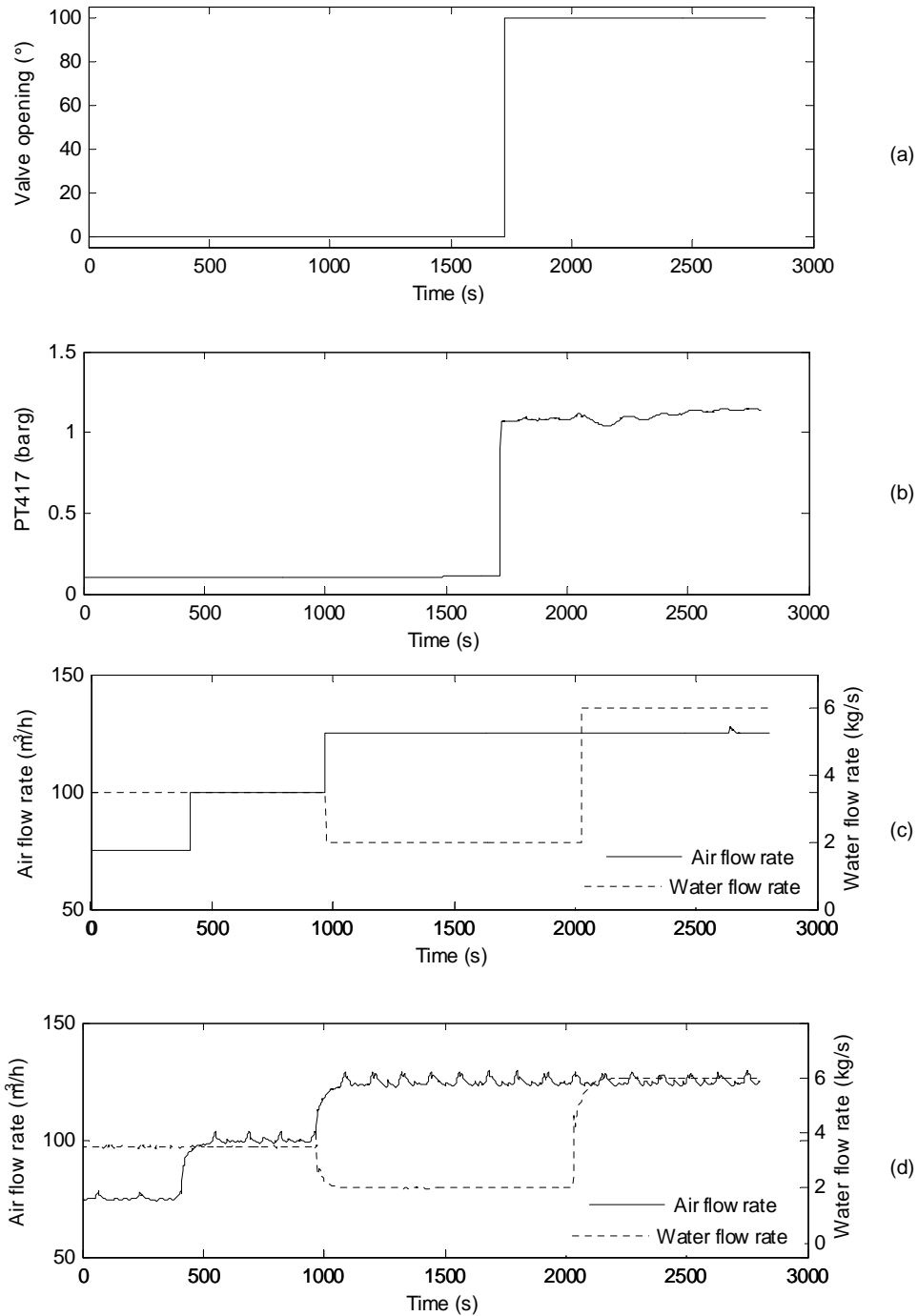
Fig. 27: Fault evolution (a), 2" line pressure PT417 (b), flow rate set points (c) and measured flow rates (d) for data set 6.1

Data set 6.2
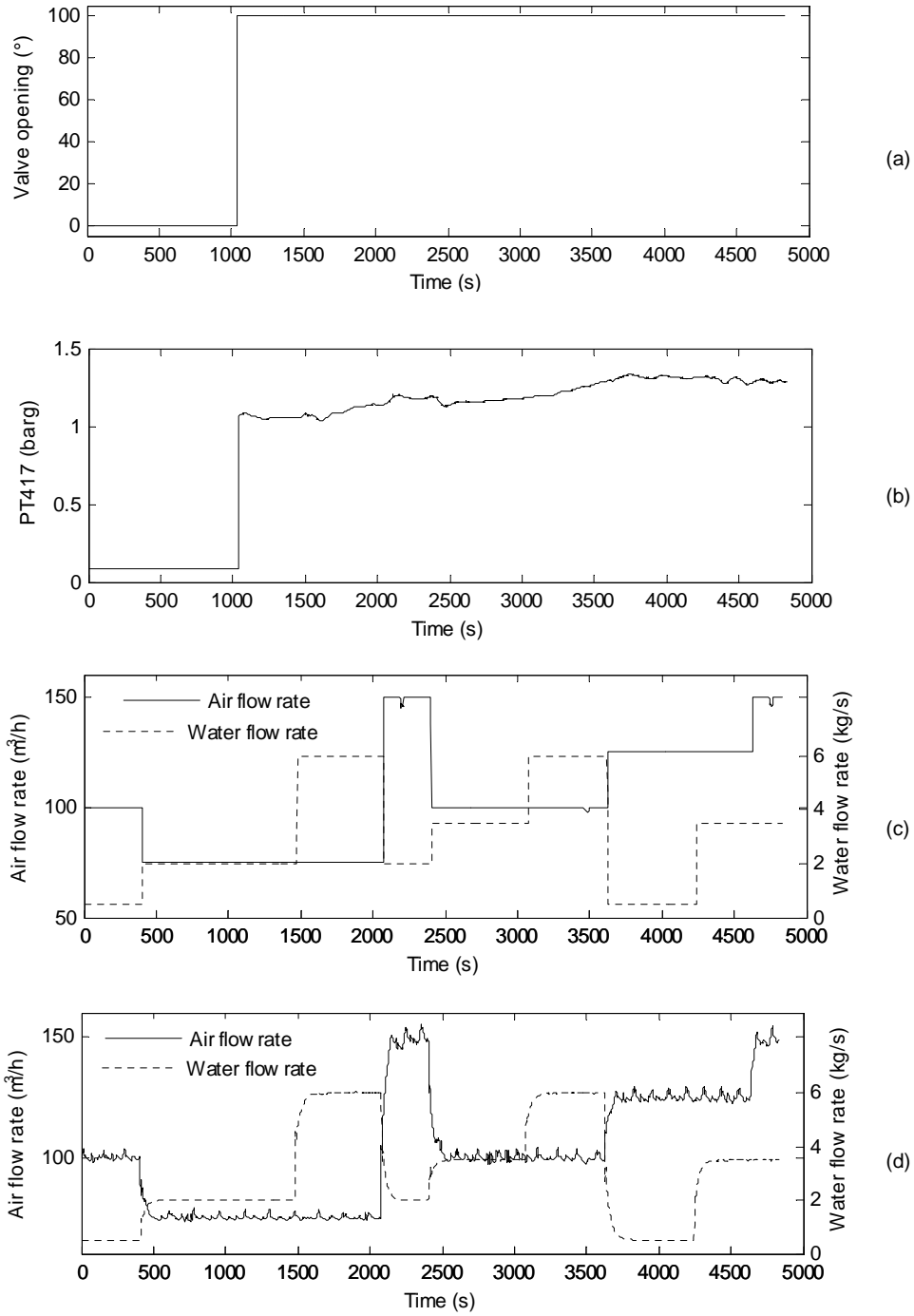
Fig. 28: Fault evolution (a), 2" line pressure PT417 (b), flow rate set points (c) and measured flow rates (d) for data set 6.2

3.3 Data analysis software

A special software tool was designed in Matlab (version R2010A) for the analysis of the data sets described in 3.2 using different multivariate algorithms. It is a graphic user interface where the user can easily select the set of data to analyse, the process variables included in the analysis and the tuning parameters for the selected algorithm. The objective of this tool is to allow the user to run multiple analyses with different configurations in a fast and organized manner. In this way it is easier to observe how changes in the training data sets selected, the variables included in the analysis, or the tuning parameters affect the final results in terms of fault detection and diagnosis. Fig. 29 shows the software interface and the different sections of the main window. The following subsections describe in detail the different parts of the main menu.
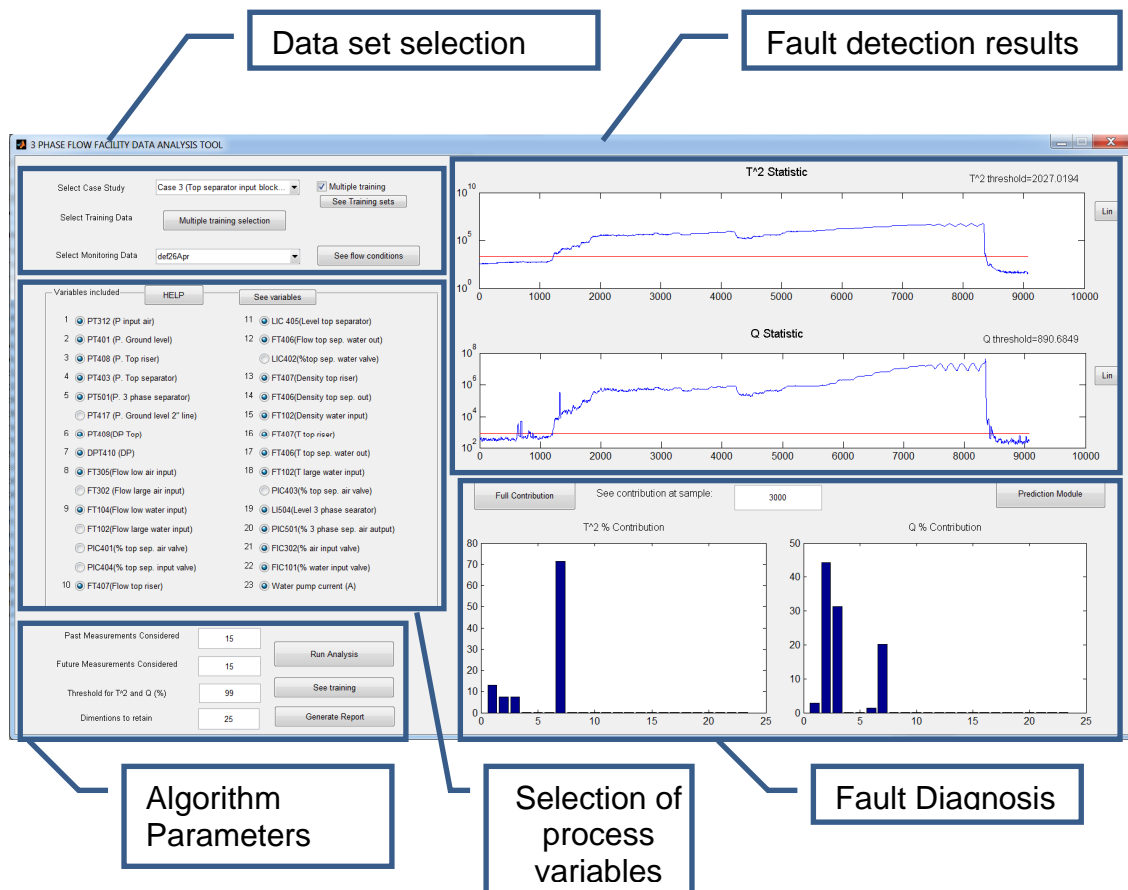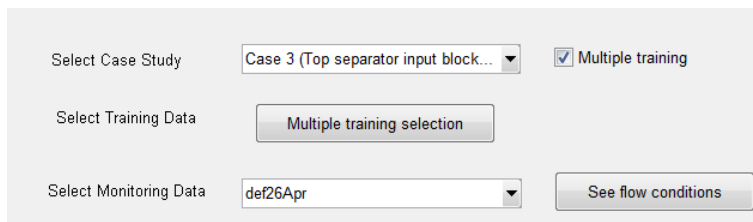
Fig. 29: Software tool main menu

69

3.3.1 Data set selection

In this section the user can select the data sets to be analysed. Initially, it is necessary to select the data set used for training the algorithm and calculate the transformation matrices and the thresholds for the health indicators. Normally, the data sets acquired under normal operational conditions will be used for training. To obtain good results in terms of fault detection and diagnosis it is important to select a training data set where the operational conditions are representative of the conditions that will be found in the data set to analyse. In this section of the program it is also possible to obtain a plot of the operational conditions (air and water flow rate) to help the user to decide if a particular data set should be included or not.

Fig. 30: Detail of data set selection section

3.3.2 Selection of process variables

The variables selection section is designed to allow the user to observe how the monitoring results are affected by the process variables selected for the analysis. Originally 30 different process variables were measured during the tests, but only 24 were included in the final analysis presented in chapter 4. For this case study it was particularly important to remove the valve position measurements of the valves which were not manipulated during the tests. The addition of measurements consisting in a constant value (e.g. 100% for valves fully opened during the analysis or 0% for those fully closed) without any deviation due to noise or changes in the operational conditions causes problems in the mathematical manipulation of the data. The presence of several columns consisting in constant values can affect the internal process of matrix inversion, essential in many multivariate algorithms. This was the case of the two outlet valves in the top and bottom of the top separator (VC401 and VC402) and the top separator input valve (VC404). Two flow meters situated in in

70

alternative pipe branches that were not used in this case study were also excluded (FT302 and FT102). The pressure measured in the bottom of the riser in the 2" line (PT417) was only included in Case 6 for both the normal and faulty operation data.
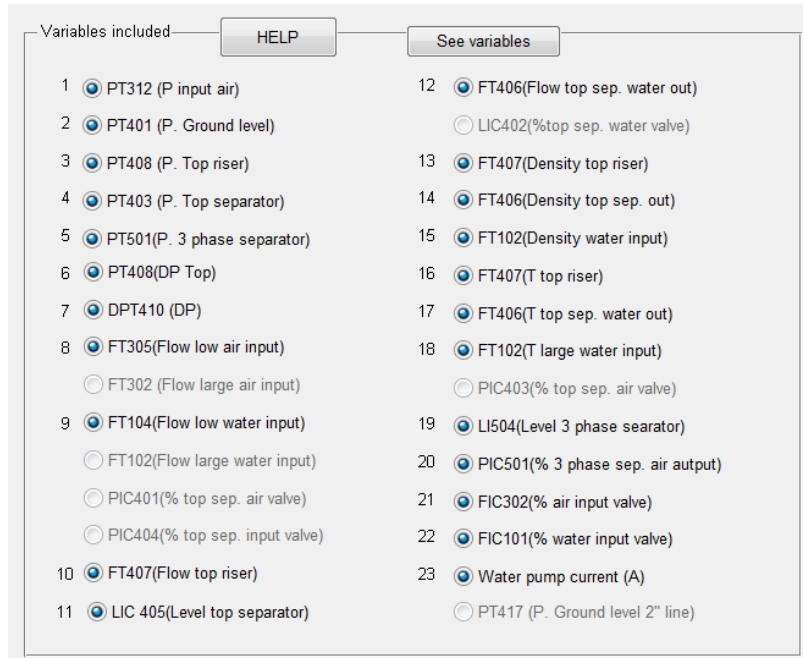
Fig. 31: Detail of the process variables selection section

3.3.3 Algorithms Parameters

The application of multivariate algorithms for process monitoring normally requires the selection of certain parameters such as the number of lags to be considered, the confidence bound for the thresholds of the health indicators or the number of dimensions to be retained. In this part of the program (Fig. 32) the user can select the algorithm to be used for the data analysis and introduce these parameters. In addition it is possible to obtain a plot of the statistical distribution of the health indicators calculated during the training period, which can be helpful for the selection of the desired confidence bound (see Fig. 33).
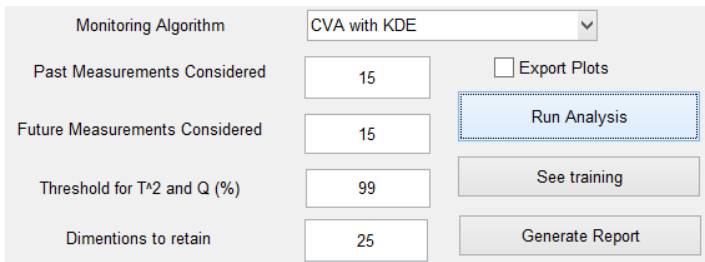
Fig. 32: Detail of the algorithms parameters section

From this section the calculation can be run, obtaining the results in the "Fault detection results" section, and it is also possible to obtain a report of the violations of the health thresholds pushing the button "Generate report".
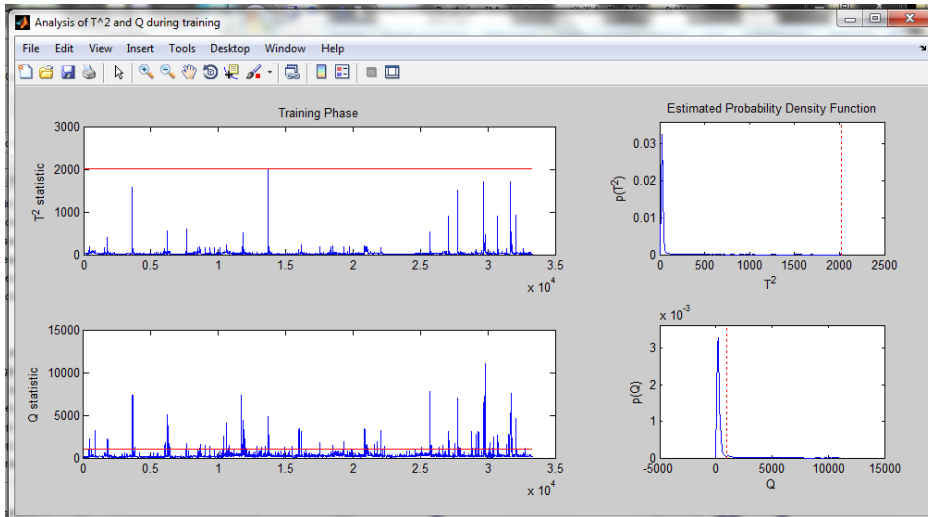
Fig. 33: Example of health indicators during training phase (left) and their statistical distribution (right)

3.3.4 Fault detection results

After running the analysis, the T^2 and Q indicators for the monitoring data set are plotted in this section, as well as the thresholds represented by a red line in each plot. In these charts the user can see if the threshold has been violated at any time, indicating that a fault was detected (see Fig. 34). The y axis can be represented either in linear or logarithmic scale to provide a better visualization of the health indicator against the threshold.
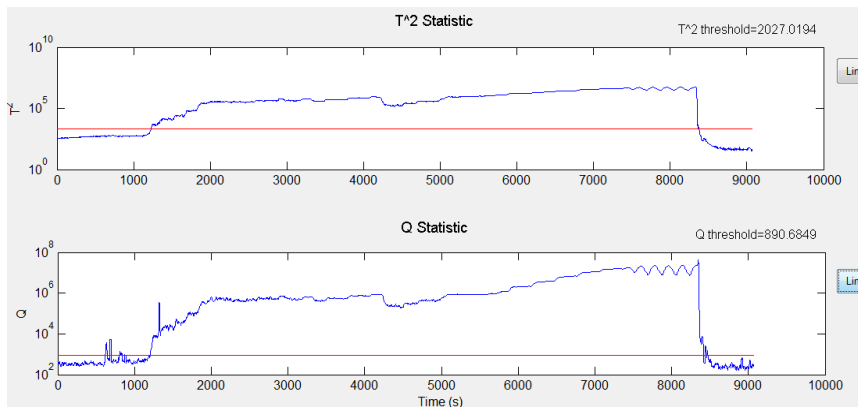
Fig. 34: Detail of fault detection section

3.3.5 Fault diagnosis

Once the health indicators have been calculated for the monitoring data set, the contribution plots for the statistic indicators can be plotted at any time point selected by the user. These charts represent how much is contributing to the final value of the indicators each one of the process variables included in the analysis, helping the user to identify the origin of the fault (Fig. 35).
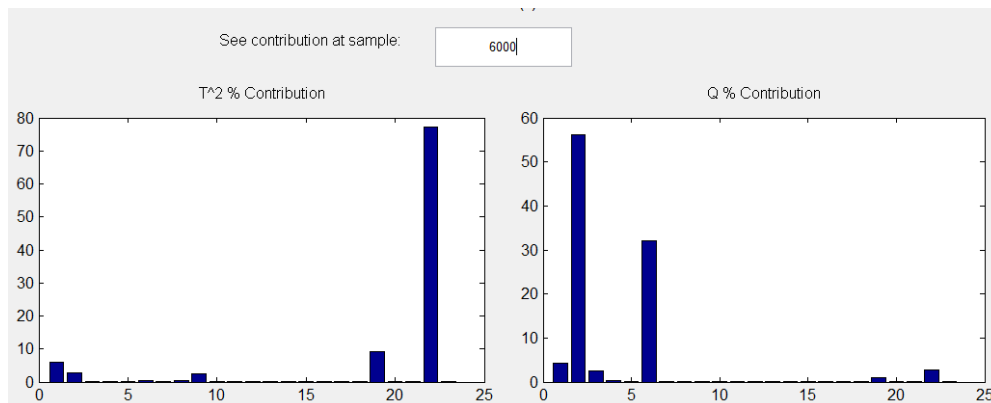
Fig. 35: Detail of fault diagnosis section

3.4 Application example: Fault detection using PCA

This section shows an example of application of the data sets described in the previous sections. The PCA was used to detect the fault introduced in data set 1.1 after the computation of the transformation matrices and the thresholds for the indicators using the normal operation data sets T1, T2, and T3.

73

3.4.1 Introduction to PCA

PCA can characterize the state of a process by projecting the measured data into a lower-dimensional space. This dimensionality reduction technique preserves the correlation between the process measurements capturing the data variability in an optimal way [2]. Given an $n\,x\,m$ data set matrix X containing n observations and m variables it is possible to obtain a set of loading vectors V by solving the eigenvalue decomposition of the covariance matrix S:

$$S = \frac{1}{n-1} X^T X = V \Lambda V^T \tag{3-1}$$

where the loading vectors V are ordered by the amount of variance expressed by the corresponding eigenvalues in the diagonal matrix Λ. The loading vectors attached to the a largest singular values are retained in the loading matrix P $\in \mathbb{R}^{m x a}$. These vectors will be used to produce a lower dimensional representation of the measured data that captures systematic trends of the process, separating it from the part of the data which contains basically random noise. The score matrix W contains the projection of the observed data into the lower-dimensional space, while the residual matrix E represents the difference between the observations and the projection of W back into the m-dimensional space:

$$W = XP \tag{3-2}$$

$$E = X - WP^T \tag{3-3}$$

The indicators most commonly used for the detection of faults are the Hotelling T^2 indicator (which represents major variations in the data) and the squared prediction error Q (representing variations in the residual space). These indicators can be computed for each observation x as follows:

$$T^2 = y^T y \tag{3-4}$$

$$Q = x^T (I - PP^T) x \tag{3-5}$$

where the principal components $y = \Lambda^{1/2} P^T x$ have been re-scaled to have unit variance [15].

Assuming that the observations are sampled randomly from a multivariate normal distribution, the threshold for the T^2 indicator can be estimated as [78]:

$$T_\alpha^{\,2} = \frac{a(n-1)(n+1)}{a(n-a)} F_\alpha(a, n-a) \tag{3-6}$$

where $F_\alpha\,(a,\,n\text{-}a)$ is an F-distribution having degrees of freedom a and $n\text{-}1$ with a significance level of α. The control limit for the Q index was computed as defined by Jackson and Mudholkar [79]:

$$Q_\alpha = \theta_1 \left[\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \tag{3-7}$$

where $\theta_j = \sum_{a+1}^{n} \left(\lambda_{ii} \right)^j$ for $\mathrm{j} = 1, 2, 3$ $\mathrm{h}_0 = 1 - \dfrac{2\theta_1 \theta_3}{3\theta_2^{\,2}}$, c_α is the normal deviate that

corresponds to the (1-α) percentile.

3.4.2 Results and discussion

Training

The three data sets acquired during normal operation were joined to give a continuous run of 33397 observations and 23 variables that captured all normal operating conditions. This was then used as the training data set for building the PCA model. Five principal components were retained based on the number of principal components whose cumulative variance accounted for 86.76% of the total variance (Table 10).

Table 10: Relative variance explained by the first 10 principal components

PC	Eigenvalue	Percent variance	Cumulative percent variance
1	8.14	35.38	35.38
2	4.95	21.53	56.91
3	3.98	17.29	74.20
4	1.97	8.54	82.74
5	0.92	4.01	86.76
6	0.84	3.65	90.41
7	0.81	3.50	93.91
8	0.50	2.18	96.09
9	0.23	1.02	97.11
10	0.19	0.84	97.95

The thresholds for the T^2 and Q statistics were calculated using (3-6) and (3-7) for a confidence bound of 99%, obtaining a result of 15.09 and 9.69 respectively.

Monitoring

The transformation matrices calculated during the training period were used to compute the T^2 and Q statistics for the data set 1.1. Fig. 36 shows the monitoring charts for Case 1.1, where the T^2 and Q indicators are plotted in black and the threshold of these indicators is represented by a dashed grey line. The fault start and end points are marked as vertical dashed lines. It can be seen that both the T^2 and Q values exceeded their control limits at some points which indicates the detection of a fault. The Q index detected the fault earlier than the T^2. However, the Q index is not robust in this case as it records a high false alarm rate. That is, its value exceeded the control limit at several points that had no programmed fault (Fig. 36(b)).
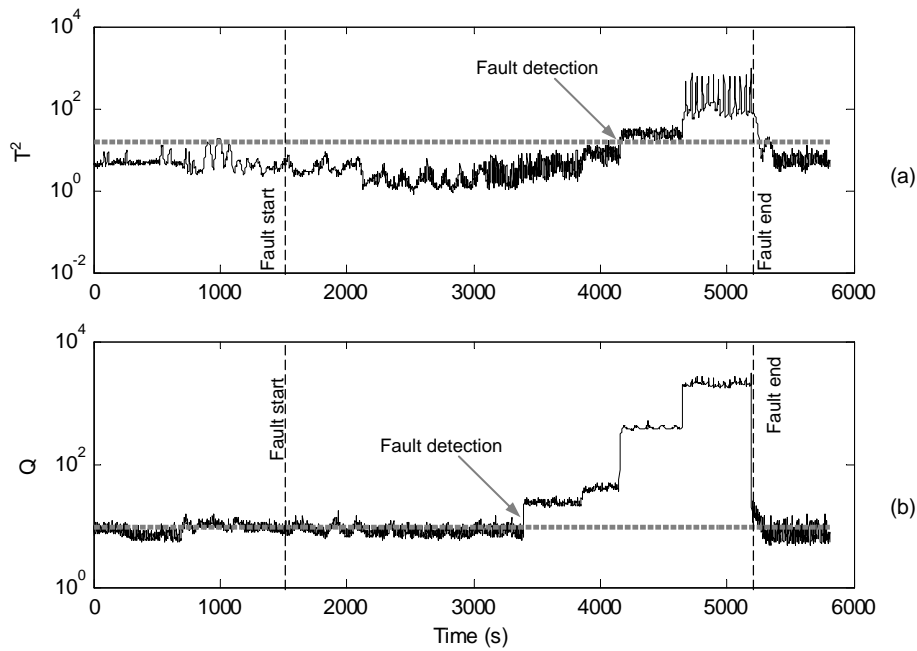
Fig. 36: On-line monitoring charts for data set 1.1 (air-line blockage, changing operational conditions)

Fig. 37 shows the contribution plots obtained from data set 1.1 for the T^2 and Q indices respectively. At sample 5000, both contribution plots point at variable 1 (air delivery pressure PT312) as the main contributor to the deviation observed during the analysis. This result correctly indicates that a fault related to the air delivery pressure occurred, showing that this fault can be detected and diagnosed using multivariate statistical methods.
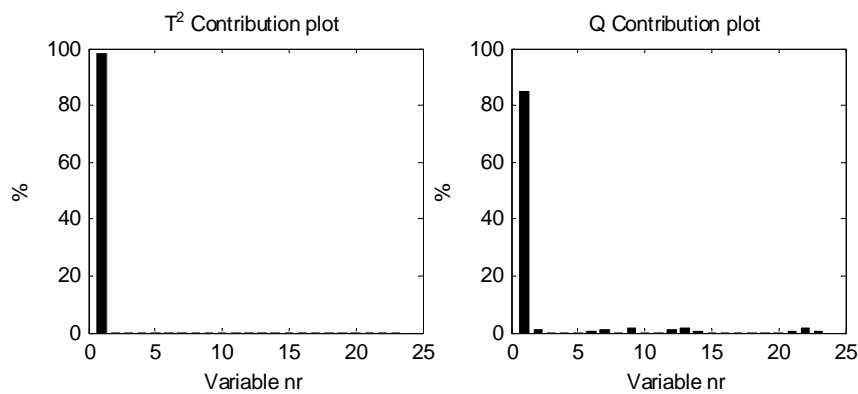
Fig. 37: Contribution plots of T^2 and Q at sample number 5000 for data sets 1.1 (air-line blockage, changing operational conditions)

3.5 Conclusion

Despite the success reported by many researchers in the application of different data driven methods for fault detection and diagnosis using computer simulated data and real data acquired in different experimental rigs and industrial facilities, there is a lack of a common benchmark case study for the comparison of the performance of this methods using real process data. This work is presented with the aim of providing a benchmark case study to prove the effectiveness of different methodologies for the detection and diagnosis of faults in processes working under varying operational conditions. In this chapter the test rig and the different case studies have been presented. This cases are composed of three data sets acquired under normal operational conditions to be used for training purposes and six cases were different faults were artificially introduced to assess the capabilities of process monitoring algorithms in detecting and diagnosing these faults. In order to provide an example of application of this case study, PCA was used for the detection and diagnosis of one of the faults introduced in the rig. The results showed that it is effectively possible to detect the fault, but the capabilities of PCA in this particular case were relatively weak due to the dynamic and non-linear nature of the process. This caused a high false alarm rate (especially in the Q indicator) and a long detection time. Consequently, there are still lots of opportunities to improve the result obtained applying more advanced algorithms that can deal with these particular challenges. The results obtained from the analysis of the data applying CVA and other multivariate algorithms will be presented in chapter 4, including a detailed description of the methodology used and a discussion about these results.

4 STATISTICAL PROCESS MONITORING OF A MULTIPHASE FLOW FACILITY, PART II: APPLICATION OF CANONICAL VARIATE ANALYSIS FOR FAULT DETECTION AND DIAGNOSIS

Abstract

Industrial needs are evolving fast towards more flexible manufacture schemes. As a consequence, it is often required to adapt the plant production to the demand, which can be volatile depending on the application. This is why it is important to develop tools that can monitor the condition of the process working under varying operational conditions. CVA is a multivariate data driven methodology that can be applied to detect and diagnose faults in industrial systems. This method has the ability to capture the process dynamics more efficiently than other similar data driven algorithms and its superior performance has already been demonstrated by several researchers using computer simulated data.

The aim of this study is to demonstrate the ability of CVA to detect and diagnose artificially seeded faults in a complex large scale test rig working under variable operating conditions. Chapter 3 described in detail the experimental test rig from which the data was collected for different cases. This chapter reports on the results obtained by applying the CVA to acquired data sets for fault detection and diagnosis. The CVA results are compared with other dynamic approaches to demonstrate its superior performance.

4.1 Introduction

Data driven methods are widely used for fault detection and diagnosis applications in real industrial systems. In particular, multivariate monitoring techniques such as the PCA or the PLS can take into account the correlation between the different variables measured in the process, and they show advantages against the traditional univariate methods [17]. However, there is a need for more effective techniques that can deal with problems like changing operational conditions or nonlinear systems [3; 17; 22] . Ku et al. [15] proposed

the use of lagged variables to take into account time correlation to extend PCA to dynamic system monitoring (DPCA). Similarly, a dynamic version of PLS was proposed by Komulainen [16] . Despite of their success, DPCA and DPLS have been reported not to be as efficient as other state-space based methodologies such as CVA when applied to systems working under variable loading conditions, principally due to the representation of the system dynamics [17], [18; 19].

Canonical Variate Analysis is a data driven methodology which maximizes the correlation between two sets of variables [2]. The literature gives examples of successful application of CVA for fault detection and diagnosis using computer simulated data [2; 17-22] or data acquired in small test rigs [51] or particular parts of a system [52] but there are no examples of application of CVA in real and complex systems working under varying operational conditions. The aim of this work is to demonstrate the ability of CVA to detect and diagnose artificially seeded faults in a large scale test rig similar to a real process.

This chapter shows the results obtained from the application of CVA to experimental data acquired in the three-phase flow facility at Cranfield, which is comparable to a real small multiphase flow separation process.   In this case study different faults were introduced deliberately in the system, simulating typical faults that can be expected in real plants such as blockages, incorrect system operation or non-conventional operating conditions. The data sets were acquired under changing operational conditions, modifying the flow rate set points to ensure that the fault detection was undertaken not only in the steady-state regime. The objective of this case study is to assess the performance of CVA as a method capable of detecting faults in real systems working under variable operating conditions. Additional tests were carried out using data sets acquired under steady operational conditions, and the performance of CVA for the detection and diagnosis of faults was compared with other methodologies such as PCA, DPCA, PLS and DPLS using the data sets acquired under single or multiple steady operational conditions.

4.2 Methodology

4.2.1 CVA for fault detection in industrial processes

The application procedure of CVA is similar to other multivariate algorithms for condition monitoring: it requires an initial data set to train a model, which is used to calculate the transformation matrices and the thresholds for the health indicators. In application, it is possible to monitor the process by simply converting the high-dimensional acquired data into the indicators using the transformation matrices. The value of these indicators compared with the threshold calculated during the training stage will determine the presence or absence of detected faults.

The objective of CVA is to find the linear combinations that maximize the correlation between two sets of variables. In order to take into account time correlations, the observation vector y is expanded at each time point k by considering p previous and f future measurements (each one containing m variables), generating the past and future observation vectors $y_{p,k}$ and $y_{f,k}$ respectively:

$$y_{p,k} = \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ \vdots \\ y_{k-p} \end{bmatrix} \in \Re^{mp} \qquad y_{f,k} = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+f-1} \end{bmatrix} \in \Re^{mf} \tag{4-1}$$

The data are normalized to 0 mean in each different variable to avoid domination of those variables with higher absolute values measured:

$$\hat{y}_{p,k} = y_{p,k} - \bar{y}_{p,k} \tag{4-2}$$

$$\hat{y}_{f,k} = y_{f,k} - \bar{y}_{f,k} \tag{4-3}$$

where $\bar{y}_{p,k}$ and $\bar{y}_{p,k}$ represent the sample means of $y_{p,k}$ and $y_{f,k}$ respectively. The optimal number of past and future lags (p and f) considered in the analysis can be calculated computing the autocorrelation function of the summed squares of

all measurements [17]. All the past and future vectors are arranged together in different columns generating the past and future matrices Y_p and Y_f:

$$Y_p = \left[\hat{y}_{p,p+1}, \hat{y}_{p,p+2}, ... \hat{y}_{p,p+M}\right] \in \Re^{mp \times M} \tag{4-4}$$

$$Y_f = \left[\hat{y}_{f,p+1}, \hat{y}_{f,p+2}, ... \hat{y}_{f,p+M}\right] \in \Re^{mf \times M} \tag{4-5}$$

where $M=n\text{-}f\text{-}p+1$ for a data set of n observations.

The covariance and cross-covariance matrices of past and future matrices can be estimated as follows:

$$\Sigma_{pp} = \frac{1}{M-1} Y_p Y_p^T \tag{4-6}$$

$$\Sigma_{ff} = \frac{1}{M-1} Y_f Y_f^T \tag{4-7}$$

$$\Sigma_{fp} = \frac{1}{M-1} Y_f Y_p^T \tag{4-8}$$

The correlation between two linear combinations of future and past vectors ($a^T(\hat{y}_{f,k})$ and $b^T(\hat{y}_{p,k})$) can be calculated as:

$$\rho_{fp}(a,b) = \frac{a^T \Sigma_{fp} b}{\left(a^T \Sigma_{ff} a\right)^{1/2} \left(b^T \Sigma_{pp} b\right)^{1/2}} \tag{4-9}$$

Using the variable changes $u=\Sigma^{1/2}_{ff}a$ and $v=\Sigma^{1/2}_{pp}b$ the optimization problem can be rewritten as:

$$\max_{u,v} \quad u^T \left(\Sigma^{-1/2}_{ff} \Sigma_{fp} \Sigma^{-1/2}_{pp}\right) v \tag{4-10}$$
$$s.t. \quad u^T u = 1$$
$$v^T v = 1$$

The solution u and v for the optimization problem can be calculated decomposing the scaled Hankel matrix using Singular Value Decomposition (SVD):

$$H = \Sigma^{-1/2}_{ff} \Sigma_{pf} \Sigma^{-1/2}_{pp} = UDV^T \tag{4-11}$$

where:

$$U = [u_1, u_2 \cdots u_{mf}] \in \Re^{mf \times mf} \qquad V = [v_1, v_2 \cdots v_{mp}] \in \Re^{mp \times mp}$$

$$D = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_x \end{bmatrix} \in \Re^{mf \times mp}$$

U and V are orthogonal matrices of eigenvectors and D is a diagonal matrix, indicating that U and V are only pairwise correlated. The degree of pairwise correlation between U and V is indicated by the diagonal elements γ_i in D. Reordering the elements in D ($\gamma_1 > \gamma_2 > ... > \gamma_{mp}$) and the attached eigenvectors in U and V it is possible to select the first r columns of V which best correlate U and V, generating a new dimensionally reduced matrix V_r.

The transformation matrices J and L which convert the $m \cdot p$-dimensional past measurements to the r-dimensional canonical variates and residuals can be calculated as:

$$J = V_r^T \Sigma_{pp}^{-1/2} \tag{4-12}$$

$$L = (I - V_r \cdot V_r^T) \Sigma_{pp}^{-1/2} \tag{4-13}$$

The canonical variates z and the residuals ε are calculated projecting the acquired data into the low-dimensional space:

$$z = J \cdot Y_p \tag{4-14}$$

$$\varepsilon = L \cdot Y_p \tag{4-15}$$

The statistical indicators that provide information about the health of the system can be obtained converting the available past observations into a lower dimensional data. The statistical indicators more frequently used are the Hotelling T^2 statistic and Q statistic.

$$T_j^2 = \sum_{i=1}^{r} z_{i,j}^2 \tag{4-16}$$

$$Q_j = \sum_{i=1}^{r} \varepsilon_{i,j}^2 \tag{4-17}$$

The upper control limits (UCL) for T^2 and Q can be calculated for a given significance level α such that $P(T^2 > T^2_{UCL}(\alpha)) = \alpha$ and $P(Q > Q_{UCL}(\alpha)) = \alpha$ respectively. Normally these control limits are calculated assuming that the probability density functions of the T^2 and Q statistics are Gaussian, but system nonlinearities can derive into modelling errors which are not Gaussian, making this assumption invalid.

Odiowei and Cao [17] developed a methodology to solve this issue by estimating the actual probability density function of the statistical indicators using Kernel Density Estimations (KDE). The probability of a random variable y (with a probability density function $p(y)$) to be smaller than a certain value b is defined as:

$$P(y < b) = \int_{-\infty}^{b} p(y)dy \tag{4-18}$$

Where $p(y)$ can be calculated through the kernel function K:

$$p(y) = \frac{1}{Mh} \sum_{k=1}^{M} K\left(\frac{y - y_k}{h}\right) \tag{4-19}$$

$$K(z) = e^{-z^2/2} \Big/ \sqrt{2\pi} \tag{4-20}$$

Where h is the selected bandwidth (see [17]) and y_k is each one of the M samples in y.

The T^2 metric represents the variation of the state variables indicating changes in the retained space. Q is also known as the squared prediction error and it represents the sum of the squared variation error in the residual space. Both indicators are complementary; some faults can cause an increment in the variability of the system states whilst others will be manifested as an increment in the residual space variability. For this case study the event of fault detection will be considered every time any of the indicators exceed the respective UCL. This makes the monitoring performance insensitive to the number of states retained r [17].

In the event of fault detection it is crucial to locate the source of the fault in order to understand the origin of the problem and undertake the best maintenance

action. Chiang et al. [2] suggested the use of contribution plots to estimate how much each one of the variables contributed to the final value of the statistical indicator. In CVA it is easy to calculate the contribution c of each variable y_j to the final value of each of the canonical variates z_i:

$$c_{i,j} = \frac{z_i}{\gamma_i} J_{i,j} \cdot \hat{y}_j \tag{4-21}$$

where γ_i is the singular value corresponding to the loading vector J_i. The total contribution C of the j_{th} process variable y_j can be obtained as:

$$C_j = \sum_{i=1}^{r} c_{i,j} \tag{4-22}$$

This technique has been successfully applied by other researchers for fault identification [80-83]. Once a fault has been detected, it is possible to prioritize the variables responsible for fault identification based on their individual contributions, and the plant engineers can use this information together with their plant knowledge to determine the origin of the fault.

4.2.2 Experimental set up

The data sets used in this investigation were acquired from the 3 phase flow facility at Cranfield University, which is a large scale test rig for multiphase flow research designed to provide a controlled and measured amount of air oil and water. All the data were captured by the Emerson Delta V Digital Automation System at a sampling rate of 1 Hz. Mixture of only air and water was used in all the experiments, and the three phase separator was always pressurized to 1.0 barg. The rig description and a detailed   explanation of the data sets acquired and the faults introduced were presented in chapter 3. Three data sets (T1, T2 and T3) were acquired from the system working under normal operating conditions (no faults). In these data sets the air and water flow rate set point was modified to obtain data from the system working under different conditions, trying to cover all the available spectrum of flow rate combinations. Additionally, six different faults were introduced in the system to study the capabilities of different monitoring algorithms in terms of fault detection and diagnosis. The faults introduced were summarized in Table 3. These faults were introduced

gradually when possible so as to observe how the severity of the fault affects the health indicators. The faults were introduced after a period of normal operation to visualize the change in the indicators from normal operation to faulty conditions. After reaching the maximum fault severity, the faults were removed from the test rig allowing the operation to return to normal operational conditions. For each one of the cases (except Case 6) 2 different data sets were analysed using CVA in order to observe how the algorithm performs under different conditions. Table 1 summarized the different measurements acquired from the system, although variable 24 was only included in the analysis of Case 6.

4.3 Results and discussion

The results provided by the application of CVA to the training and monitoring data sets introduced above are presented in this section. Each one of the faults is explained briefly again here, but additional details can be found in chapter 3 of this document. For each data set analysed the information provided here includes the results obtained from the CVA application: T^2 and Q indicators and contribution plots at the moment of fault detection. The plots representing CVA fault detection results for each analysis contain the T^2 and Q indicators plotted as solid black lines, while the threshold is plotted as a grey dashed line. The fault starting and ending points in each case are represented by a vertical dashed black line. It is important to notice that all the faults cases except Case 6 were introduced gradually, and the fault severity increases with time after the introduction of the fault.

4.3.1 Training data sets and selection of tuning parameters

The first step for the application of CVA to the "normal operation" data sets in order to train the algorithm and obtain the transformation matrices and UCL for the thresholds is to select the number of past and future lags considered (p and f) and the number of estates retained (r). The optimal number of past and future lags considered in the analysis can be calculated computing the autocorrelation function of the summed squares of all measurements [17]. This function measures the cross correlation between a signal and a delayed version of itself at different lags. In this way it is possible to determine for how long the correlation of the signal with past lags is significant, and thus only the lags which are relevant are selected. Fig. 38 shows an example of autocorrelation function for the training data set T1 against a confidence bound of ±5% (solid black line).
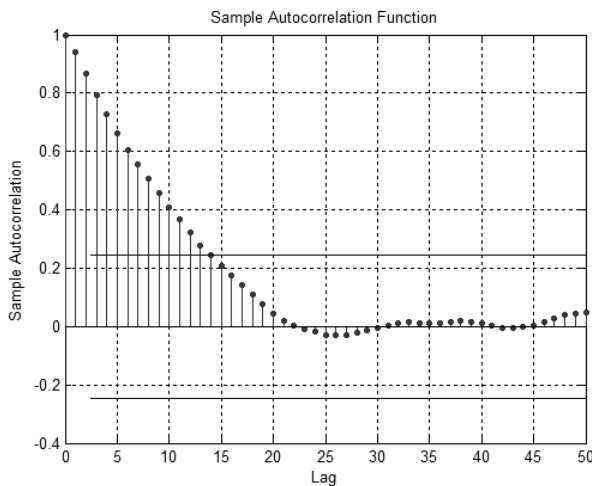
Fig. 38: Autocorrelation function of the summed squares of all measurements for data set T1

For this study p and f were set to 15 according to the results obtained from the analysis of the autocorrelation function of the three training data sets.

Different methodologies have been suggested for the calculation of the optimal number of states retained r, among which those based on considering the dominant singular values in the matrix D [84] and methodologies based on the Akaike Information Criterion (AIC) [2] are the most common.

Fig. 39 shows the normalized singular values obtained from (4-11). In this particular case, the singular values decrease slowly and setting the number of retained states based on the dominant singular values will derive in an unrealistic model [17]. In addition, the number of states retained is not especially relevant for this study because both statistical indicators (T^2 and Q) are used at the same time for fault detection. It means that those system variations not captured in the retained space will be captured by the residual space and vice versa. A confidence bound of 99% was considered for the calculation of the UCL.
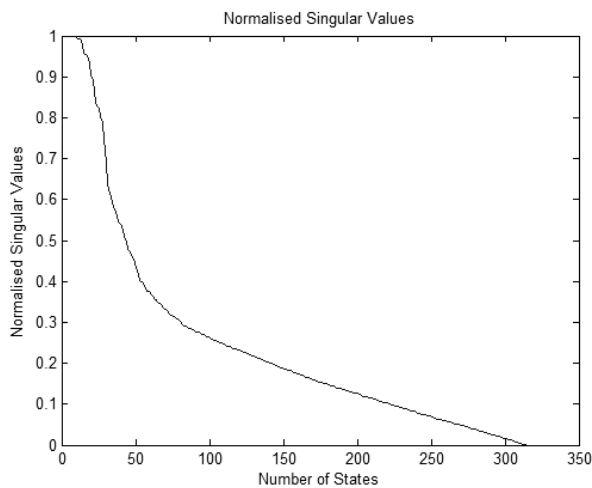
Fig. 39: Normalized singular values for T1

In order to obtain a richer and more varied training set that covers all the spectrum of operational conditions, the data sets acquired under normal operational conditions were combined. The combination was done by calculating the past and future matrices individually for each data set according to (4-4) and (4-5) and then the matrices obtained were merged. The original length of the data sets was 10372 s for T1, 9825 s for T2 and 13200 s for T3. In order to check the capacity of these data sets to represent the system dynamics accurately producing a low number of false alarms, the three data sets were mixed in pairs generating three different combined sets. The objective of analysing these three combinations is to see which of them produces a lower false alarm rate when the remaining data set is used during the monitoring

period. In order to select the optimal number of dimensions retained r, CVA was performed for each one of the three combined training data sets using a range of values for this parameter. For low values of r the number of false alarms is high because the retained space is not able to represent accurately the states of the system and consequently the number of the T^2 threshold violations increases. On the other hand if the state order selected is too high it results in the model underfitting the data [2], increasing again the false alarm rate. The training data set combination which produced lower false alarm rates was T2 and T3. After several analyses testing different values for r for each combination of data sets finally $r=25$ was adopted in order to minimize the false alarm rate in normal conditions (see Fig. 40).
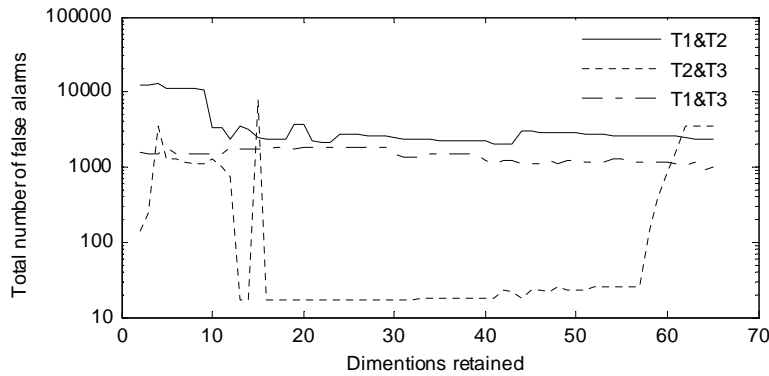
Fig. 40: Analysis of the influence of the number of states retained

The T^2 and Q plots obtained for the different combined training sets using the selected number of states retained are represented in Fig. 41, Fig. 42 and Fig. 43 including the UCL calculated using KDE for a 99% confidence bound plotted as a grey dashed line.
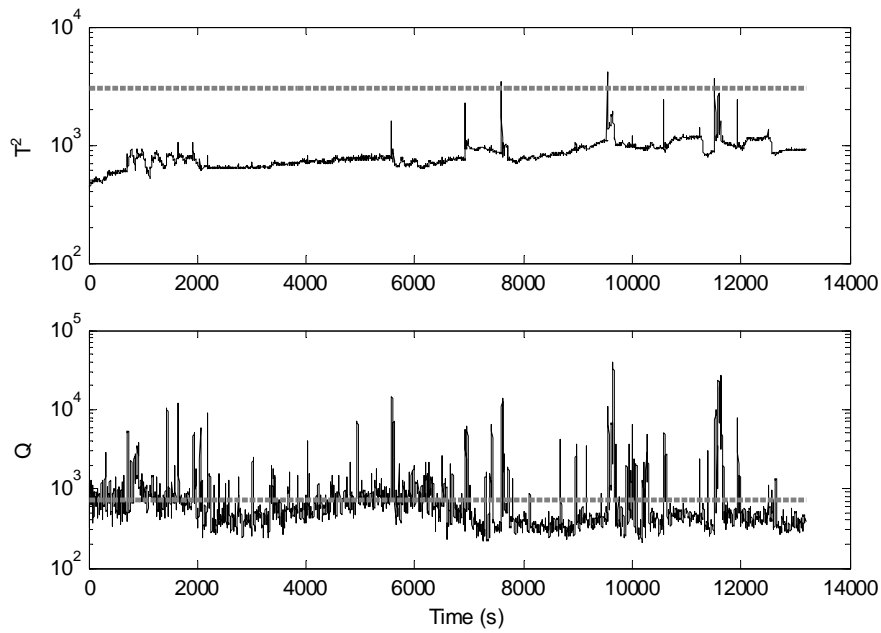
Fig. 41: T^2 and Q statistics for data set T3 using T1 and T2 combination for training
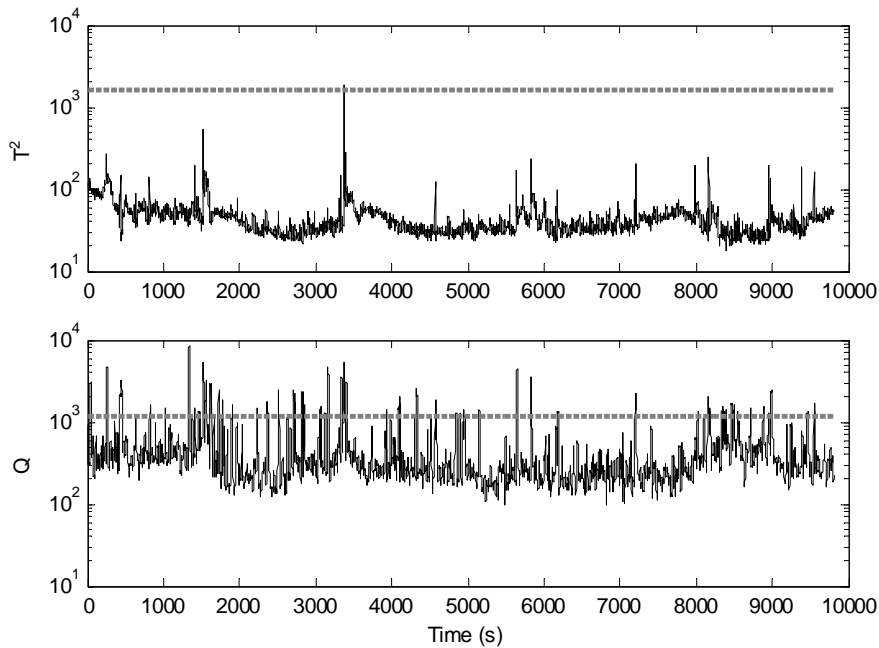
Fig. 42: T^2 and Q statistics for data set T2 using T1 and T3 combination for training
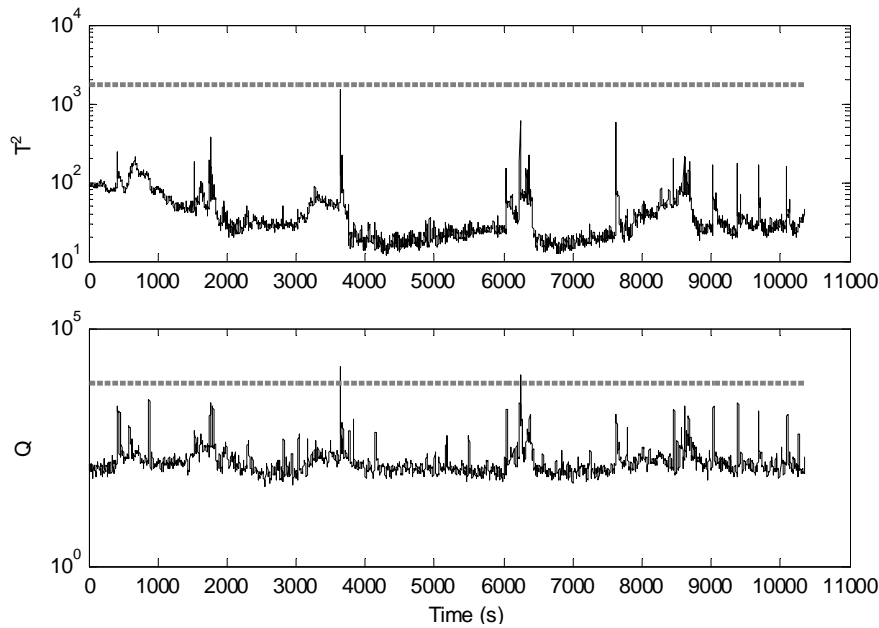
Fig. 43: T^2 **and** Q **statistics for data set T1 using T2 and T3 combination for training**

Table 11 summarises the results obtained from this analysis for the parameter *r=25*:

Table 11: Summary training data sets analysis

Training	Monitoring	T^2 Threshold	Q threshold	False alarm rate T^2	False alarm rate Q
T1&T2	T3	3036.81	734.94	0.0378%	29.1158%
T1&T3	T2	1620.63	1158.71	0.0102%	7.9601%
T2&T3	T1	1753.97	6940.73	0%	0.2121%

This results show that the combination of T1 and T2 for training is not able to capture accurately all the systems variations that occurred during T3, obtaining a high number of points over the limit especially for the Q indicator. The number of points over the threshold using the combination T1 and T3 for the training stage and T2 for monitoring produced slightly better results, but it is obvious that the combination which produced the best results capturing the system dynamics and consequently reducing the number of threshold violations was T2 and T3. For this reason, in the analysis of the different data sets acquired under faulty operating conditions the transformation matrices and indicators thresholds obtained using this combination of data sets will be used. The length of the

91

combined training data set is 22967 s, taking into account that after the calculation of the past and future matrices the length of each data set is reduced by $p\text{-}f\text{+}1$ samples. Fig. 42 shows the probability density functions obtained applying KDE to the T^2 and Q statistic values calculated for the data set combination T2 and T3.
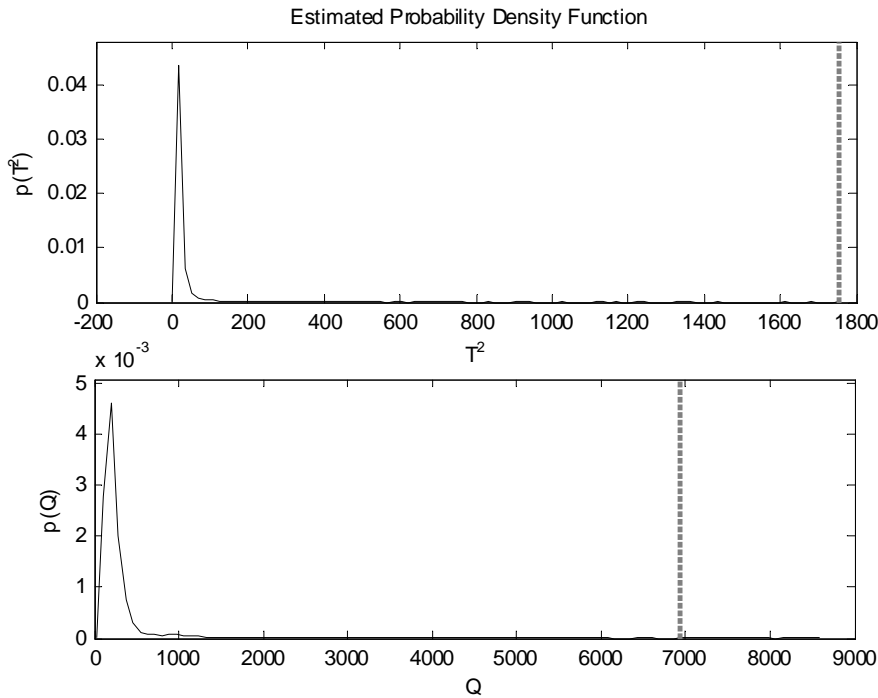
Fig. 44: Probability density functions of T^2 and Q statistics during training (99% confidence bound)

The UCL calculated for the T^2 and Q statistics was 1753.98 and 6940.73 respectively.

4.3.2 Case 1: Air line blockage

The fault introduced in this case was to gradually close the manual valve of the air line simulating a blockage that develops over time. Two data sets were analysed, identified in chapter 3 as data sets 1.1 and 1.2. In data set 1.1 the fault was introduced while the system was operated in changing loading conditions but in data set 1.2 the air and water flow rates set points were constant, 150m^3/hr of air and 2kg/s of water.

The results obtained in terms of fault detection and diagnosis for data set 1.1 are represented in Fig. 45. The first fault detection occurred at sample 2992 for the T^2 statistical indicator and in sample 2985 for the Q indicator. Those samples correspond to a valve opening of 60° in both cases. Both indicators fall below the UCL when the fault is removed opening the air valve completely. The contribution plots at the fault detection time (sample 2985) are represented in Fig. 45(b). In this case the variable contributing more to the final value of the T^2 statistical indicator is the air delivery pressure (PT312), pointing at a conflict with the pressure in the air line section before the mixing point. The Q indicator is not as clear as the T^2 for this test, and other variables like the pressure at the bottom of the riser (PT401) or the differential pressure between the top and the bottom of the riser (PT401-PT408) are more significant.
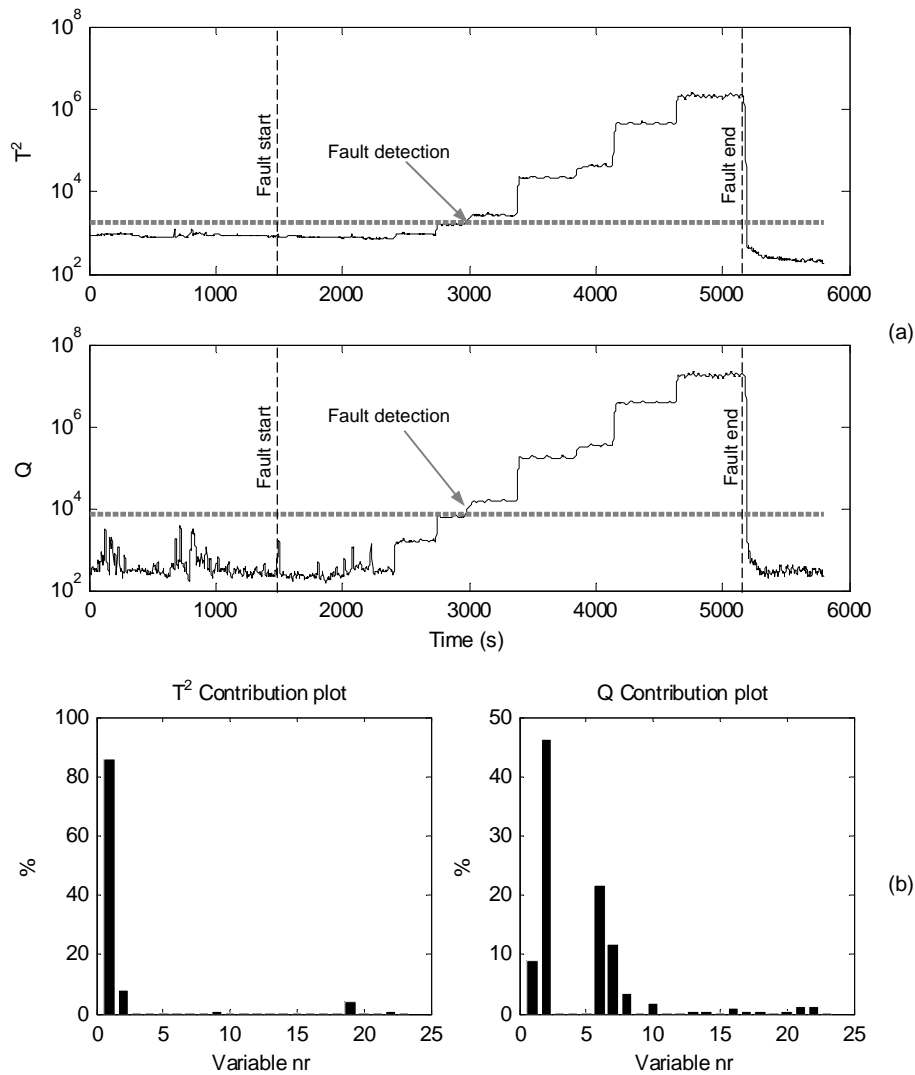
Fig. 45: Results from data set 1.1: T^2 and Q indicators (a) and contribution plots at sample 2985 (b)

The results obtained in terms of fault detection and diagnosis for data set 1.2 are represented in Fig. 46. For this data set the first fault detection happened in sample 2163 for the T^2 statistical indicator and in sample 1870 for the Q indicator without short false alarms in any case. Those samples correspond to a valve opening of 35° and 40° respectively. Both indicators fall below the UCL when the fault is removed opening the air valve completely. The contribution plots at the fault detection time (sample 1870) are represented in Fig. 46(b). The variable contributing more to the final value of the T^2 statistical indicator is again the air delivery pressure PT312. The Q indicator is not as clear as the T^2,

and other variables like the pressure at the bottom of the riser (PT401) and the differential pressure between the top and the bottom of the riser (PT401-PT408) are more significant.
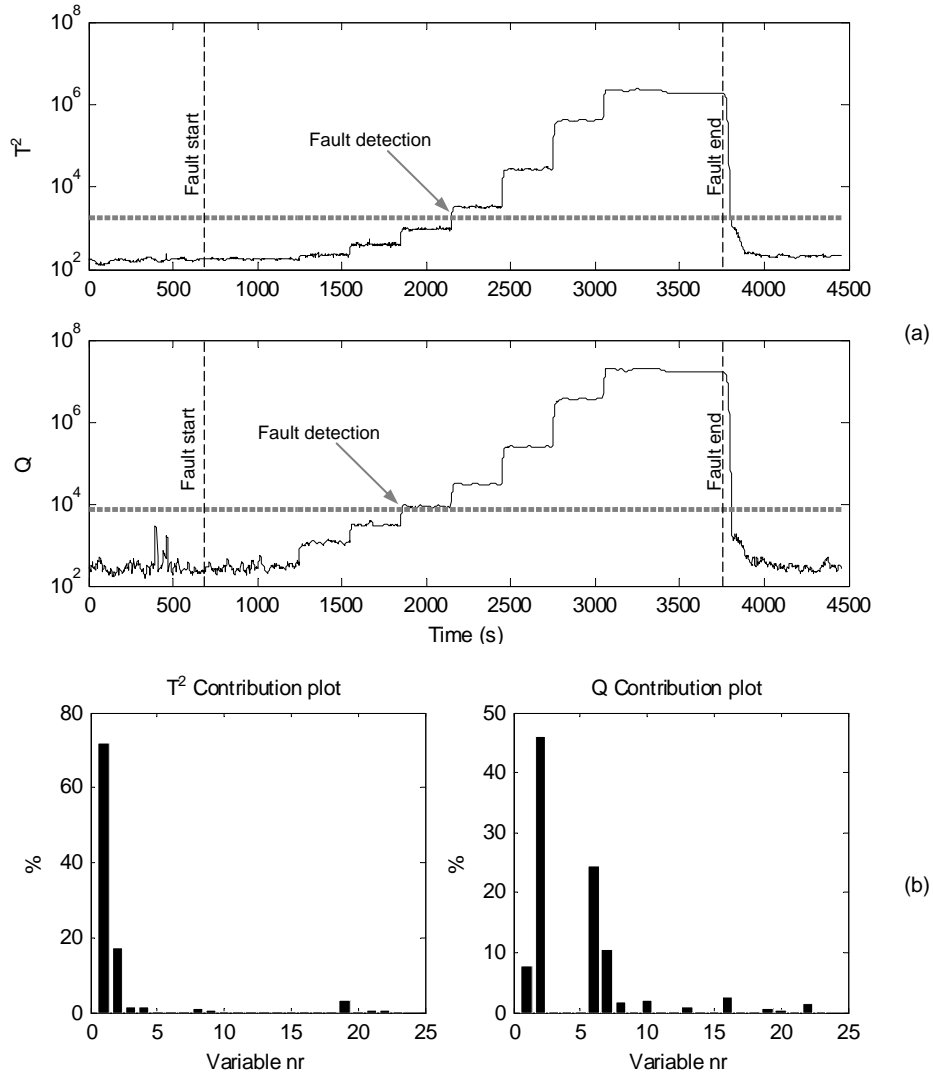
Fig. 46: Results from data set 1.2: T^2 and Q indicators (a) and contribution plots at sample 1870 (b)

4.3.3 Case 2: Water line blockage

Similar to the fault case 1, in fault case 2, the water line valve was closed gradually to simulate the blockage instead of the air line valve. Two data sets were analysed, identified in chapter 3 as data sets 2.1 and 2.2. In data set 2.1 the fault was introduced while the system was operated in changing loading conditions and in data set 2.2 the air and water flow rates set points were

constant, 100m³/hr of air and 2kg/s of water. The results obtained in terms of fault detection and diagnosis for data set 2.1 are represented in Fig. 47.
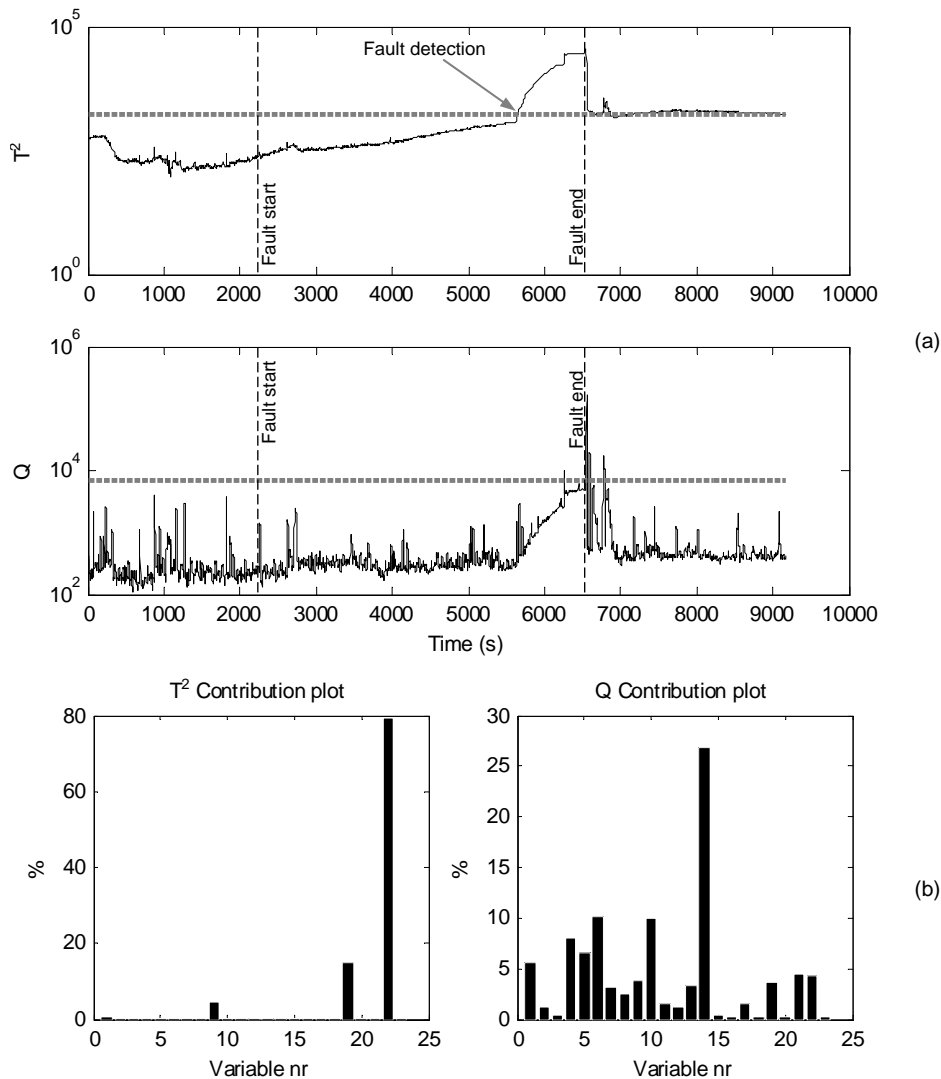
Fig. 47: Results from data set 2.1: T^2 and Q indicators (a) and contribution plots at sample 5660(b)

In this case the first fault detection happened in sample 5660 for the T^2 statistical indicator, while in the Q indicator there is no clear sign of fault detection apart from 3 short alarms around sample 6600. Sample 5660 corresponds to a valve opening of 25°. It is important to notice that in this particular case the water flow rate was seriously affected by the fault introduced instants after the fault detection, which gives a reduced amount of reaction time before the fault becomes critical. The T^2 indicator falls below the UCL when the fault is removed opening the water valve completely but generates several short

false alarms before the end of the test. The poor performance of the Q indicator in this particular fault is attributed to the threshold value calculated: observations of the Q statistic seem to indicate it may have been possible to detect the fault using a lower threshold. The contribution plots at the fault detection time (sample 5660) are represented in Fig. 47 (b). In this case the variable contributing more to the final value of the T^2 statistical indicator is the position of the valve VC101 which corresponds to the valve that controls the water flow rate. This is an indication that the control loop is reacting to the increment in the pressure losses generated by the blockage by increasing the valve opening to keep the desired flow rate. For the Q indicator the main contributions are split between different variables, which makes it difficult to identify the source of the fault.

The results obtained in terms of fault detection and diagnosis for data set 2.2 are represented in Fig. 48. In this case the first fault detection happened in sample 2288 for the T^2 statistical indicator, while in the Q indicator again there is no clear sign of fault detection apart from 3 short alarms around sample 2700. Sample 2288 corresponds to a valve opening of 20°. The T^2 indicator falls below the UCL when the fault is removed opening the water valve completely. For the Q statistic it could have been possible to detect the fault using a lower threshold. The contribution plots at the fault detection time (sample 2288) are represented in Fig. 48 (b). In this case the variable contributing more to the final value of the T^2 statistical indicator is again the position of the valve VC101. For the Q indicator the main contributions are split between different variables, which makes it difficult to identify the source of the fault. The repeated lack of precision of the Q indicator locating the source of the fault and the high number of false alarms can be an indication of oversensitivity in the residual space.
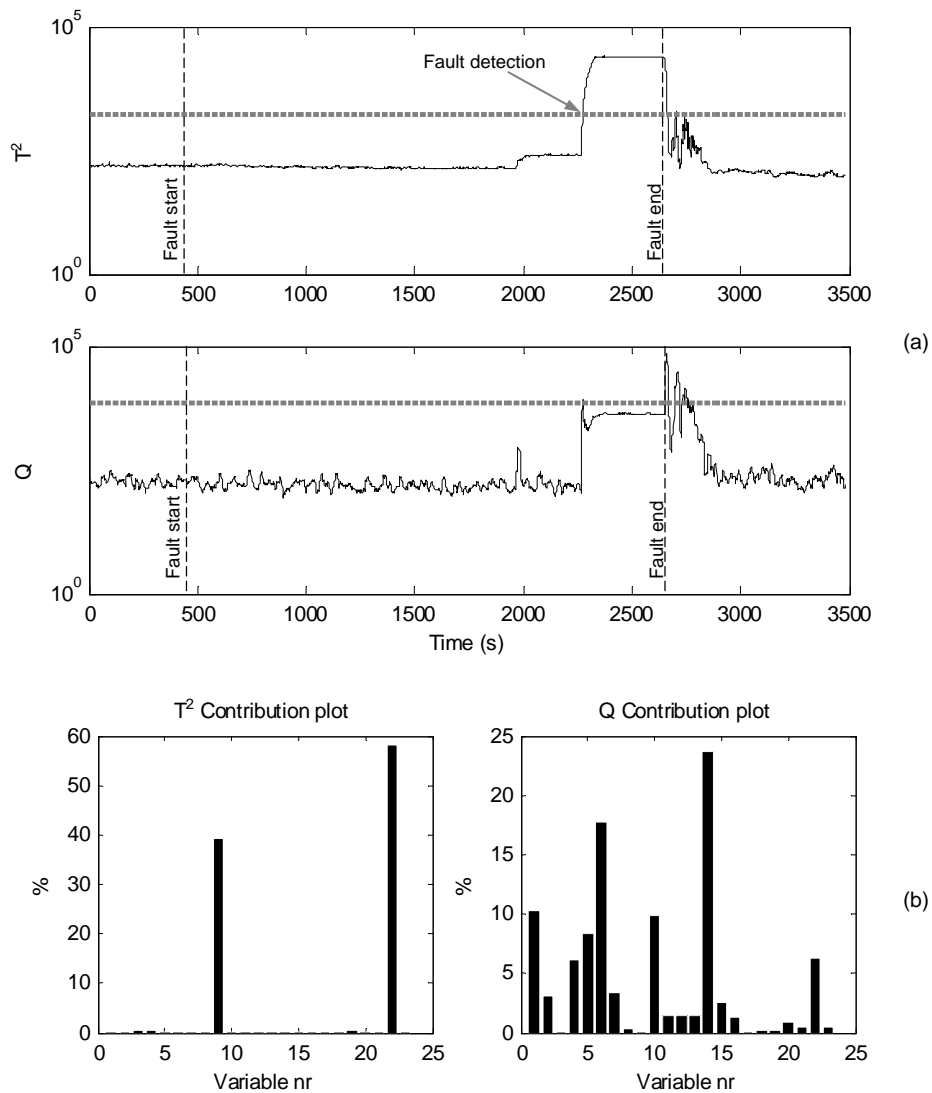
Fig. 48: Results from data set 2.2: T^2 and Q indicators (a) and contribution plots at sample 2288 (b)

4.3.4 Case 3: Top separator input blockage

In this case the blockage was introduced manipulating the valve VC404 (top separator input) which is operated pneumatically, offering higher precision in the measurement of the valve position. Two data sets were analysed, identified in chapter 3 as data sets 3.1 and 3.2. In data set 3.1 the fault was introduced while the system was operated in changing loading conditions. In data set 3.2 the air and water flow rates set points were constant, 100m³/hr of air and 2kg/s of water. The results obtained in terms of fault detection and diagnosis for data set 3.1 are represented in Fig. 49.
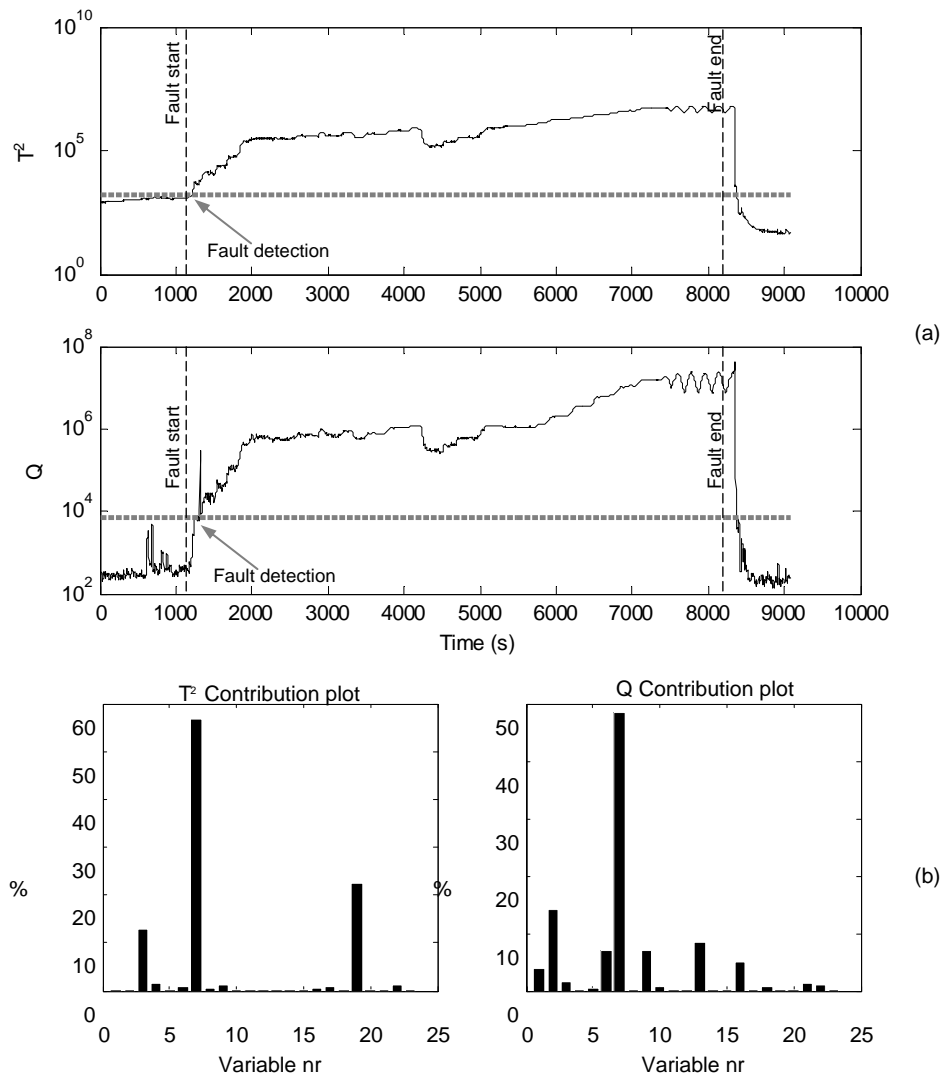
Fig. 49: Results from data set 3.1: T^2 **and** Q **indicators (a) and contribution plots at sample 1230 (b)**

The first fault detection happens in sample 1230 for the T^2 statistical indicator and 1324 for the Q indicator. These time points correspond to a valve opening of 60% and 45% respectively. Both indicators fall below the UCL values when the fault is removed opening the valve completely. In this case the recovery to normal conditions takes slightly longer than in the previous cases because the air and water flow rates were seriously affected in the last stages of the fault. The contribution plots at the fault detection time (sample 1230) are represented in Fig. 49(b). In this case the variable contributing more to the final value of the T^2 statistical indicator is the differential pressure over the valve VC404 (PT408-PT403), pointing precisely to an excessive pressure loss in this valve. It is

important to mention that the position of the valve VC404 was not included in the analysis. The contribution of the level in the 3 phase separator (LI504) is also significant. For the Q indicator, the most significant variables are the differential pressure over the valve VC404 (PT408-PT403), and the pressure measured at the bottom of the riser (PT401).

The results obtained in terms of fault detection and diagnosis in data set 3.2 are represented in Fig. 50.
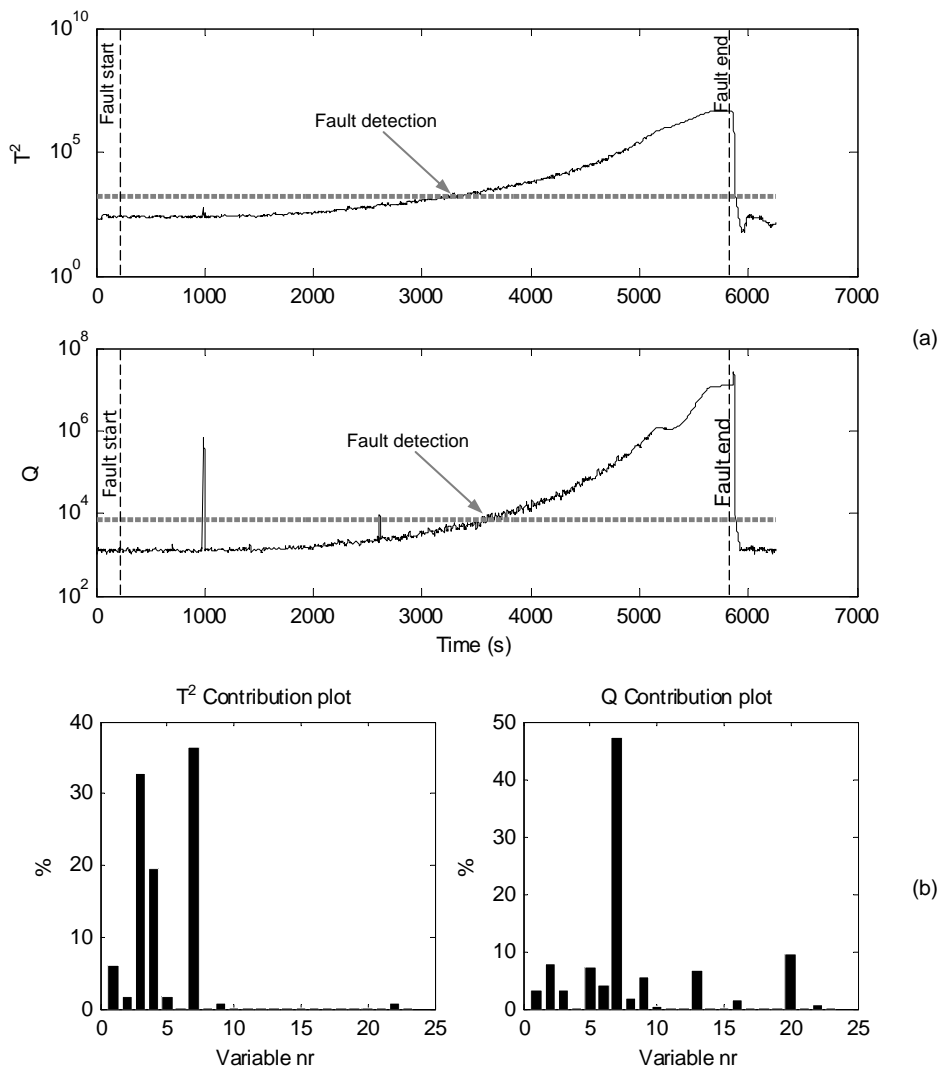
Fig. 50: Results from data set 3.2: T^2 and Q indicators (a) and contribution plots at sample 3419 (b)

The first fault detection happens in sample 3419 for the T^2 statistical indicator after 8 short fault alarms (difficult to see due to the logarithmic scale) while for

the Q indicator the detection happens at sample 3704 after 8 short false alarms. Sample number 3419 corresponds to a valve opening of 48% and sample 3704 corresponds to 43%. Both indicators fall below the UCL values when the fault is removed opening the valve completely. The contribution plots at the fault detection time (sample 3419) are represented in Fig. 50 (b). In this case the variables contributing more to the final value of the T^2 statistical indicator are the differential pressure over the valve VC404 (PT408-PT403) and the pressure at the top of the riser (PT408). For the Q indicator the most significant variables are the differential pressure over the valve VC404 and the position of the air outlet valve in the 3 phase separator.

4.3.5 Case 4: Open direct bypass

In the rig there is an alternative 4" line which can carry the flow directly from the mixing point to the 3 phase separator, bypassing the riser and the top separator. There are valves at the beginning and at the end of this alternative line to isolate it from the rest of the system, and in normal conditions they are closed. The fault introduced in this case consisted in the gradual opening of this valves, simulating a leakage where part of the flow is lost or a wrong system operation. Two data sets were analysed, identified in chapter 3 as data sets 4.1 and 4.2. In data set 4.1 the fault was introduced while the system was operated under changing loading conditions. In data set 4.2 the air and water flow rates set points were constant, 150m^3/hr of air and 2kg/s of water. The results obtained in terms of fault detection and diagnosis for fault 4.1 are represented in Fig. 51.

The first fault detection occurs at sample 1501 for the T^2 statistical indicator and 1531 for the Q indicator. Those time points correspond to a valve opening of 10° in both cases. The behaviour of the indicators in this case is completely different from the behaviour observed in the previous cases. The T^2 indicator shows an oscillating behaviour from the point of fault detection. Secondly both indicators reach a maximum value around sample 3500 and next their value decreases despite the fault severity continuously increased until sample 6293.
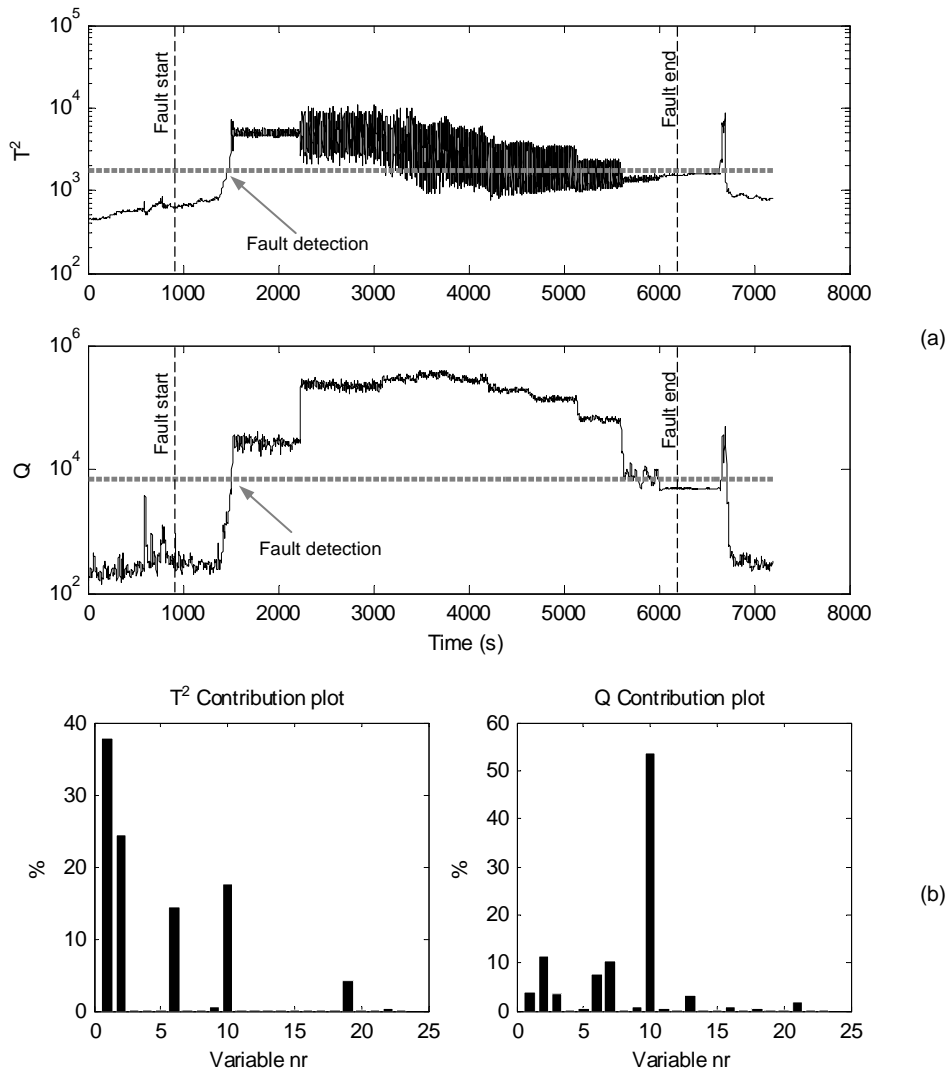
Fig. 51: Results from data set 4.1: T^2 and Q indicators (a) and contribution plots at sample 1501 (b)

The introduction of a leakage at the bottom of the riser generated slugging (which will be studied in the next fault case) due to the geometry of the bypass valve manipulated to simulate this leakage. As it can be seen in Fig. 52, the derivation for the bypass pipe line is located at the top part of the 4" line, situating the bypass line around 45cm higher than the main 4"line. When the bypass valve was initially opened, the gas flowing inside the 4" line on top of the water escaped through this alternative line, which affected the multiphase flow through the riser. The water remained at the bottom of the riser generating a blockage which reached the bypass valve region, increasing the pressure in all

the pipe lines before that point. When the pressure was sufficiently high as to flush the liquid through the 4" line and the bypass line the water was discharged, clearing the blockage and restarting the cycle again. This explains the oscillations in the T^2 indicator as a consequence of the oscillations in the pressure measured in the air line (PT312) and the pressure at the bottom of the riser PT401. Fig. 53 shows the effect generated by the fault over the pressure in the air supply line (PT312) and its consequence in the measured flow rate at the top of the riser (FT407).
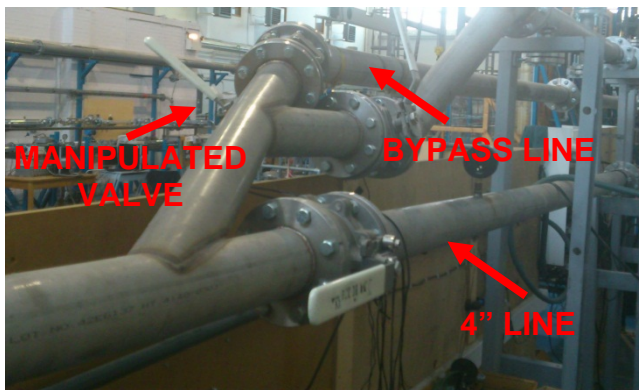
Fig. 52: Bypass derivation detail
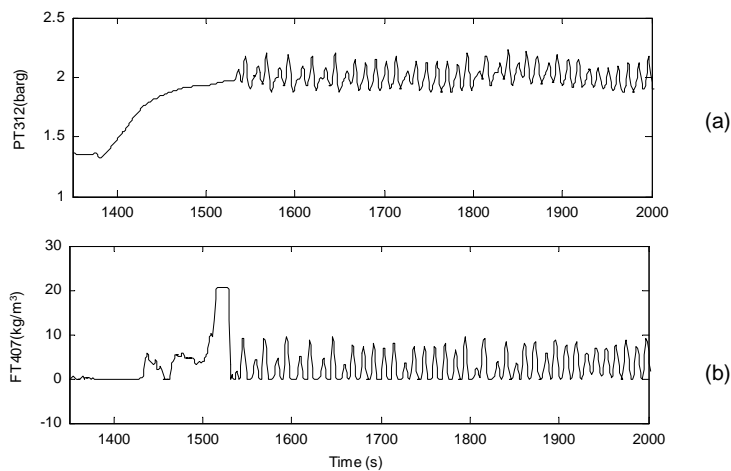
Fig. 53: Effect of fault 4 on the air supply pressure (a) and top riser flow rate (b)

As the bypass valve is opened, a higher percentage of the flow is carried through this alternative pipe line. This mitigates the liquid blocking effect and reduces the severity of the slugging. That is the reason why both indicators decrease after reaching a maximum around sample 3500 despite the fault continuously increased in severity. The contribution plot for the T^2 indicator at

the detection time (sample 1501) shows a higher contribution of the air supply line pressure (PT312) and the pressure measured at the bottom of the riser (PT401) due to the effect of the liquid blockage. For the Q indicator the most important contribution comes from the flow rate measured at the top of the riser (FT407), due to the lack of flow rate caused by the derivation of part of the air and water through the bypass line and the liquid blockage at the bottom of the riser.

The results obtained in terms of fault detection and diagnosis for fault 4.2 are represented in Fig. 54.
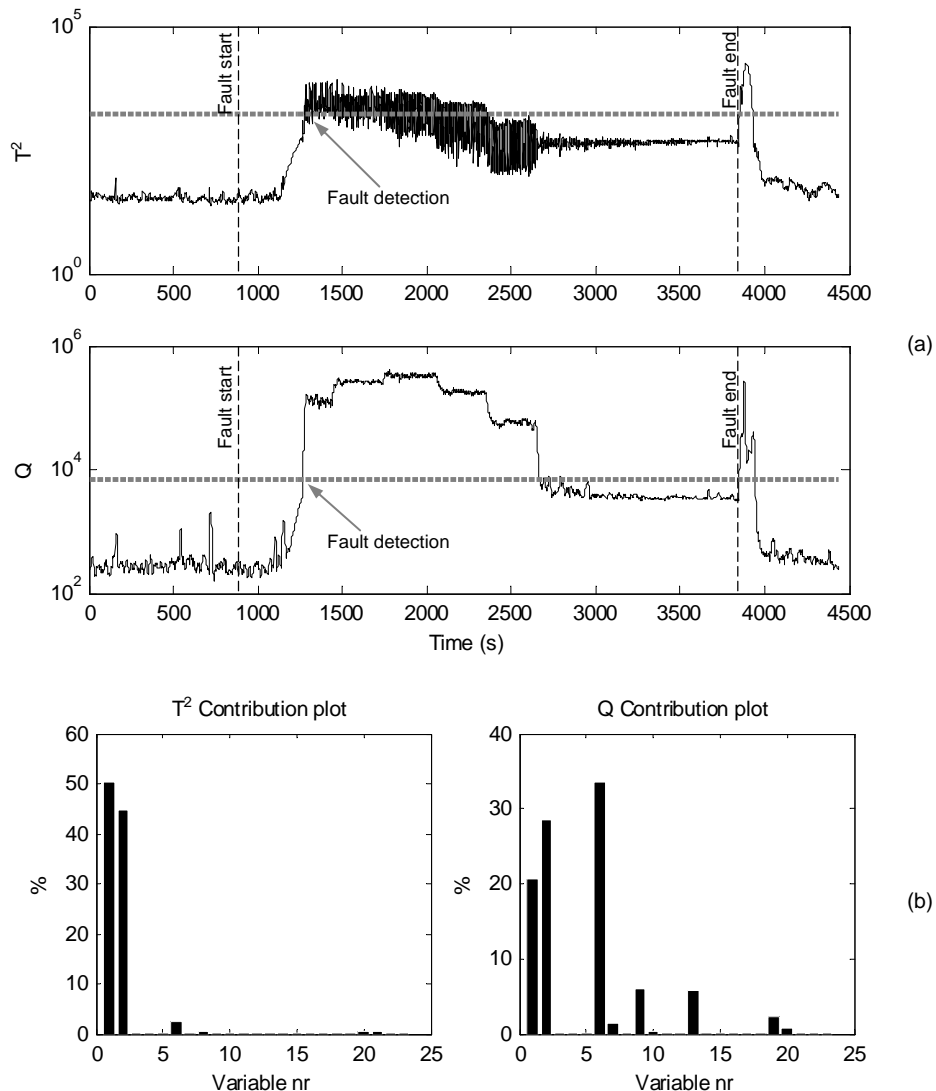
Fig. 54: Results from data set 4.2: T^2 and Q indicators (a) and contribution plots at sample 1281(b)

The first fault detection occurs at sample 1287 for the T^2 statistical indicator and 1281 for the Q indicator. These time points correspond to a valve opening of 10° in both cases. The behaviour of the indicators in the same observed in data set 4.1, the T^2 indicator oscillates after the point of fault detection and both indicators reach a maximum value after which their value decreases despite the fault severity keeps increasing. The contribution plot for the T^2 indicator at the detection time (sample 1281) shows a higher contribution of the air supply line pressure (PT312) and the pressure measured at the bottom of the riser (PT401) due to the effect of the liquid blockage. For the Q indicator the most significant variables are the differential pressure between the top and the bottom of the riser, the bottom riser pressure and the air delivery pressure.

4.3.6 Case 5: Slugging conditions

Slugging is a transient phenomenon that can occur in risers with multiphase flow when the speed of the gas and the liquid are relatively small. The liquid tends to accumulate in the base of the riser blocking the flow. Due to this blockage, the pressure rises until it is sufficient to flush the liquid (and the gas) out of the riser. After this surge, the liquid remaining in the riser falls down, creating a new blockage and starting the cycle again. The fault was introduced by reducing the air and water flow rates to regimes where slugging is produced. During the tests, the flow rate was varying at different points from normal to slugging conditions. Two data sets were analysed, identified chapter 3 as data sets 5.1 and 5.2. Fig. 55 represents the operational conditions selected for data set 5.1. The flow rate combinations where slugging is expected are shaded in grey.
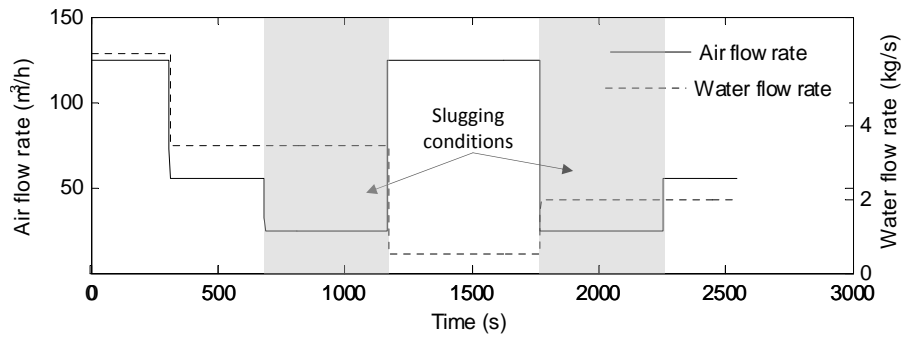
Fig. 55: Flow rate set points for data set 5.1

The results obtained in terms of fault detection and diagnosis for data set 5.1 are represented in Fig. 56.
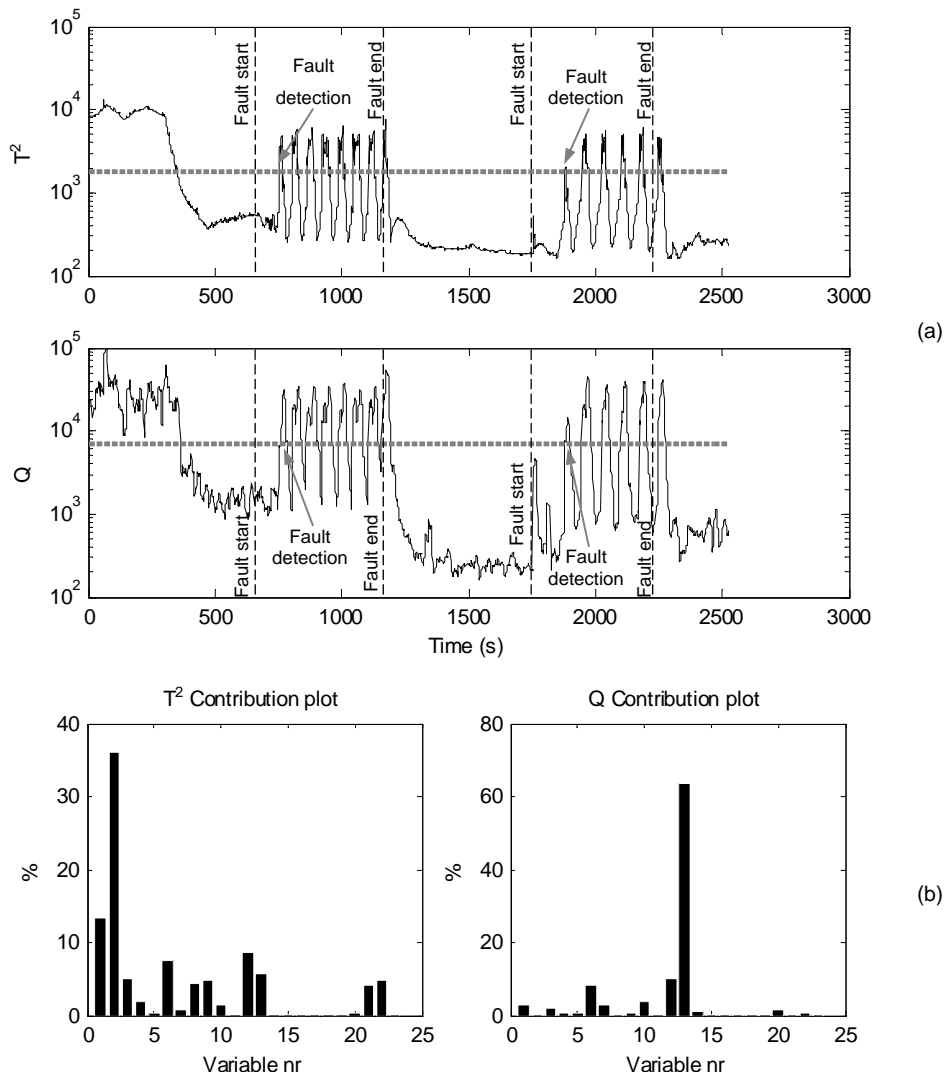
Fig. 56: Results from data set 5.1: T^2 and Q indicators (a) and contribution plots at sample 769(b)

In this data set both indicators are above the threshold at the beginning of the experiment. This was not an expected result as the flow rates selected at the beginning of the experiment ($125m^3$/hr of air and 6kg/s of water) are not supposed to produce slugging and are already included in the training data sets. The value of both indicators fluctuates with time during slugging conditions but the threshold is only reached when slugging is produced. The fault detection produced by the introduction of slugging conditions happens at samples 769 and 1895 for the T^2 statistic and samples 773 and 1895 for Q. In this case the variables contributing most to the final value of the T^2 statistical indicator for the first detection caused by slugging (sample 769) are the differential pressure between the top and the bottom of the riser and the air delivery pressure (PT312). The most significant variable for the value of the Q indicator at this time point is the density measured at the top of the riser (FT407). As it can be seen in Fig. 57 which represents the pressure measured at the bottom of the riser during the experiment (PT401), slugging was not produced at the beginning of the experiment but the values observed are significantly high. The fluctuations observed in the T^2 and Q indicators correspond with the pressure fluctuations generated by the slugging effect, returning momentarily to normal values after each discharge.
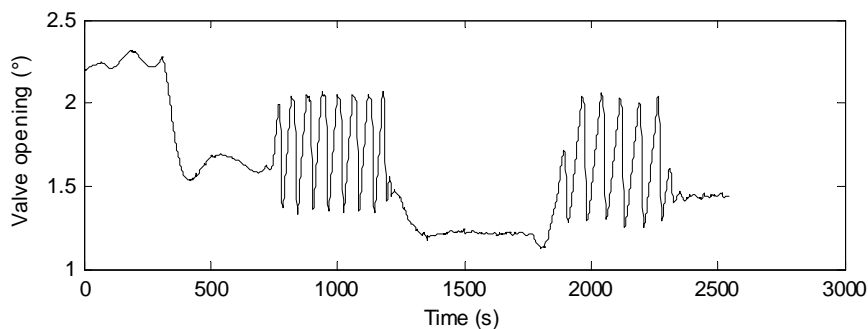
Fig. 57: Riser bottom pressure (PT401) for data set 5.1

Fig. 58 represents the operational conditions selected for data set 5.2.

107
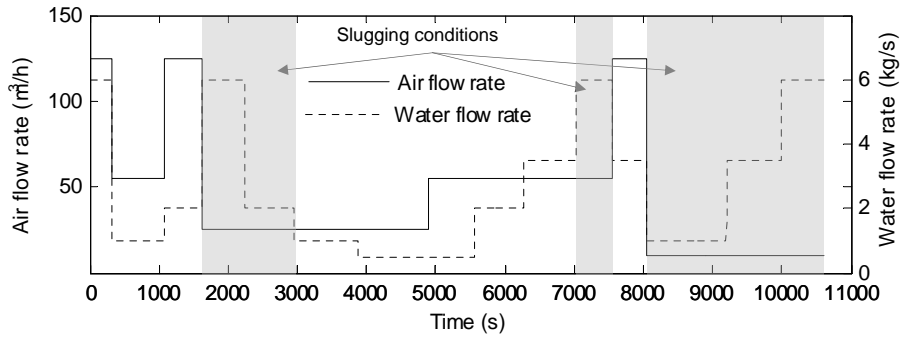
Fig. 58: Flow rate set points   for data set 5.2

The results obtained in terms of fault detection and diagnosis for data set 5.2 are represented in Fig. 59.
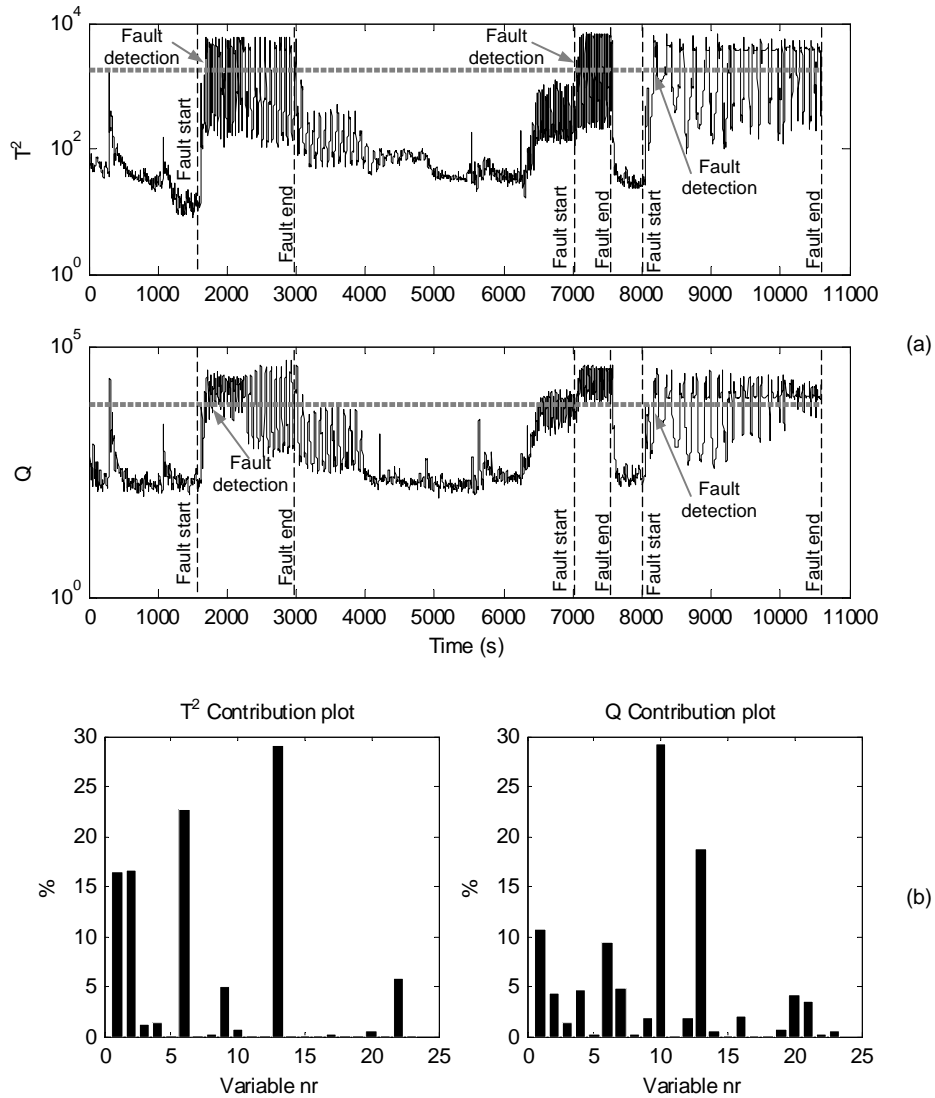
Fig. 59: Results from data set 5.2: T^2 and Q indicators (a) and contribution plots (b)

The fault detection caused by slugging happens at samples 1716, 7075 and 8171 for the T^2 statistical indicator and 1643 and 8103 for the Q indicator after one short false alarm. The detection of the second period of slugging is complicated in the Q indicator because it was already over the threshold limit before the introduction of slugging, where the flow rate conditions were in a transition zone between slugging and normal. In this particular case the value of both indicators fluctuates with time during slugging conditions due to the repetitive nature of the phenomenon, but the threshold is only reached during slugging conditions and also for transition zones in case of the Q indicator. In this case the variables contributing more to the final value of the T^2 statistical indicator at the first detection point (sample 1643) are the density of the fluid at the top of the riser, differential pressure between the top and the bottom of the riser, and the bottom riser pressure (PT401). These are the key variables normally used to characterise slugging. The most significant variables for the value of the Q indicator at this time point are the flow rate at the top of the riser and the density measured in the same point.

4.3.7 Case 6: Pressurization of the 2" line

In the previous cases the 4" line was used to carry the flow to the top of the riser. In these conditions, the 2" line is totally isolated from the rest of the system and thus it should not be pressurized. At the top of the riser, just before the input of the top separator there is a bridge with a valve that connects the 4" and the 2" line. Fig. 60 shows a detail of this valve:

Fig. 60: Bridge valve between the 4" line and the 2" line

The fault introduced consisted basically in opening this bridge valve keeping both sides of the 2" line isolated from the rest of the installation. The objective of this fault is to simulate a wrong operation of the system that in a real process can cause degradation in the performance or in the output product quality or even safety issues. This fault should not affect the flow conditions in the 4" line or other parts of the rig, but it will pressurize the 2" line. In this particular case, an additional variable was measured and included in the analysis in both the training and the monitored data set. This variable is the pressure measured at the bottom of the riser in the 2" line by PT417. The addition of this variable caused a change in the estimated threshold for the T^2 and Q indicators, obtaining new values of 1755.16 and 3774 respectively. The valve was opened fast as the pressurization is expected to happen almost instantaneously. One data set was analysed, identified in chapter 3 as data set 6.1.

The results obtained in terms of fault detection and diagnosis are represented in Fig. 61. The fault detection was noted at sample 1724 for both statistical indicators, showing a fast reaction to the fault. Both indicators remained over the UCL until the end of the experiment as expected. In the instant of fault detection, the contribution plots of both indicators show clearly that the fault is related with pressure measured in the 2" line (PT417). Without this additional measurement included in the analysis it was impossible to detect the fault, and both indicators remained below the UCL for the whole experiment with the exception of one short false alarm in the Q indicator (see Fig. 62).
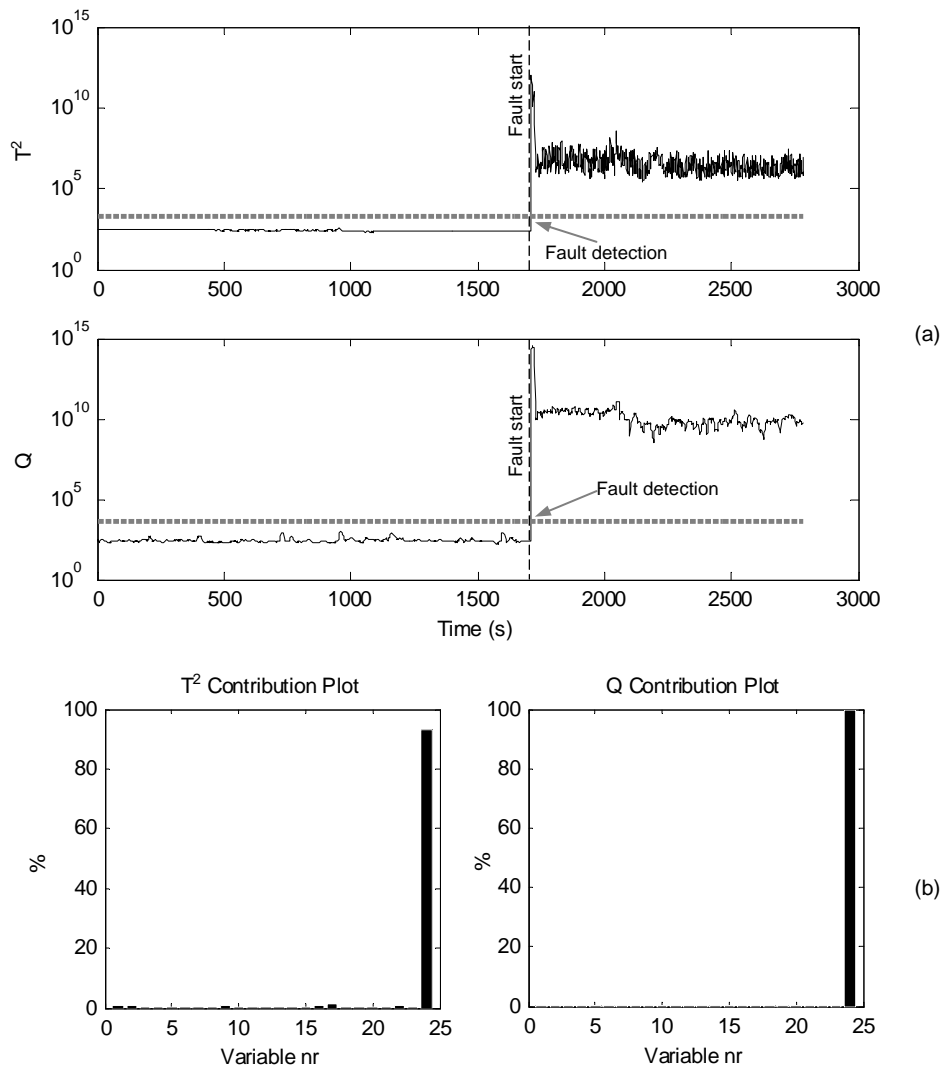
Fig. 61: Results from data set 6.1: T^2 and Q indicators (a) and contribution plots (b)
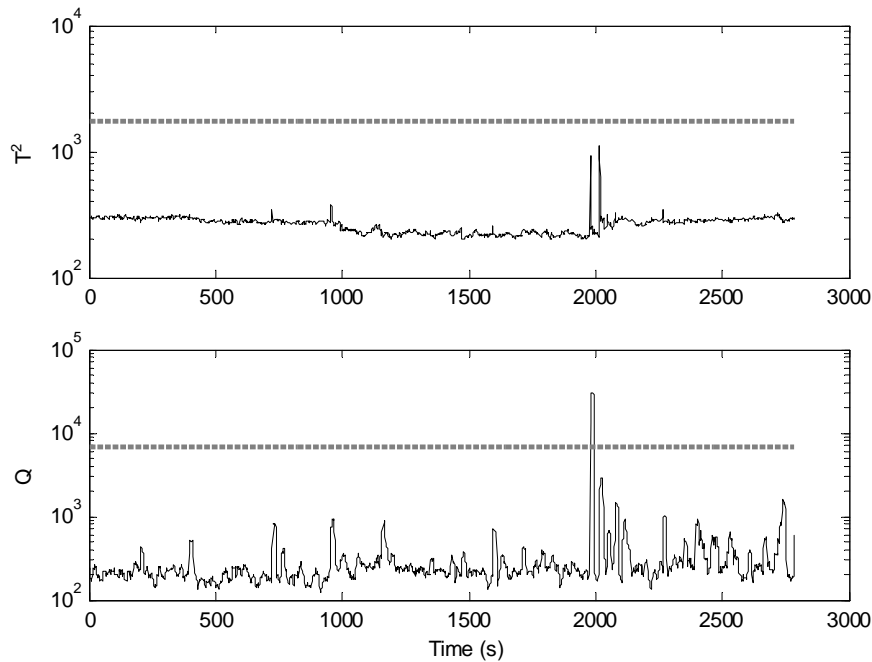
Fig. 62: T^2 **and** Q **indicators for Case 6 without PT417**

4.4 Results summary

This section summarizes the results obtained in terms of fault detection. The results shown previously obtained by the application of CVA with KDE are compared with the results obtained by the application of PCA, DPCA, PLS and DPLS with and without KDE for the calculation of the indicators' thresholds. The results obtained from the analysis of additional data sets described in chapter 3 but not included in 4.3 are also summarized here. For the PCA, DPCA, PLS and DPLS analysis the number of retained states was set to 5 in order to cover 85% of the covariance, and the number of lags considered in DPCA and DPLS was set to 5. For the PLS and DPLS analysis the air and water set points were selected to construct the predictor matrix, while the measurements taken from the system were used in the predicted matrix. For the calculation of the UCL a confidence bound of 99% was used in all cases.

Table 13 and Table 14 show the results obtained in terms of detection rate, which is calculated as the percentage of observations between the fault starting and the ending point whose indicator value is over the corresponding threshold. Table 12 shows the start and end points considered for each one of the data sets.

112

The false alarm rates presented in Table 15 and Table 16 are calculated as the percentage of observations outside the fault region for which the indicator exceeds the threshold. The fact that most faults were introduced gradually makes it more difficult to identify the start and end of the fault. For this performance analysis it was established that the starting point of each fault will be the instant where the system was first manipulated, independently of the degree of manipulation. In the particular case of data sets 5.1 and 5.2 where the faulty conditions were introduced and removed 2 and 3 times respectively in each experiment, the start and end of the faulty conditions is considered as the instant when the flow rate set point expected to produce slugging was introduced or changed.

The detection times presented in Table 17 and Table 18 are calculated as the difference between the fault detection time and the fault starting time. The event of fault detection was considered independently for the T^2 and Q indicators as the first point after which it was found a significant amount of consecutive observations above the threshold, trying to avoid short false alarms happening inside the fault zone.

Table 12: Fault starting and ending points

Data set	Start	End
1.1	1566	5181
1.2	657	3777
1.3	691	3691
2.1	2244	6616
2.2	476	2656
2.3	331	2467
3.1	1136	8352
3.2	333	5871
3.3	596	9566
3.4	452	8731
4.1	953	6294
4.2	851	3851
4.3	241	3241
5.1 (1)	686	1172
5.1 (2)	1772	2253
5.2 (1)	1633	2955
5.2 (2)	7031	7553
5.2 (3)	8057	10608
6.1	1723	2800
6.2	1037	4830

Table 13: T^2 detection rate (%)

Case	Data set	PCA	DPCA	PCA KDE	DPCA KDE	PLS	DPLS	PLS KDE	DPLS KDE	CVA	CVA KDE
	Threshold	15.09	15.09	21.37	21.88	15.09	15.09	81.14	1239.32	44.38	1753.98
1	1.1	22.65	23.10	14.72	14.69	36.49	-	27.11	20.97	-	60.66
	1.2	22.91	22.85	22.85	22.79	32.47	-	22.91	22.98	-	51.73
	1.3	19.63	19.60	19.56	19.56	31.2	-	19.7	19.77	-	39.73
2	2.1	-	-	-	-	1.87	59.42	-	-	-	21.87
	2.2	-	-	-	-	-	15.45	-	-	-	16.88
	2.3	-	-	-	-	-	-	-	-	-	29.21
3	3.1	98.37	98.44	97.69	97.81	99.36	-	96.71	96.72	-	98.69
	3.2	81.41	82.81	37.14	36.53	48.23	92.79	32.95	32.39	-	45.75
	3.3	97.22	97.73	95.57	95.37	99.36	-	96.21	96.62	-	98.81
	3.4	97.09	97.15	96.67	96.65	99.33	-	96.80	96.98	-	98.06
4	4.1	34.50	35.61	30.425	29.43	40.34	-	30.94	30.08	-	57.74
	4.2	9.5	9.87	6.4	4.70	15.07	-	5.97	3.44	90.53	21.50
	4.3	18.43	17.23	15.47	12.3	25.43	-	14.8	15.10	-	29.03
5	5.1	70.63	71.87	53.98	54.91	94.72	-	29.69	28.74	-	25.33
	5.2	83.34	84.19	67.51	66.49	98.68	-	57.73	52.77	100	43.17
6	6.1	99.72	99.53	99.72	99.44	67.13	-	-	-	-	99.91
	6.2	99.86	99.78	99.84	99.76	-	-	-	-	100	100

The results from Table 13 about the detection rate observed using the T^2 statistic show that PCA and DPCA have a similar rate of detection success. The detection rate observed by adding KDE for the calculation of the indicator threshold is also very similar. The performance of PLS was slightly better than PCA and DPCA, but in most cases it was impossible to determine the fault detection using DPLS due to the unrealistic value obtained from the threshold calculation. The DPLS results were massively improved adding KDE for the calculation of a more realistic threshold. Similarly, it was impossible in most cases to determine the success rate of CVA due to the low threshold value estimated using the Gaussian assumption, but the results were improved adding KDE, obtaining better detection rates than with any other method.

Table 14: Q detection rate (%)

Case	Data set	PCA	DPCA	PCA KDE	DPCA KDE	PLS	DPLS	PLS KDE	DPLS KDE	CVA	CVA KDE
Threshold		9.77	42.34	10.73	50.15	46.71	226.21	41.53	456.75	343.29	6940.73
1	1.1	52.72	60.17	50.32	49.68	30.23	50.34	34.80	29.90	81.24	61.41
	1.2	42.92	43.33	42.46	42.48	32.56	32.69	32.59	32.63	83.75	61.13
	1.3	38.23	42.97	33.53	36.37	19.83	87.83	20.26	29.9	64.73	39.73
2	2.1	62.96	69.47	58.30	60.43	-	6.84	1.21	3.79	46.75	-
	2.2	24.86	45.22	18.67	18.35	-	12.16	-	-	17.84	-
	2.3	32.54	15.87	30.43	29.35	-	12.68	-	-	34.59	-
3	3.1	99.72	100	99.84	100	98.67	99.42	98.75	98.65	99.91	97.87
	3.2	-	-	96.87	99.89	88.01	93.82	96.17	56.62	-	41.71
	3.3	99.65	99.70	99.63	99.61	99.25	99.88	99.26	99.38	99.60	96.55
	3.4	99.72	99.94	99.69	99.59	99.19	99.89	99.23	99.32	99.57	97.01
4	4.1	92.64	94.14	92.14	91.98	43.02	57.59	45.22	42.85	92.53	82.32
	4.2	31.63	37.20	28.47	30.77	15.30	61.53	16.17	15.73	90.97	47.63
	4.3	94.86	99.43	92.23	94.1	22.67	54.1	25.00	25.83	92.3	58.43
5	5.1	-	-	-	-	58.63	96.48	65.77	78.80	-	16.23
	5.2	89.41	93.28	87.34	89.57	71.51	97.64	74.13	85.89	100	67.49
6	6.1	99.81	99.81	99.81	99.81	99.81	99.72	99.81	99.72-	99.91	99.91
	6.2	99.94	100	100	100	99.92	99.89	100	99.89	100	100

The results obtained from the analysis of the detection rate using the Q statistic indicate a similar trend. DPCA shows a better performance than PCA and the addition of KDE reduces slightly the detection rate of both methods due to the higher threshold value obtained. DPLS performance is better than PLS, but again this performance is slightly reduced by the addition of KDE, although the real benefit of KDE will be seen in the false alarm rate. The performance of CVA is better than the performance observed for other techniques. Again the addition of KDE to CVA reduces slightly the detection rate due to the higher threshold estimated, but this small decrement will be rewarded by a large reduction of the false alarm rate.

Table 15: T^2 false alarm rate (%)

Case	Data set	PCA	DPCA	PCA KDE	DPCA KDE	PLS	DPLS	PLS KDE	DPLS KDE	CVA	CVA KDE
1	1.1	5.91	5.69	3.09	3.18	23.99	-	0.63	0.45	-	1.14
	1.2	13.51	13.95	10.84	11.14	18.48	-	5.79	1.78	-	2.89
	1.3	9.53	10.14	6.73	6.89	12.11	-	1.59	0.98	-	2.42
2	2.1	-	-	-	-	9.58	95.71	-	-	-	41.76
	2.2	-	-	-	-	-	18.90	-	-	-	1.67
	2.3	-	-	-	-	-	-	-	-	-	1.47
3	3.1	15.95	13.32	10.62	10.83	22.62	-	2.18	1.81	-	2.13
	3.2	51.77	51.77	38.96	37.60	10.08	56.57	4.08	2.60	-	3.54
	3.3	22.29	22.30	17.94	17.78	18.95	-	2.28	1.89	-	2.78
	3.4	20.80	21.24	13.49	13.72	26.11	-	0.22	0.22	-	0.22
4	4.1	3.43	3.64	1.49	1.49	22.33	-	1.50	1.50	-	2.78
	4.2	9.17	9.65	6.96	6.96	14.81	-	5.65	5.04	39.29	5.31
	4.3	29.35	29.65	21.48	21.33	32.22	-	9.23	4.39	-	13.31
5	5.1	70.08	69.82	65.31	65.05	50.25	-	23.95	23.32	-	24.46
	5.2	11.47	11.15	3.03	2.68	64.44	-	3.02	3.10	57.91	0.48
6	6.1	1.27	1.45	0.29	-	0.29	-	-	-	-	0
	6.2	0	0	0	0	-	-	-	-	37.84	0

In terms of false alarms, PCA and DPCA showed a similar performance, which really was improved by the introduction of KDE .This improved performance caused by the addition of KDE is even more evident in the case of PLS and DPLS, which show the best results together with CVA including also KDE for the estimation of the threshold. The results obtained from this analysis evidence the major improvements obtained when KDE is used to estimate the threshold of the indicators.

Table 16: Q false alarm rate (%)

Case	Data set	PCA	DPCA	PCA KDE	DPCA KDE	PLS	DPLS	PLS KDE	DPLS KDE	CVA	CVA KDE
1	1.1	16.98	37.56	8.51	11.20	3.051	34.17	4.00	2.73	37.84	1.34
	1.2	35.26	50.26	25.24	34.04	11.14	12.28	12.03	8.04	32.74	3.34
	1.3	10.22	10.74	8.25	9.47	6.89	66.81	7.00	6.91	20.06	3.10
2	2.1	53.58	53.46	53.53	53.47	-	22.80	0.91	1.57	60.49	-
	2.2	23.48	56.91	11.02	14.06	-	3.20	-	-	18.77	-
	2.3	11.59	15.87	7.47	5.83	-	8.89	-	-	30.66	-
3	3.1	32.97	57.79	20.43	33.17	11.41	54.17	13.39	10.00	33.83	1.54
	3.2	-	-	71.93	86.64	48.36	51.50	52.17	16.43	-	3.67
	3.3	12.31	15.73	9.98	9.92	16.94	35.59	19.40	15.92	25.25	5.51
	3.4	23.23	42.03	12.61	16.37	15.49	56.63	18.36	8.18	31.19	0.22
4	4.1	47.30	50.51	38.29	47.93	1.93	56.93	2.25	2.30	46.27	2.78
	4.2	25.91	49.62	18.12	20.88	7.03	8.62	7.17	6.77	41.35	6.75
	4.3	42.97	60.96	31.32	39.93	18.46	75.34	20.12	19.06	41.30	12.10
5	5.1	-	-	-	-	35.77	73.06	41.99	41.11	-	23.76
	5.2	22.53	26.77	19.90	61.64	7.86	32.34	9.02	21.73	40.97	5
6	6.1	5.97	8.87	3.19	4.87	0	0.81	0	0	8.067	0
	6.2	0.28	0	0.28	0	0	0.38	0	0	13.02	0

In the case of false alarms observed in the Q statistic, PCA showed a better performance than DPCA, but those differences are minimal after the addition of KDE, which caused a massive improvement in the false alarm rate. Similarly PLS presents lower rates off false alarms than DPLS and the latter one improved its performance remarkably with the addition of KDE. CVA shows a relatively high rate of false alarms, but this rate is reduced when adding KDE, obtaining the best results among all the algorithms studied.

Table 17: T^2 detection time (s)

Case	Data set	PCA	DPCA	PCA KDE	DPCA KDE	PLS	DPLS	PLS KDE	DPLS KDE	CVA	CVA KDE
1	1.1	3080	3080	3083	3084	2296	-	2698	3080	-	1426
	1.2	2405	2407	2407	2409	2107	-	2405	2407	-	1506
	1.3	2411	2412	2413	2415	2105	-	2409	2411	-	1808
2	2.1	-	-	-	-	4378	4023	-	-	-	3416
	2.2	-	-	-	-	-	1849	-	-	-	1812
	2.3	-	-	-	-	-	-	-	-	-	1512
3	3.1	207	112	374	211	52	-	399	235	-	94
	3.2	3266	3219	3745	3733	3240	1405	3967	3788	-	3086
	3.3	434	307	453	450	90	-	451	307	-	106
	3.4	299	243	303	305	72	-	303	250	-	160
4	4.1	585	586	585	588	512	-	584	585	-	548
	4.2	493	494	507	443	457	-	442	448	291	436
	4.3	400	393	416	404	375	-	400	400	-	416
5	5.1 (1)	-	-	-	-	-	-	65	64	-	69
	5.1 (2)	40	37	47	45	28	-	103	121	29	109
	5.2 (1)	27	28	42	40	32	-	50	47	2	83
	5.2 (2)	-	-	-	-	-	-	-	-	-	44
	5.2 (3)	16	14	19	18	17	-	36	86	83	114
6	6.1	3	1	3	2	356	-	-	-	-	1
	6.2	5	4	6	5	-	-	-	-	1	1

In terms of detection time using T^2, DPCA shows a faster response in some cases compared with PCA. The detection time is slightly increased for both algorithms after the addition of KDE due to the higher threshold estimated. Similarly, DPLS with KDE presents lower detection times than PLS with KDE in most cases, improving also the performance observed in DPCA, but the algorithm which showed the fastest reaction to the faults was CVA with KDE.

Table 18: Q detection time (s)

Case	Data set	PCA	DPCA	PCA KDE	DPCA KDE	PLS	DPLS	PLS KDE	DPLS KDE	CVA	CVA KDE
1	1.1	1829	1830	1830	1831	2584	1834	2583	2585	644	1419
	1.2	1803	1804	1804	1804	2104	2104	2103	2106	590	1213
	1.3	2103	1809	2103	2101	2405	1713	2404	2107	1195	1808
2	2.1	1852	1783	2214	1857	-	4073	4199	4249	3431	-
	2.2	1816	1803	1828	1817	-	1912	-	-	1791	-
	2.3	1536	1536	1542	1537	-	1918	-	-	1481	-
3	3.1	17	0	74	0	99	46	94	101	34	188
	3.2	-	-	2067	1173	1834	1405	1409	2637	-	3371
	3.3	34	34	50	35	93	31	92	65	22	309
	3.4	32	24	33	34	153	59	63	63	21	247
4	4.1	426	422	436	428	480	461	467	476	455	578
	4.2	376	348	410	369	455	436	440	442	292	430
	4.3	302	162	313	302	379	371	378	3737	293	371
5	5.1 S1	-	-	-	-	-	-	-	-	-	73
	5.1 S2	-	-	-	-	47	32	42	60	23	109
	5.2 (1)	36	31	21	34	39	29	38	34	1	87
	5.2 (2)	-	-	-	-	-	-	-	-	-	-
	5.2 (3)	85	13	85	16	29	23	29	27	36	46
6	6.1	2	1	2	1	2	1	2	1	1	1
	6.2	2	1	2	1	3	1	2	1	1	1

Using the Q indicator the fault reaction time is similar for PCA and DPCA, but DPCA produced slightly better results. The addition of KDE in these cases again increases the detection time. DPLS performs better than PLS but after the addition of KDE the performance of both algorithms is similar, and slower compared with the Gaussian assumption. CVA shows the fastest performance, which is slightly reduced by the addition of KDE.

4.5 Conclusion

CVA is a well-known technique for dynamic process monitoring which performance has already been tested by several researchers using computer simulated data. Nevertheless, examples of application using real data are anecdotic. In this work CVA was applied to sets of data obtained in a large

scale test rig where different process faults were seeded during changing operational conditions. The faults were successfully detected using the T^2 and Q metrics. In some cases the fault detection happened earlier for the Q statistic but in the other hand, Q produced a higher number of false alarms than T^2. This indicates that the Q statistic was slightly oversensitive for these particular conditions. This phenomenon has already been reported by other researchers [2] and it has been attributed to the fact that most faults tend to create new states rather than magnifying the states based on normal operation. Contribution plots were used to locate those measurements which are more severely affected by the fault in order to help in the process of fault diagnosis once a fault has been detected. The results obtained for the T^2 contribution plots were always related with the fault introduced and, combined with knowledge about the process, could have been used for fault identification. In the other hand the Q contribution plots were not always as clear as the T^2 plots. This was probably caused by the mentioned oversensitivity of the Q indicator; as a consequence any small variation in the residual space is represented in the contribution plots, sharing the contribution between several variables which complicates the fault detection. These results demonstrate that CVA can be effectively applied for the detection and diagnosis of faults in real complex systems working under variable operating conditions.

When comparing the performance of CVA including KDE with other methodologies it is clear that this method is superior to the rest in terms of detection rate, false alarm rate and detection time. For fault detection, CVA with KDE produced a higher detection rate in most cases and, when it was not the case, the higher rate obtained using other method had a much higher false alarm rate as a consequence (see for example data sets 3.2 and 4.2). It is noticeable that the use of KDE for the calculation of the thresholds improved the performance of all the algorithms in terms of false alarms rate. In some of the studied cases it was not possible to use the results obtained for fault detection if the thresholds were calculated using the Gaussian assumption because the indicator value was higher than the threshold during the whole experiment. This evidences that the Gaussian assumption in this case was unrealistic due to high

nonlinearities in the system, which made almost essential the use of KDE for threshold estimation.

5 ESTIMATION OF PROCESS PERFORMANCE DEGRADATION UNDER FAULTY CONDITIONS USING CANONICAL VARIATE ANALYSIS

Abstract

Condition monitoring of industrial processes can minimize maintenance and operating costs of industrial systems while increasing the process safety and enhancing the quality of the product. In order to achieve these goals it is necessary not only to detect and diagnose process faults but also to react to them by scheduling the maintenance and production according to the condition of the process.

In this investigation, process data was acquired from an experimental large-scale multiphase flow facility operated under changing operational conditions to test the capabilities of canonical variate analysis (CVA) to detect and diagnose process faults, as well as to estimate performance degradation and predict the behavior of the faulty system. This information could be used to improve the maintenance and production schedules attending to two criteria: firstly the impact of the fault in terms of operation safety, product quality, loading conditions, energy consumption, etc. compared with normal operation; and secondly the performance of the faulty system working under the operating conditions expected in the future.

The results suggest that CVA can be effectively used to estimate how faults affect the process performance in comparison to normal operation and to predict future process performance after the appearance of the fault by using data collected during the early stages of degradation.

5.1 Introduction

Modern industrial facilities are highly complex systems composed by interconnected devices of different nature (process, electrical and mechanical interfaces). Physics-based methods have been traditionally used for process monitoring [4]. These methods model the process behaviour using first-principle

equations and the fault detection and diagnosis is performed by observing differences between the model predictions and actual process measurements. These types of methodologies provide accurate results as long as the model is reliable. However, due to the complexity of modern facilities the characterization by using first-principle models is difficult and sometimes impossible [4]. Due to these difficulties and the high degree of instrumentation and automation in modern industrial processes data-based methods have gained popularity. Some of the main challenges associated with the data-driven methods are high-dimensional data, non-Gaussian distributions, non-linear relationships and dynamically varying operational conditions [3; 4; 17; 22].

As it was mentioned in previous chapters, data driven methods are widely used for fault detection and diagnosis applications in real industrial systems. Multivariate algorithms such as the Principal Component Analysis (PCA) or the Partial Least Squares (PLS) can account for the correlation between the different variables measured in the process, which is an advantage over the traditional univariate methods [2]. The use of lagged variables to take into account time correlation to extend PCA to dynamic system monitoring (DPCA) was proposed by Ku et al. [15]. Similarly, a dynamic version of PLS was proposed by Komulainen [16] . Despite of their success, DPCA and DPLS are known to be not as efficient as other state-space based methodologies such as Canonical Variate Analysis (CVA) when applied to systems working under variable operating conditions, principally due to the representation of the system dynamics [17], [18; 19].

CVA aims to maximize the correlation between two sets of variables [2]. This method has been applied in the past for fault detection and diagnosis using computer simulated data [2; 17-22] or data acquired in small test rigs [51] or particular parts of a system [52]. In chapter 4 the capabilities of CVA to detect and diagnose process faults were tested and compared with other monitoring techniques using experimental data. This investigation proved the superior performance of CVA for the detection and diagnosis of faults in a real complex system working under varying operational conditions. In addition, CVA can be

used to identify the coefficients in the linear state-space equations that define the dynamic behaviour of a system [21]. This method has been successfully applied for system identification using computer simulated data (including the Tennessee Eastman case study) and a pilot-scale distillation column. [21; 85; 86].

The objective of this investigation is to show how the system identification capabilities of CVA can be used to estimate the performance degradation in a real system affected by faults, and predict the future performance of the faulty system working under varying operational conditions. Once a fault is detected by the algorithm, if it is not critical and allows a safe process operation, the operators can use this information to improve the maintenance and production schedules based on the evaluation of the following two aspects: Firstly the impact of the fault in terms of operation safety, product quality, loading conditions, energy consumption, etc. compared with normal operation. Secondly the future performance of the faulty system for the different operating conditions expected.

The process data used in this investigation was acquired from the multiphase flow facility described in 3.2, where different process faults were deliberately introduced. The changing operational conditions and the non-linear nature of the multiphase flow process, together with the size and complexity of the test rig make this rig an ideal candidate for the case studied. CVA was applied in first place for fault detection and diagnosis and secondly to estimate the process performance under normal and faulty conditions using process data. The results suggest that CVA can be effectively used to estimate how the fault affects the process performance in comparison to normal operation and to predict future process performance after the appearance of the fault by using data collected during the early stages of degradation.

The rest of the chapter is structured as follows: 5.2.1 describes the CVA methodology for system identification, and a brief description of the test rig configuration for this study is presented in 5.2.2. The different data sets acquired are explained in detail in 5.3. The results obtained from the analysis

are presented in section 5.4, including algorithm training in 5.4.1, fault detection and diagnosis in 5.4.2, performance degradation in 5.4.3 and prediction of faulty system behaviour in 5.4.4. Finally the work is concluded in 5.5.

5.2 Methodology

5.2.1 CVA for system identification

The application procedure of CVA for fault detection and diagnosis was described in detail in 4.2.1. In addition to the detection and diagnosis of process faults, CVA can be used to identify the coefficients in the linear state space equations that define the dynamic behaviour of the system. This method has been used in the past for system identification using simulated data or data acquired from pilot-scale rigs [21; 85; 86]. In particular Juricek et al. [86] presented a successful application of CVA for identification of the Tennessee Eastman challenge, which simulates a nonlinear, open-loop unstable and large-dimensional process which contains a mixture of fast and slow dynamics. This method is more computationally efficient than PLS as it is based on generalized singular value decomposition instead of iterative estimation of residuals from the previous step. The procedure of system identification using CVA requires a sequence of steps [21]:

1) Selection of the input (manipulated) and output (measured) variables for the model
2) Selection of the input excitation sequence
3) Collection of data using the selected input
4) Selection of the model structure (number of states and number of past and future lags considered)
5) Estimation of the parameters in the state-space equations
6) Model validation

Given a set of inputs u and outputs y, the model that represents the linear state-space is given by:

126

$$x_{t+1} = \Phi x_t + G u_t + w_t \qquad \textbf{(5-1)}$$

$$y_t = H x_t + A u_t + B w_t + v_t \qquad \textbf{(5-2)}$$

where x_t is a r-order state vector, w_t and v_t are white noise and Φ, G, H, A and B represent the state-space matrices. If the order of the model r is selected to be equal or greater than the actual order of the system, the state vectors x_t can be replaced by the state estimates z_t [19; 87]:

$$z_t = J \cdot x_t \qquad \textbf{(5-3)}$$

 Assuming y_t and u_t are known and z_t can be obtained from CVA analysis, the only unknowns of the system are the matrices Φ, G, H, A and B. Larimore [87] proposes the next method for the calculation of those matrices using multivariate regression:

$$\begin{bmatrix} \Phi & G \\ H & A \end{bmatrix} = \sum \left[\begin{pmatrix} z_{t+1} \\ y_t \end{pmatrix}, \begin{pmatrix} z_t \\ u_t \end{pmatrix} \right] \cdot \sum {}^{-1} \left[\begin{pmatrix} z_t \\ u_t \end{pmatrix}, \begin{pmatrix} z_t \\ u_t \end{pmatrix} \right] \qquad \textbf{(5-4)}$$

where $\sum(a,b)$ denotes the sample covariance matrix for variables a and b.

This procedure was used in this investigation, in the first instance, to capture the behaviour of a system working under normal operating conditions. After a fault was introduced in the system, the performance degradation was estimated as the difference between the actual measurements and the estimations produced by the model assuming normal operation. Secondly, data was acquired from the system working under faulty conditions during the early stages of degradation to be able to predict the performance of the system for the different operating conditions forecasted in the future. The objective of this analysis is to observe how the fault affects the process when working under different loading conditions assuming that the severity of the faults allows the system to continue operating in suboptimal circumstances.

The accuracy of the model was assessed by looking at the average normalized error e for each one of the measured variables y_j:

$$e_j = \frac{1}{N} \sum_{t=1}^{T} \frac{y_{t,j} - \hat{y}_{t,j}}{y_{t,j}} \qquad \textbf{(5-5)}$$

where $\hat{y}_{t,j}$ represents the estimated value of the measured variable y_j at time t, and N denotes the total number of observations measured during a period of time T.

The total averaged model error E is computed as:

$$E = \frac{1}{m} \sum_{j=1}^{m} e_j \qquad \textbf{(5-6)}$$

5.2.2 Experimental set up

The data sets used for this investigation were acquired from the 3-phase flow facility described in 3.2.1. In this study all the data was captured at a sampling rate of 1 Hz. The variables used include 17 different process variables (see Table 19) and two process inputs (air and water flow rate set point, see Table 20). Only air and water were used in all the experiments and the three phase separator was always pressurized to 1.0 barg in normal operation. The variety in the nature of the selected variables (including pressure, flow rate, level, density, temperature, valve position and current measurements) and the different dynamics of these variables make the problem more challenging from the identification point of view. The changing operational conditions and the non-linear nature of the multiphase flow process, together with the size and complexity of the test rig make this rig an ideal candidate for the case studied.

Table 19: List of process variables used in this study

Variable nr	Location	Measured Magnitude	Unit
y_1	PT312	Air delivery pressure	barg
y_2	PT401	Pressure in the bottom of the riser	barg
y_3	PT408	Pressure in top of the riser	barg
y_4	PT403	Pressure in top separator	barg
y_5	PT501	Pressure in 3 phase separator	barg
y_6	PT408	Diff. pressure (PT401-PT408)	barg
y_7	PT403	Diff. pressure (PT408-PT403)	mbarg
y_8	FT305	Flow rate input air	Sm^3/h
y_9	FT104	Flow rate input water	kg/s
y_{10}	LI405	Level top separator	m
y_{11}	FT407	Density top riser	kg/m^3
y_{12}	FT104	Temperature water input	°C
y_{13}	LI504	Level gas-liquid 3 phase	%
y_{14}	VC501	Position of valve VC501	%
y_{15}	VC302	Position of valve VC302	%
y_{16}	VC101	Position of valve VC101	%
y_{17}	PO1	Water pump current	A

Table 20: List of process inputs used in this study

Input nr	Magnitude	Unit
u_1	Air flow rate	Sm^3/h
u_2	Water flow	kg/s

5.3 Cases studied

5.3.1 Normal operation

As was mentioned in section 4.2.1, it is necessary to obtain data from the system working under normal operating conditions to develop a CVA model together with transformation matrices and the thresholds of the health indicators. The procedure to apply the CVA used in this investigation the same that was used in chapter 4, where CVA was used to detect and diagnose process faults in the same test rig used here. Three new training data sets (T1, T2 and T3) were acquired from the system. In these sets the set points of air and water flow rates were deliberately varied to obtain data from the process working under varying operating conditions. The duration of these sets was 11881s, 3151s and 4550s respectively. The set T1 was acquired to train the CVA model while T2 and T3 were used for validation purposes. In order to

cover a wide spectrum of operating conditions 20 different combinations of air and water flow rates (see Table 21) were tested in T1. The flow conditions were changed, though not identically in each of the data sets, in order to obtain a good variety of large, small, long and short process changes happening in different directions (increment or decrement). The objective of this variety in the operational conditions is to ensure that the dynamics of the system are captured in all circumstances. Some of the air and water flow rate set points chosen for T2 are also shown in Table 21 as they were the same selected for T1. Nevertheless, the set points selected in T3 were different from the set points shown in Table 21 with the objective of ensuring an effective detection, diagnosis and identification under different operating points. Fig. 63 represents the measured air and water flow rate for each data set:

Table 21: Typical set point values for air and water flow rates during T1and T2

Air flow rate (m^3/h)	75	100	125	150	
Water flow rate (kg/s)	0.5	1	2	3.5	6
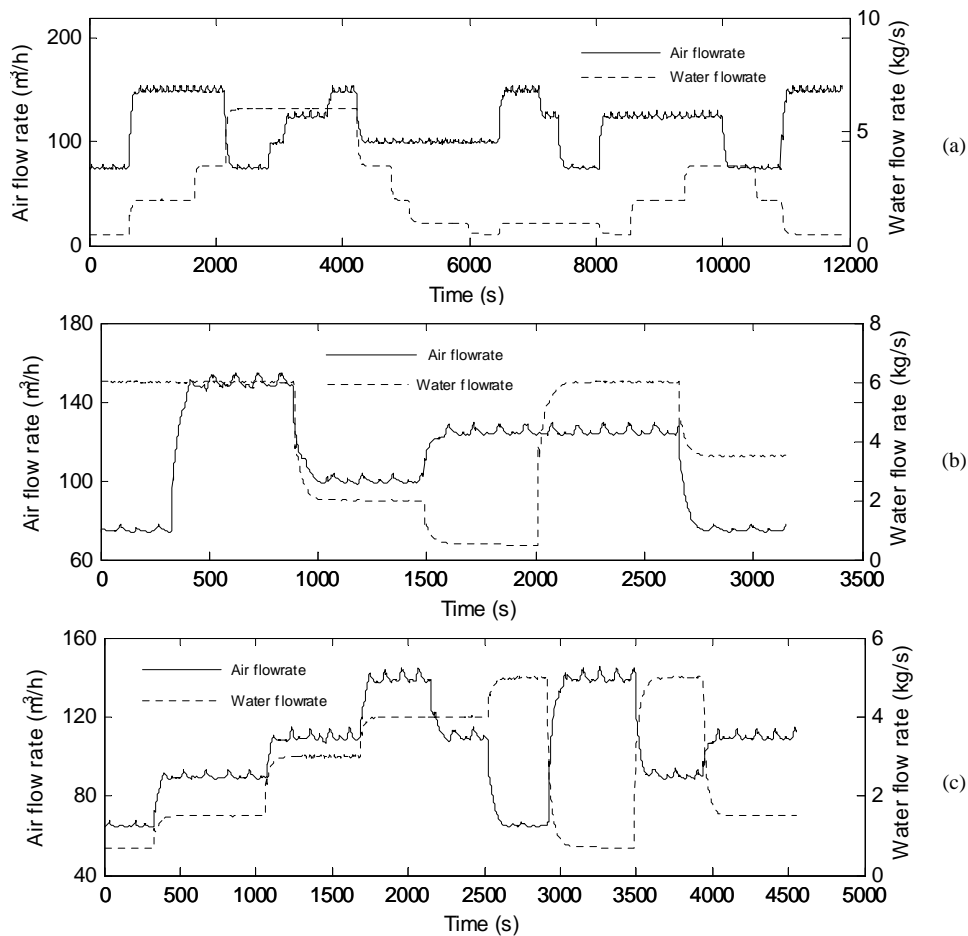
Fig. 63: Operational conditions for training data sets T1 (a), T2 (b) and T3 (c)

5.3.2 Case 1: Sensor communication error/ Stuck valve

The first fault introduced tried to simulate a sensor communication error in the pressure transducer PT501 (which measures the pressure inside the 3-phase separator) or a stuck valve in VC501 (which regulates the pressure inside the 3 phase separator). In order to simulate that fault, the control loop that links PT501 and VC51 was broken by changing the valve operation to manual mode. This test was carried out two times, the first one using the same flow rates set points represented in Table 21 (Case 1.1) and the second using different flow rates set points (Case 1.2). Fig. 64 and Fig. 65 show the operating conditions during the tests (a), the position of valve VC501 (b) and the evolution of PT501 during the tests (c) for Case 1.1 and Case 1.2 respectively.

131
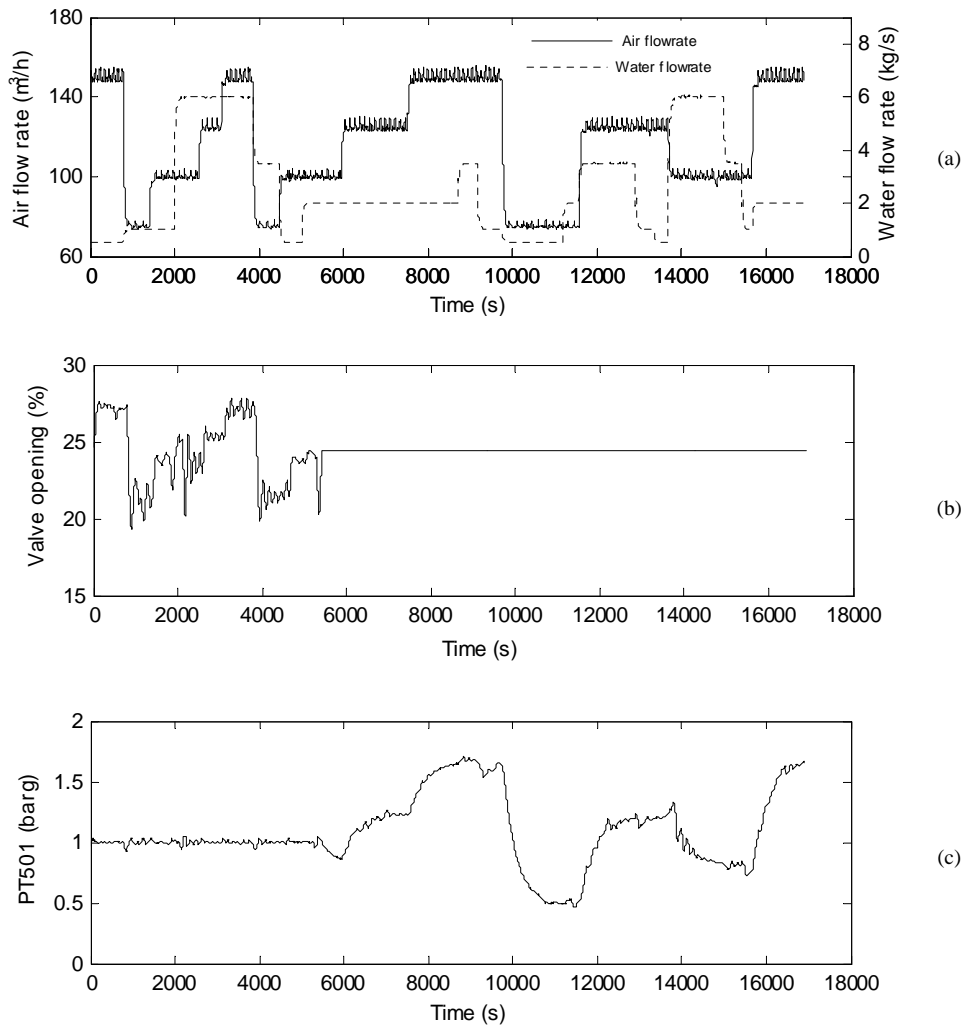
Fig. 64: Operational conditions (a), valve position VC501 (b) and fault evolution observed in PT501 (c) for Case 1.1
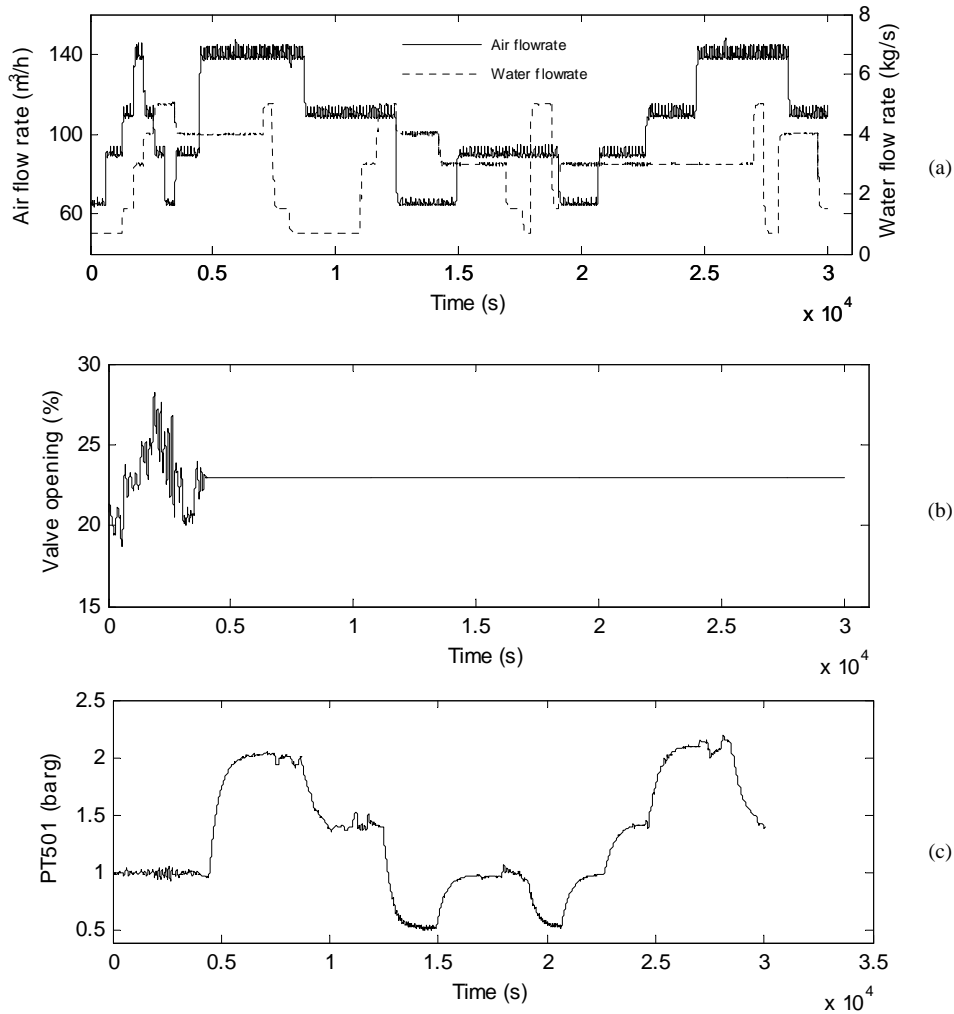
Fig. 65: Operational conditions (a), valve position VC501 (b) and fault evolution observed in PT501 (b) for Case 1.2

5.3.3 Case 2: Top separator input blockage

The second fault introduced simulates a partial pipe blockage in the inlet of the two phase separator placed on the top of the tower. The fault was simulated by closing VC404 (which is normally 100% open) to 20%. Fig. 66 shows the operating conditions during the test (a), the position of valve VC404 (b) and the evolution of the pressure measured in the top of the raiser (PT408) during the test (c).

133
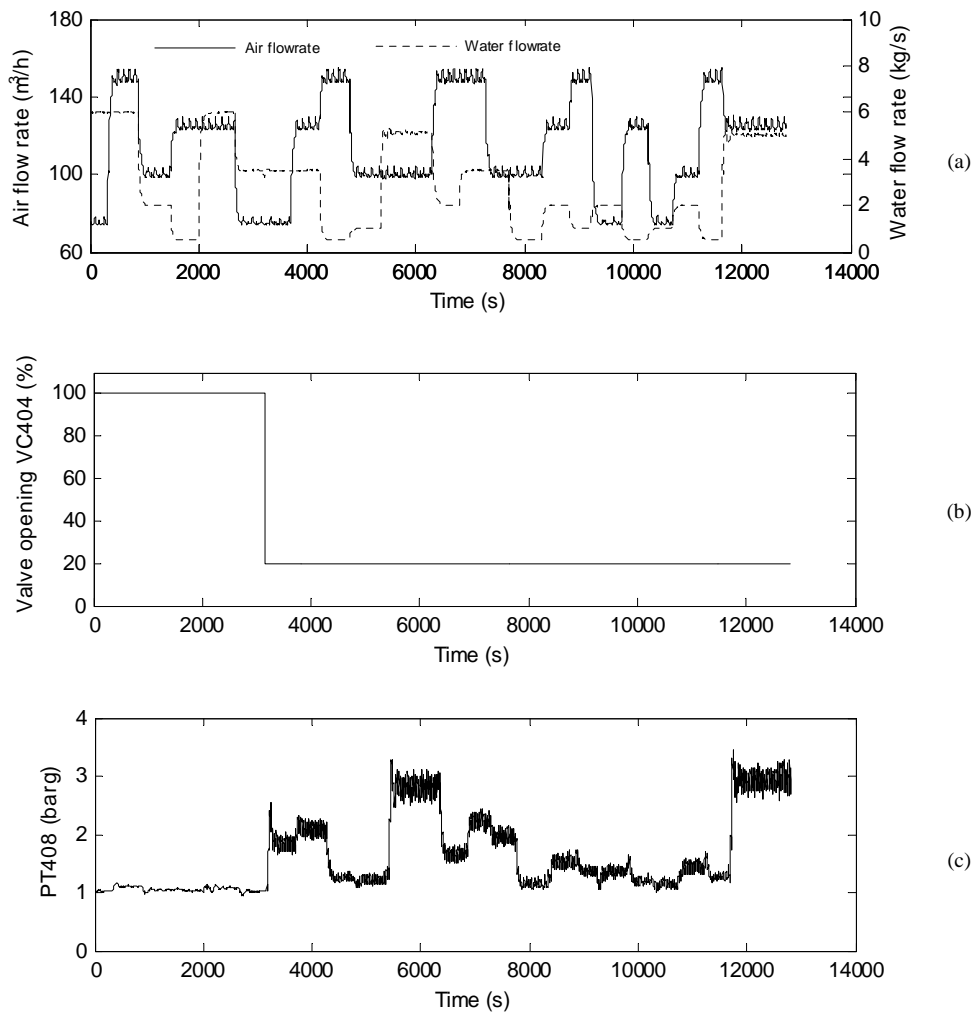
Fig. 66: Operational conditions (a), valve position VC404 (b) and fault evolution observed in PT408 (b) for Case 2

5.3.4 Case 3: Top separator air outlet blockage

The third fault introduced simulates a partial pipe blockage in the air outlet line of the two phase separator. The fault was simulated by closing VC401 (which is normally 100% open) to 45%. This test was carried out two times, the first one using the same flow rates set points represented in Table 21 (Case 3.1) and the second using different flow rates set points (Case 3.2). Fig. 67 and Fig. 68 show the operating conditions during the tests (a), the position of valve VC401 (b) and the evolution of PT403 during the tests (c) for Case 3.1 and Case 3.2 respectively.

(a)

(b)

(c)
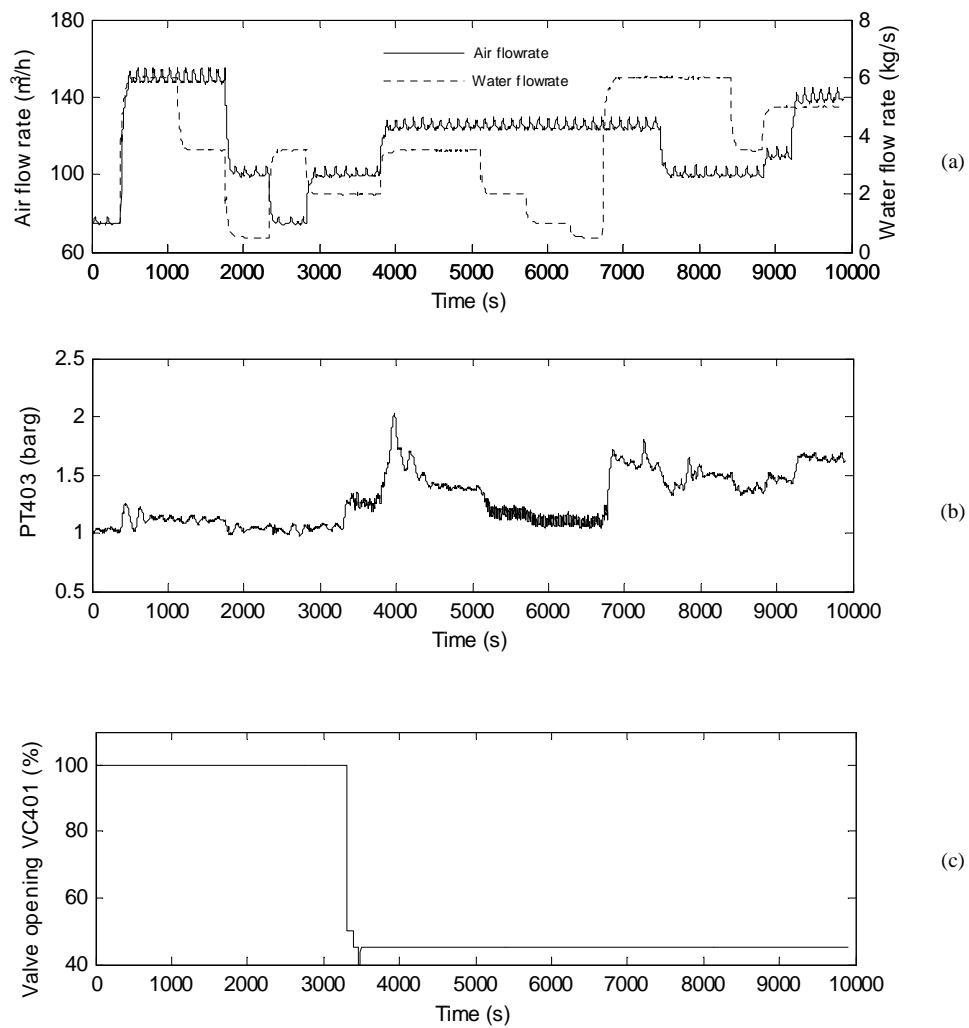
Fig. 67: Operational conditions (a), fault evolution observed in PT403 (b) and valve position VC401 (c) for Case 3.1

135
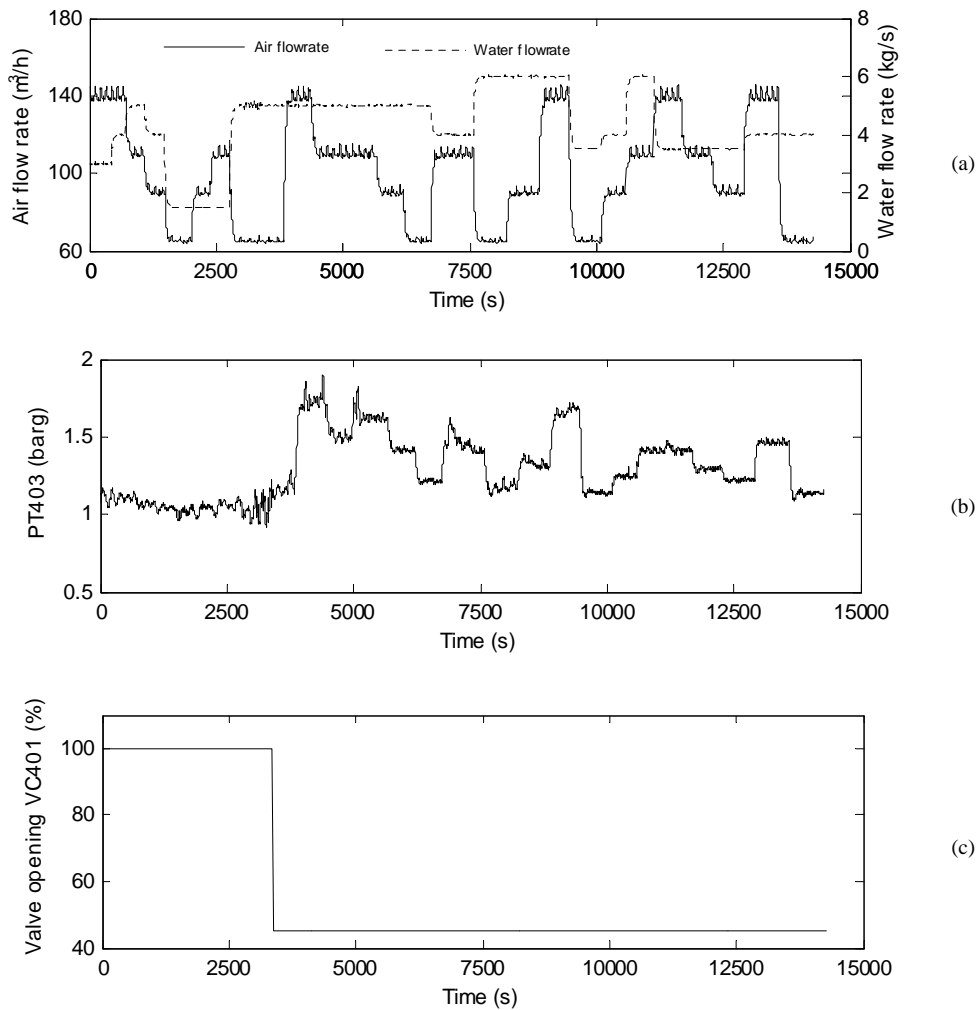
Fig. 68: Operational conditions (a), fault evolution observed in PT403 (b) and valve position VC401 (c) for Case 3.2

5.3.5 Case 4: Flow derivation through the 2" line

The last fault introduced simulates wrong system operation where part of the flow is derived through an alternative pipeline. The fault was simulated by opening the manual valves that allow the flow through the 2" line to the two phase separator in parallel with the flow in the 4" line, and some time later VC401 was fully closed to force the flow through the 2" line. Fig. 69 shows the operating conditions during the test (a), the position of the 2" line valve and VC404 (b) and the evolution of the pressure measured in the top of the raiser (PT408) during the test (c).

136
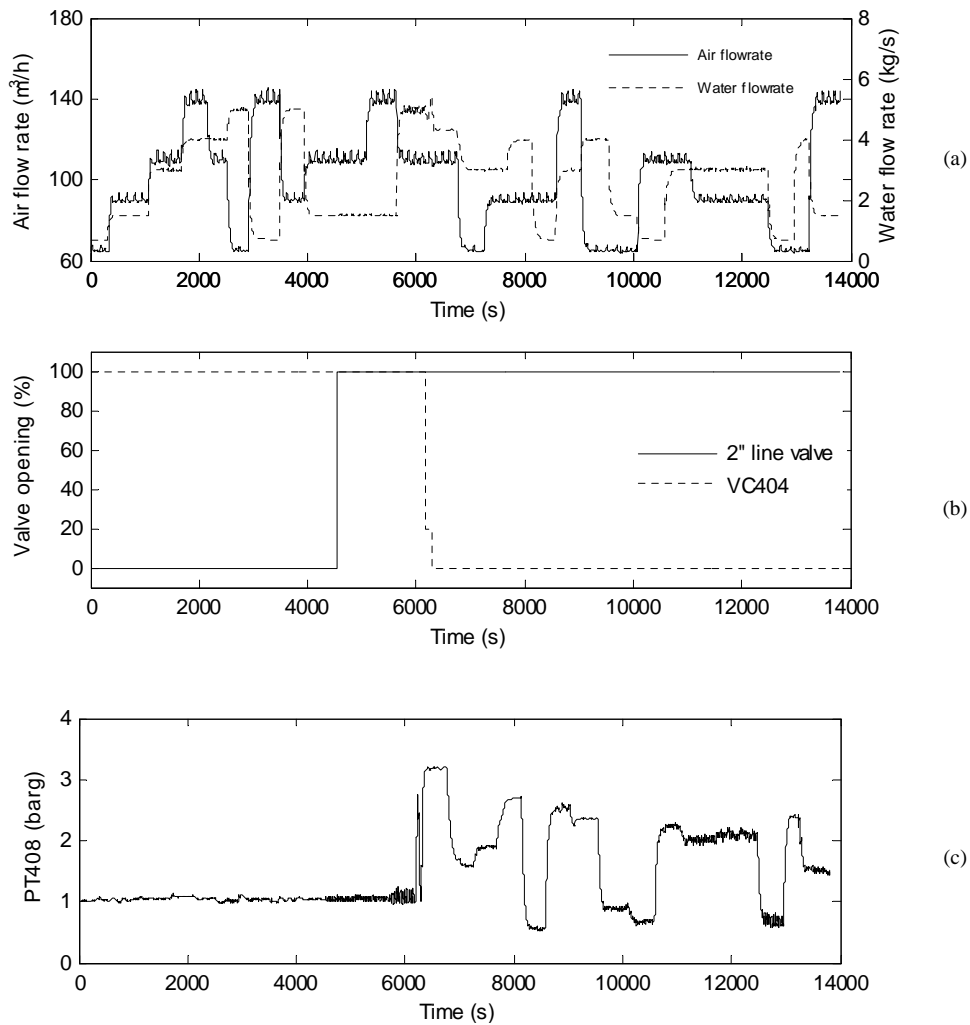
Fig. 69: Operational conditions (a), 2" line & VC404 valve position (b) and fault evolution observed in PT408 (b) for Case 4

5.4 Results

5.4.1 Algorithm training and selection of tuning parameters

As stated in section 5.3.1, data acquired from the system working under normal operating conditions is needed to obtain the transformation matrices and UCL for the thresholds. In order to build the past and future vectors in (4-1) it is necessary is to select the number of past and future lags considered (p and f) and the number of estates retained (r). The optimal number of past and future lags considered in the analysis was calculated computing the autocorrelation function of the summed squares of all measurements [17]. This function provides a measurement of the cross correlation between a signal and a

137

delayed version of itself at different lags. It is used to determine how long the correlation of the signal with past lags is significant, and thus only the number of lags which are deemed relevant are selected. Fig. 70 shows an example of autocorrelation function for the training data set T1 against a confidence bound of ±5%. According to this result the parameters p and f were set to 15 for this study.
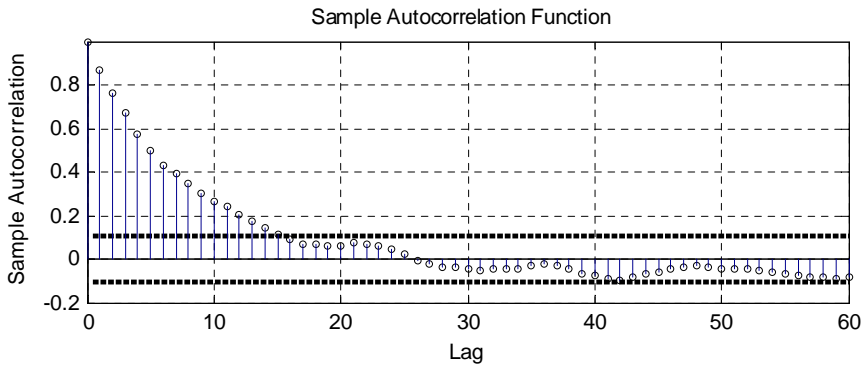
Fig. 70: Sample autocorrelation function for T1

From the different methodologies suggested for the calculation of the optimal number of states retained r, the most popular are those based on considering the dominant singular values in the matrix D [84] and methodologies based on the Akaike Information Criterion (AIC) [2].

Fig. 71 shows the normalized singular values obtained from D in (4-11) for data set T1. As it happened in the analysis in section 4.3.1, in this particular case there is not an evident number of dominant singular values as their values decrease slowly. Setting the number of retained states based on the assumption of the dominant singular values will derive in an unrealistic model [17]. The number of states retained is not especially relevant for this study due to the fault detection criterion used, where both statistical indicators (T^2 and Q) are used at the same time for fault detection. Using this assumption those system variations not captured in the retained space will be captured by the residual space and vice versa.
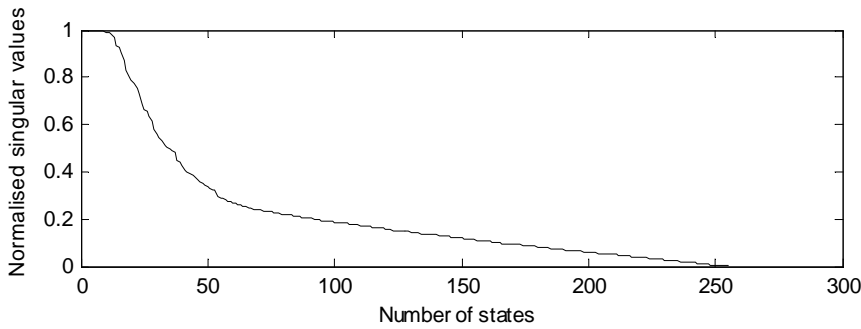
Fig. 71: Normalized singular values for T1

In order to check the capacity of T1 data set to represent the system dynamics accurately producing a low number of false alarms, this data set was used to train the CVA model while T2 and T3 were used in the monitoring phase. The objective of this analysis is to select the optimal number of dimensions retained r that minimizes the total number of false alarms found in T2 and T3. A confidence bound of 99% was considered for the calculation of the UCL. After the analysis testing different values for r for each combination of data sets finally r=25 was adopted in order to minimize the false alarm rate in normal conditions (see Fig. 72).
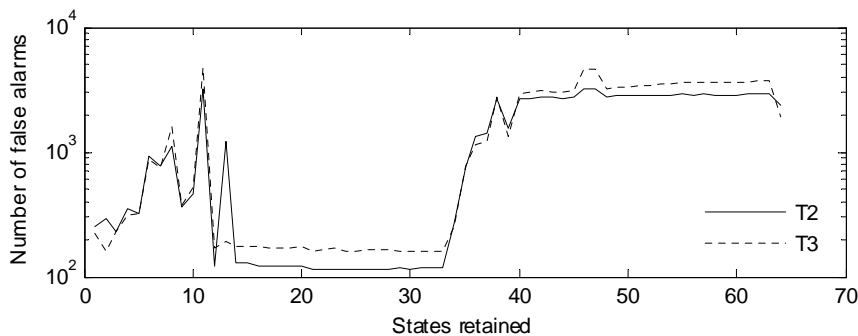
Fig. 72: Total number of false alarms for different number of states retained r

Similarly to the results obtained in 4.3.1, for low values of r the number of false alarms was high because the retained space is not able to represent accurately the states of the system and consequently the number of the T^2 threshold violations increases. On the other hand if the state order selected is too high it results in the model underfitting the data [2], increasing again the false alarm rate. This behaviour was observed in the analysis (Fig. 72) and the total number of false alarms found followed a similar pattern for T2 and T3, although this

number was typically higher for T3 probably because it is longer than T2 and the operating conditions selected were different from those chosen in T1. The T^2 and Q plots obtained from the analysis of the data sets T2 and T3 are represented in Fig. 73 and Fig. 74 respectively including the UCL calculated using KDE for a 99% confidence bound plotted as a grey dashed line.
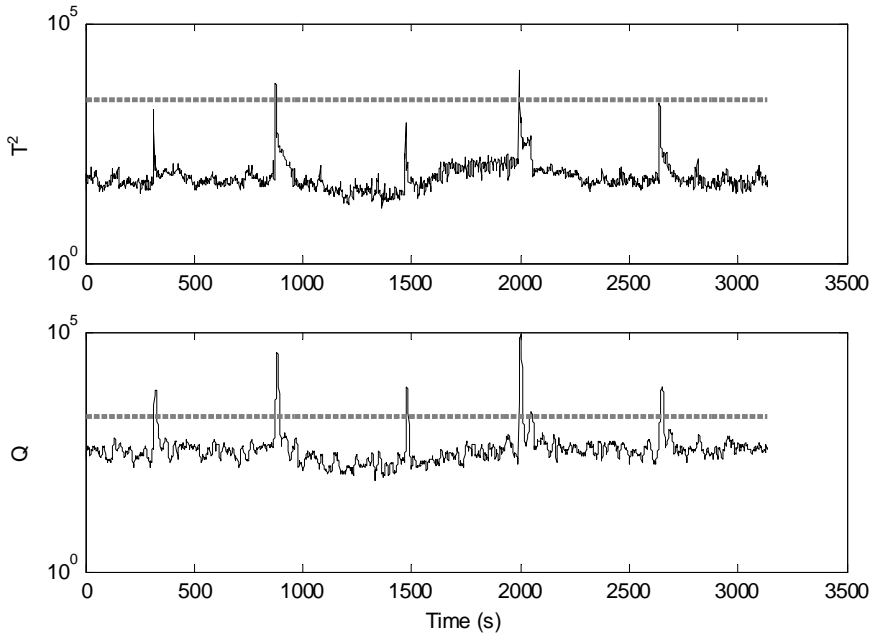
Fig. 73: T^2 and Q statistics for data set T2 using T1 for training (r=25)
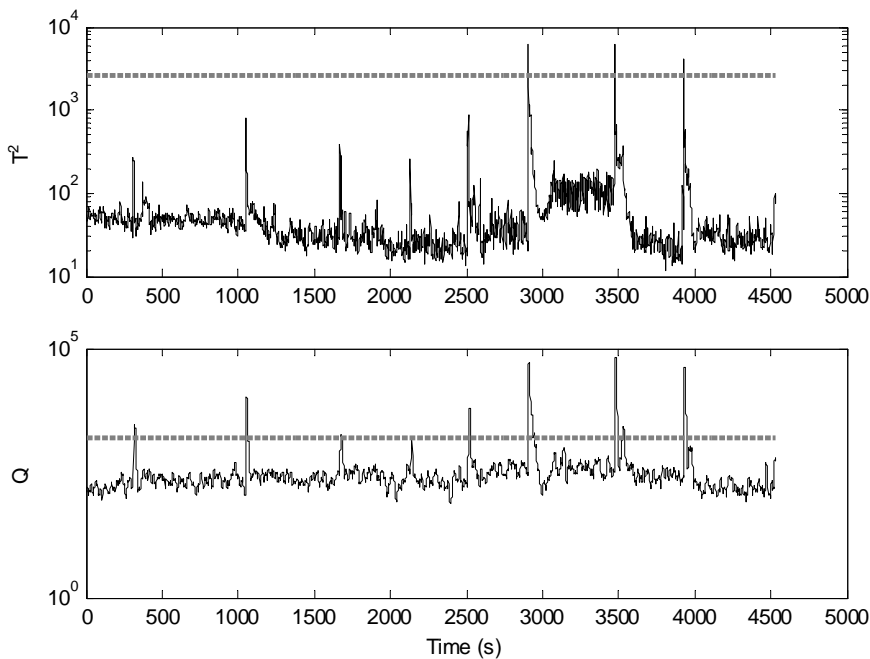
Fig. 74: T^2 and Q statistics for data set T3 using T1 for training (r=25)

140

5.4.2 Fault detection and diagnosis

To avoid the problem of short false alarms before the introduction of the fault, the event of fault detection was considered when the value of at least 60 consecutive samples of an indicator was over the UCL. For Case 4 two events of fault introduction are considered, the first when the 2" line valve was opened and the second one when VC404 was closed. Fig. 75 shows the results obtained in terms of fault detection for the six data sets analysed in their corresponding T^2 and Q plots, which are summarized in Table 22. The UCL is represented as a horizontal grey dashed line.
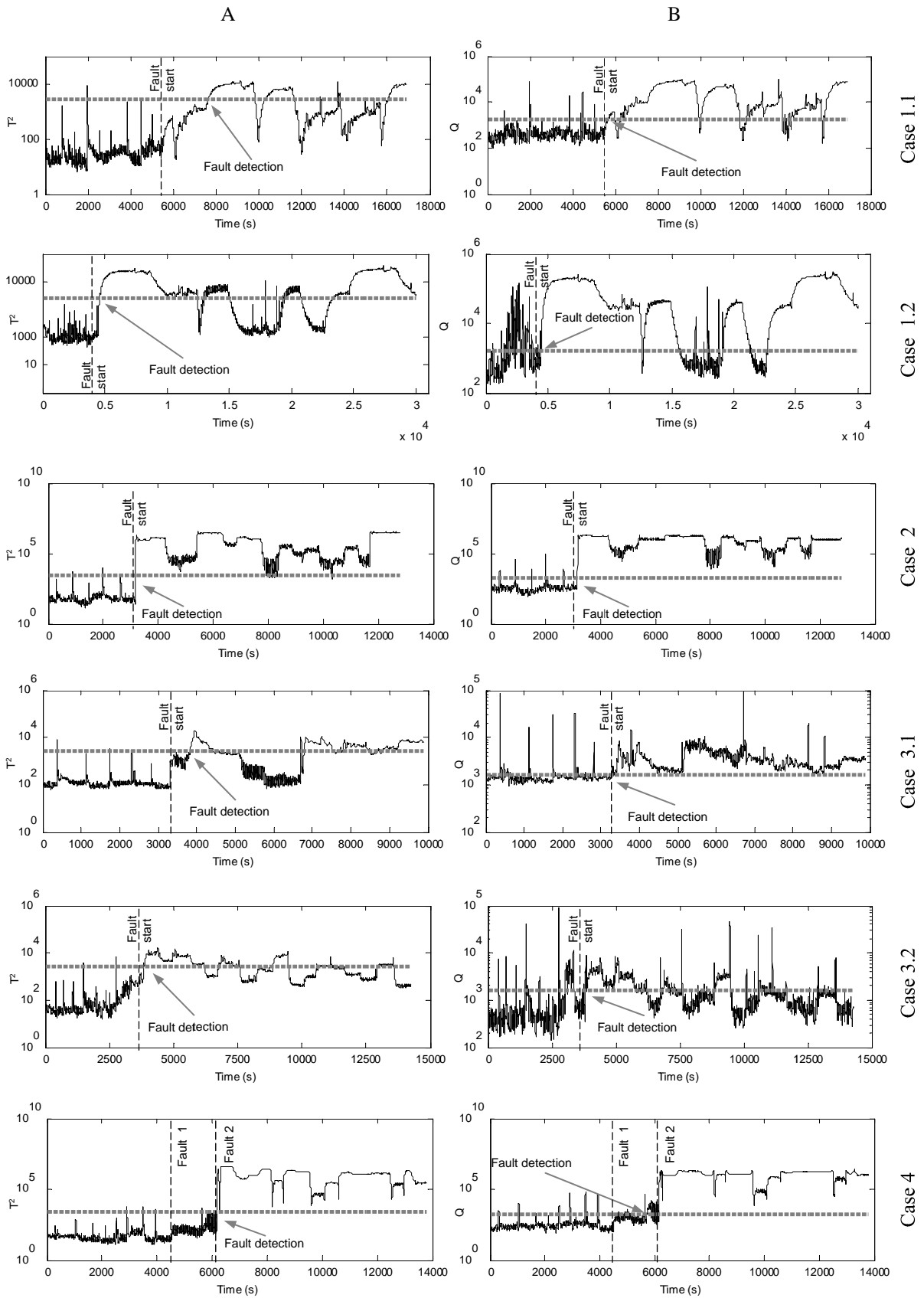
Fig. 75: T^2 (A) and Q (B) plots for the 6 data sets analysed

Table 22: *Fault detection summary*

Data set	Fault start (s)	T^2			Q		
		Detection time (s)	False alarms	False Negatives	Detection time (s)	False alarms	False Negatives
1.1	5456	2177	0.16%	60.8%	232	3.02%	10.8%
1.2	3963	655	0.02%	31.31%	515	36.23%	20.15%
2	3161	11	0.35%	0.62%	9	3.32%	0.09%
3.1	3301	540	0.45%	49.66%	91	9.39%	0.71%
3.2	3362	510	0.26%	54.34%	459	13.83	63.81%
4(1)	4550	-	-	-	1192	3.20%	11.85%
4 (2)	6179	8	0.34%	0.13%	-	-	-

The results obtained show that all the faults introduced were detected by the T^2 and Q statistics, although the changing operational conditions created fluctuations in the shape of both indicators. These fluctuations were expected as different flow rates will produce different pressures in the presence of a blockage, a reduction in the pipe diameter or lack of control action in a valve that regulates pressure in a tank. Consequently the different flow rates used in each case generated different levels of fault severity that the indicators are able to capture and represent.

All the faults were detected earlier by the Q statistic, which also produced a lower rate of false negatives in all cases except 3.2. In the other hand the T^2 statistic produced a very low rate of false alarms, showing a more reliable performance with no presence of faults. In the case of fault 4, only the Q statistic was able to detect the fault when the 2" line was opened, and it was necessary to close VC404 and derive the flow completely through the 2" line to make the fault visible for the T^2 statistic. It is important to notice that for cases 1.1 and 3.1 the results obtained were similar to the results obtained for cases 1.2 and 3.2 respectively, which were obtained under operating conditions that were not tested during the training phase. This demonstrates the capability of CVA to capture the system dynamics and detect abnormalities under different operating conditions.

Once the faults were detected, contribution plots were used at the instant of fault detection to determine the origin of the fault. The charts in Fig. 76 represent the contribution of each one of the measured variables to the final indicator value in each tested case.

In Case1.1 the most significant variable contributing to the T^2 statistic was the pressure measured in the 3-phase separator PT501, caused by the lack of control action. Other pressure measurements such as PT312, PT401, PT408 and PT403 were also affected; showing how the effects of a fault located at one point of the system affects the rest of the process. The Q contribution plot points directly to valve VC501 as the origin of the fault. Similar results were obtained in Case 1.2 where the same fault was introduced under different operating conditions.

In Case 2 the plots of both indicators show a high contribution of variable 7, which represents the differential pressure over VC404 obtained as the difference between the top riser pressure (PT408) and the top separator pressure (PT403). This result is very precise locating the origin of the fault. In addition, PT312, PT401 and PT408 show a significant contribution to the T^2 statistic, caused by the pressure increment in the pipeline before the blockage.

The contribution plots for Case 3.1 and Case 3.2 are again very similar. The most significant contribution for the T^2 indicator comes from the pressure measured inside the 2-phase separator, although the contribution of PT312, PT401 and PT408 are also significant due to the expansion of the fault in the system. In the case of Q indicator, most of the contribution comes also from PT403, although the level measured inside the 2-phase separator is also significant, pointing at a conflict in that region of the process.

The last case studied shows an important contribution of the top riser pressure (PT408) to the T^2 statistic, probably caused by flow being forced through the 2" line. The density measured in the top of the riser is the most significant variable for the Q indicator, reinforcing the results provided by the T^2 contribution plot.
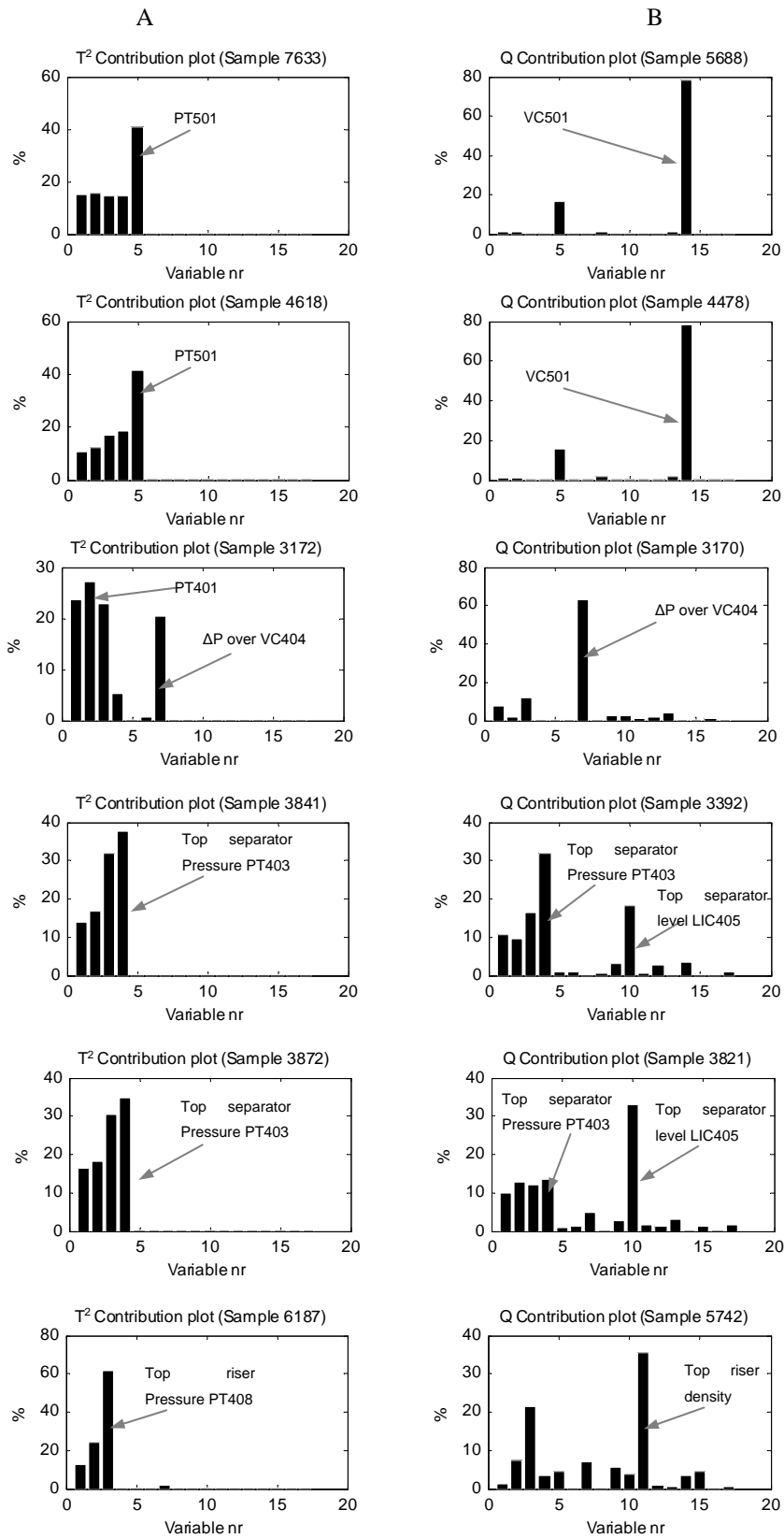
Fig. 76: T^2 **(A) and** Q **(B) contribution plots for the 6 data sets analysed**

5.4.3 Performance degradation

The data acquired under normal operating conditions in data set T1 was used to build a model of the system using the methodology presented in 5.2.1. This model was used to estimate the system response assuming normal operation for the four faulty cases studied, using the same input sequence u_t from the test to predict the response of the system using the model. The differences between the measured and the estimated system outputs can be used to estimate the degradation in the system performance caused by the fault. This information can be used afterwards to optimize maintenance and production scheduling taking into account the effect of running the system under faulty conditions.

In order to select the model order r and ensure its accuracy, the total averaged error presented in (5-6) was calculated for a range of values of r using data set T1 to build up a model and T2 and T3 to estimate the model error. The objective of this analysis is to select the model order that minimizes the prediction error. Fig. 77 shows the results obtained from this analysis.
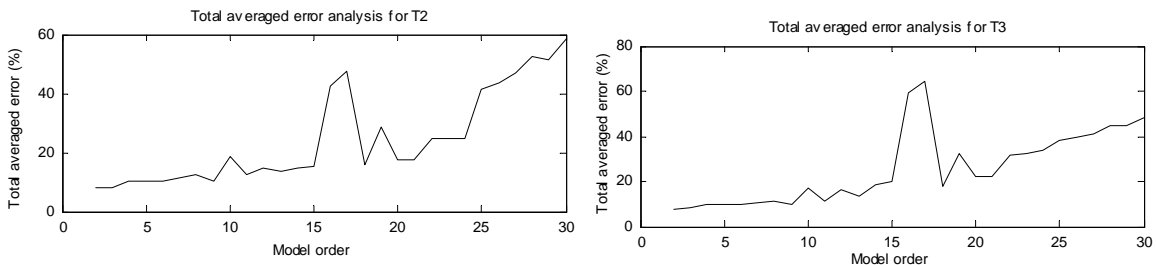
Fig. 77: Model order analysis

The results show that the averaged error for all variables tends to grow as the model order increases. This effect is caused mainly by the inaccurate initial estimation of the states. This causes oscillations in the output prediction which grow with the model order due to the increment in the number of degrees of freedom in the model. This effect can be seen in Fig. 78, which represents the prediction of PT312 for model orders 2, 3 and 4.
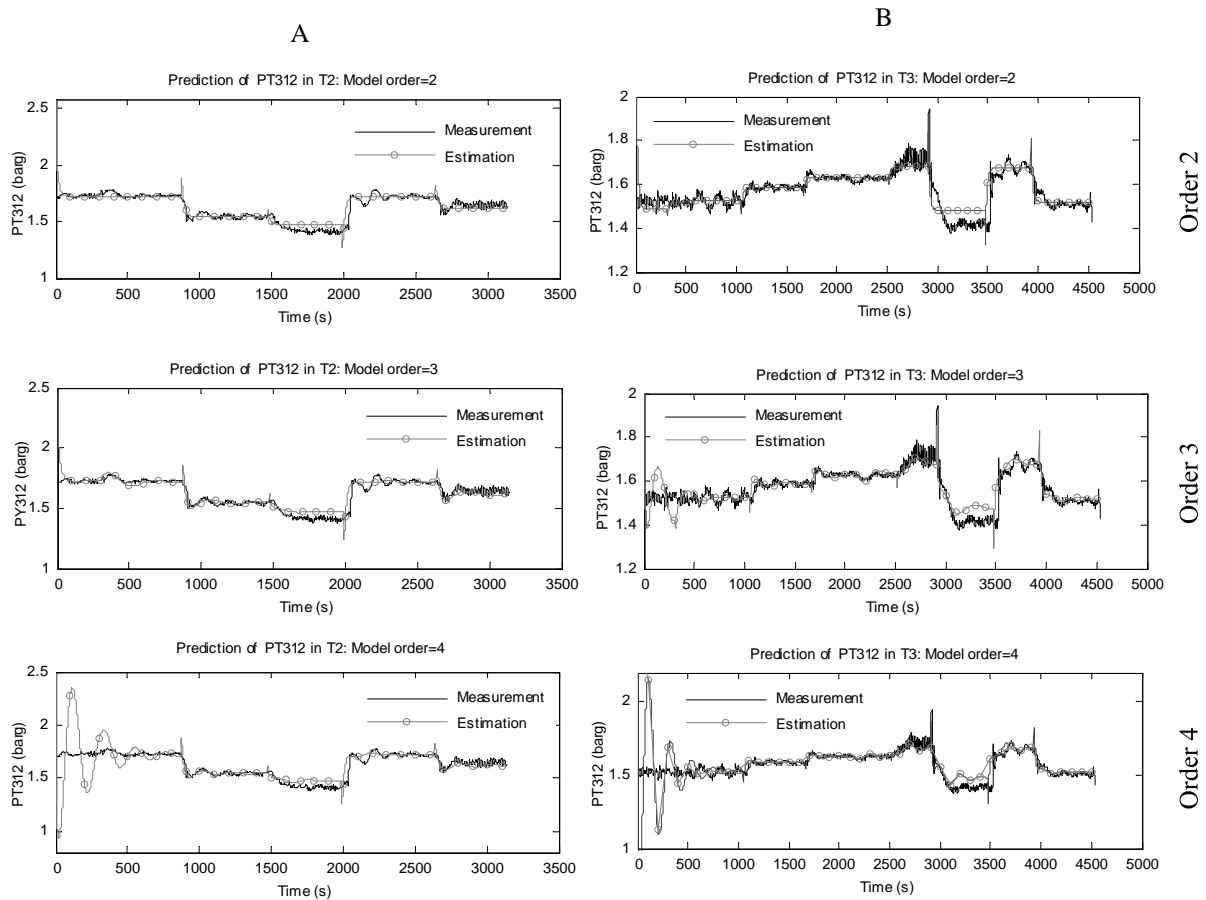
Fig. 78: Effect of model order (2, 3 and 4) over PT312 prediction accuracy for T2 (A) and T3 (B)

The results plotted in Fig. 78 show that the model of order 2 produces very low oscillations in the initial estimations, but the model ability to represent fast changes is quite limited due to the model restrictions. For model order 3 the initial oscillations are larger than for model 2, but the oscillations are dissipated relatively quickly and the estimation provided is more flexible and able to represent better the system behaviour. In model order 4 (and larger) the initial oscillations are even higher, and the attenuation takes more time. Once these initial oscillations disappear the results provided are accurate, but this transient phenomenon increases the total averaged error as shown in Fig. 77.

All the model orders tested were able to estimate the changes in the measured variables caused by changes in the operating conditions with relatively high accuracy. However, very fast random oscillations such as the first 1000

samples of T3 and around sample 2600 to 2900 in T3 cannot be represented. For this investigation, model order 3 was selected in order to combine both flexibility to represent system dynamics and accurate initial estimations. The prediction results obtained for the rest of variables (2 to 17) are summarized in Appendix A for T2 and Appendix B for T3. Using this model structure, the average normalized error calculated for each measured variable in T2 and T3 is represented in Table 23.

Table 23: Average normalized estimation error for each measured variable

Set	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	Avg
T2(%)	1.86	1.92	2.19	2.16	1.32	4.70	47.37	3.32	2.20	11.13	39.32	4.58	0.64	3.09	3.56	2.05	5.17	8.42
T3(%)	1.93	2.01	3.07	3.00	2.06	6.31	37.00	6.57	5.77	8.68	42.77	3.93	0.81	6.01	7.12	3.13	5.35	8.97

The estimation error obtained for each of the variables in T2 and T3 was of the same order of magnitude, proving that the model is able to represent the system working in conditions that were not tested during the training phase. The error is relatively much larger in variables number 7 and 11 (Differential pressure over VC404 and Density measured in FT407) due to the noise content of these variables, which cannot be accurately represented by the model.

The validated model was used to provide estimations of the process variables assuming normal operation in each one of the faulty cases studied. The model representing the system under normal operation was fed with the same input sequence used during the experiments in order to obtain estimations of its performance for the same operating conditions that were tested in the faulty cases. The objective of this analysis is to allow the process operators to estimate the impact of the fault over the system performance and take into account the effects in terms of safety, efficiency and product quality when scheduling the production and maintenance plans according to the process condition.

Based on the results provided by the contribution plots (see Fig. 76) the most significant variables in terms of fault detection in Case 1.1 and Case 1.2 were the pressure in the 3-phase separator PT501 and the position of the valve VC501. Fig. 79 represents the measurements observed for these variables in

both faulty cases and the estimation using a normal operation model. It can be seen that after the fault introduction the changes in operating conditions generate variations in PT501 (which in normal conditions is maintained at 1barg) and the algorithm was able to estimate VC501 valve position assuming normal operation. The estimations obtained for both variables before the fault introduction were accurate as expected after seeing the model validation results.
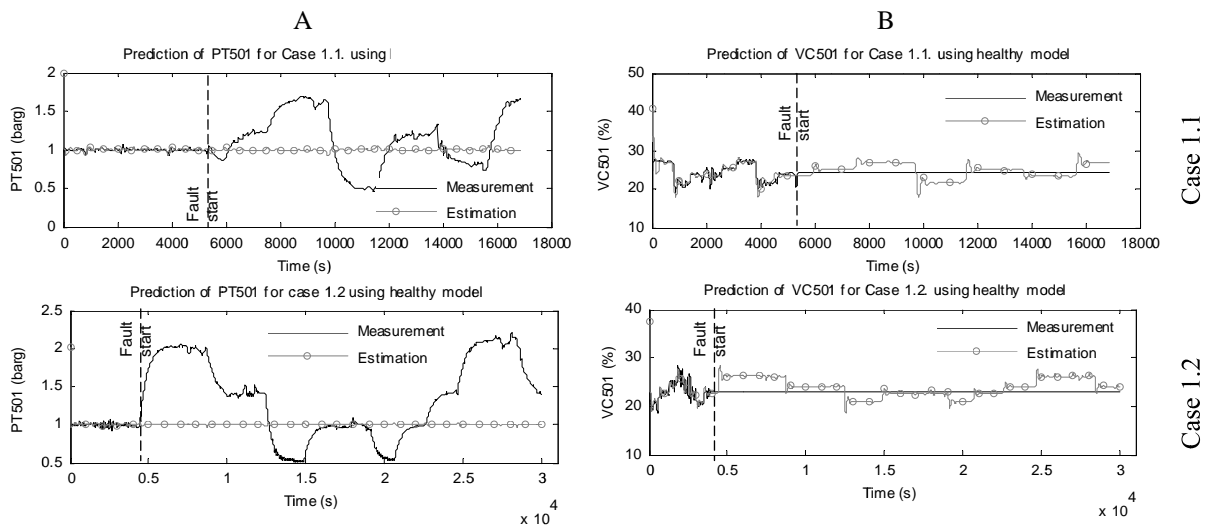
Fig. 79: Performance degradation in PT501 (A) and VC501 (B) in Case 1.1 and 1.2

For Case 2 two of the most affected variables were the pressure measured in the bottom of the riser (PT401) and the differential pressure over VC404 (PT408-PT403). The results obtained for prediction of these two variables in Case 2 assuming normal operation are represented in Fig. 80.
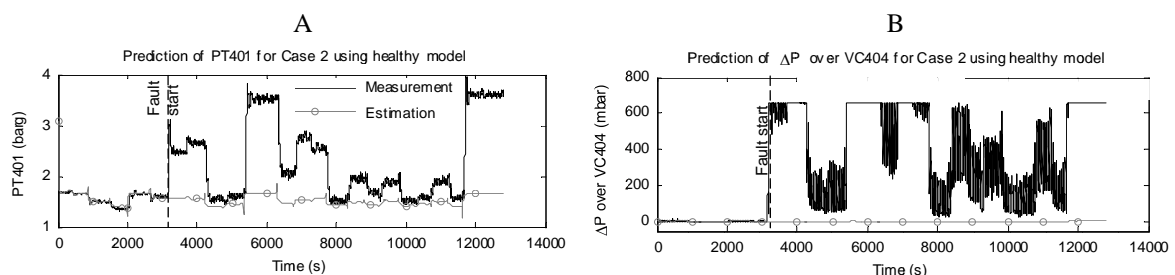
Fig. 80: Performance degradation in PT401 (A) and differential pressure over VC404 (B) in Case 2

Again the estimation of both variables was accurate until the introduction of the fault. Thereafter the differences between the measured and the estimated

values for each variable were significantly high due to the effect of the fault over the system performance. The main changes observed after the introduction of the fault were an increment of the pressure measured in the sensors placed before the blockage (PT312, PT401 and PT408) and a huge increment in the differential pressure measured between PT408 and PT403. This variable reaches values of around 5 mbar when the system is working under normal operation, but after the blockage was introduced it reached values up to 700 mbar depending on the operating conditions.

The most significant variable for the faulty cases 3.1 and 3.2 was the top separator pressure PT403 for both indicators, although the contribution of the top separator level LIC405 was also significant for Q. The results obtained for the estimation of these variables using a healthy model are represented in Fig. 81. The estimations were accurate in both cases until the fault was introduced, and the most obvious effect of the fault was an increment in the pressure measured inside the 2-phase separator (PT403). Changes in the behaviour of the level measurement inside the 2-phase separator are not that evident due to the high level of noise in this signal, but a slight increment in the average level over the estimated value can be perceived after the fault introduction.
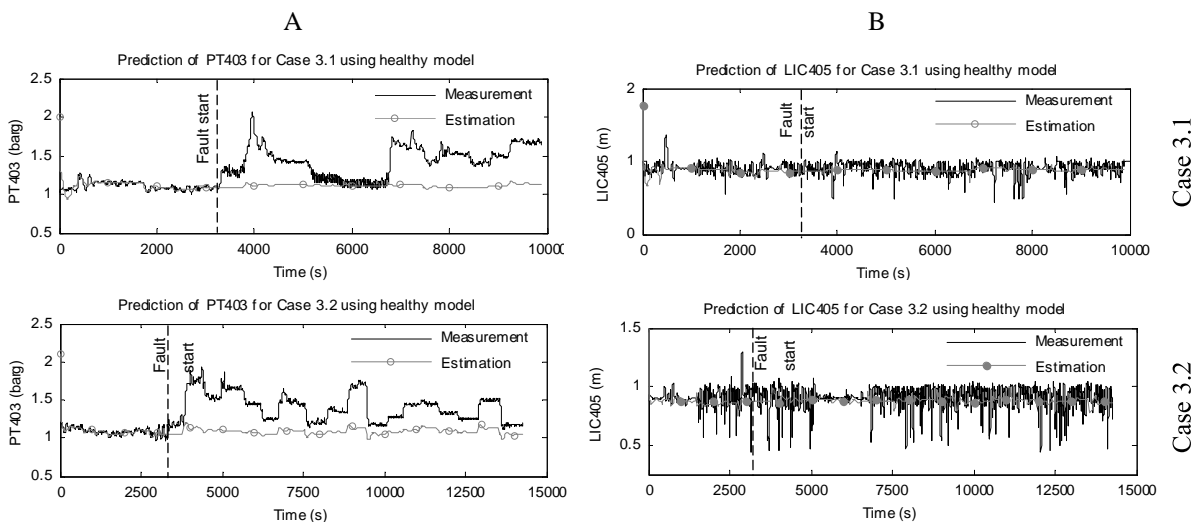
Fig. 81: Performance degradation in PT403 (A) and LIC405 (B) in Case 3.1 and 3.2

For the last case studied, the most significant variables identified for fault diagnosis were the top riser pressure PT408 and the density measured by

FT407 in the top of the riser. Fig. 82 shows the measurements of these two variables in Case 4 and the corresponding estimation using a healthy model. It can be seen that the initial opening of the 2" valve did not have a significant effect on the measured variables, except for a slight increment in the oscillation in the PT408 signal. These oscillations are more evident around sample 5800 in both measurements, probably caused by the change in the water flow rate at that point. Closing VC404 to derive all the flow through the 2" line had a considerable effect on PT408 in comparison with the estimation assuming normal operation, as well as on the density measurement which acquired an almost constant value of 998kg/m^3 due to the accumulation of liquid in the sensor region.
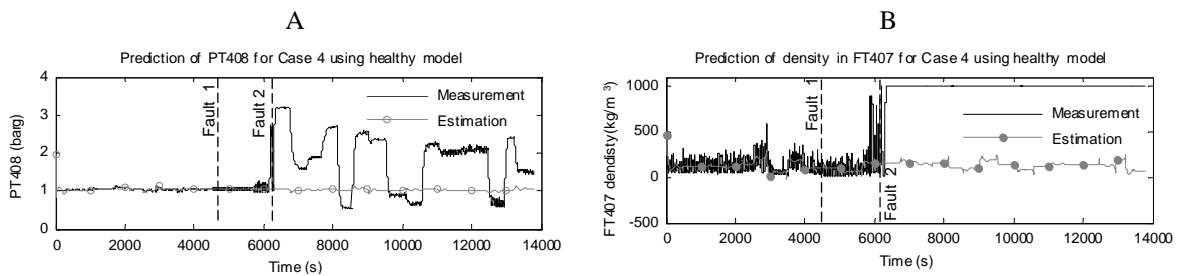
Fig. 82: Performance degradation in PT408 (A) and density measured in FT407 (B) in Case 4

5.4.4 Prediction of performance under faulty conditions

The analysis carried out in the previous subsections described how process faults can be detected and diagnosed, and how the degradation of the system performance can be estimated using CVA. If the fault severity is not critical, it can be the case that the optimal maintenance and operation strategy is to continue operating the system under faulty conditions until the next planned shutdown or until the spare parts and repairing equipment are available in the plant. In that case the plant operators will need to know how the faulty process will behave for the future expected operating conditions and determine how this behaviour affects the product quality, the safety of the plant or the energy consumed by the equipment. For example it can happen that a minor blockage has a minimal effect on the pressure inside the pipelines, the energy consumed by pumps and compressor and their capacity to deliver the desired flow rate.

151

However, if in the upcoming days or weeks the production plan requires operating the system with higher flow rates the additional pressure loses generated by the blockage can increase the pressure downstream to dangerous levels, increase the energy consumed by the devices providing raw material and affect the quality of the product if the desired flow rate is not achieved.

The faulty system can be modelled using data acquired during the early stages of degradation once a fault has been detected. The forecasted input sequences can be obtained from the production plan, from estimations based on historic data, etc. and then be used in the model to predict the system behaviour for that particular conditions. For this analysis the data acquired from the instant of fault detection was used to build a model for each one of the faulty cases studied. The amount of data used to build the model needs to be sufficient to represent the system dynamics and produce an accurate model. There is a numerical limitation on the minimum amount of data needed in relation to the matrix inversion in (5-4). If the system inputs contain constant values for each operating condition (as it happened in this case) it is necessary to capture data from at least two different operating points in order to avoid rows and columns containing zeroes in the covariance matrix $Cov([z_t,u_t],[z_t,u_t])$.

In order to ensure accuracy in the prediction the total averaged error will be computed for different lengths of training data, and the model will be considered valid once the error of the model predicting the same data segment used for training is considered acceptable. Fig. 83 shows the evolution of the total averaged error with the number of samples selected for model training in each case studied. The evolution of the error is similar in all the cases studied except in Case 2; initially the error is large and as the number of samples used for training is increased the error decreases to values of around 8.5%, similar to the error found under normal operating conditions. In Case 2 the error is initially lower probably due to the operating conditions during the fault introduction, but as the number of samples grows and the diversity of the data increased the error tended to values of around 8% as in the rest of cases. It is important to mention that for the analysis of Case 1 it was necessary to remove the variable

152

VC501 (y_{14}) from the analysis, as the series of constant values generated after switching the valve to manual mode caused numerical problems in the matrix inversion required in (4-10).
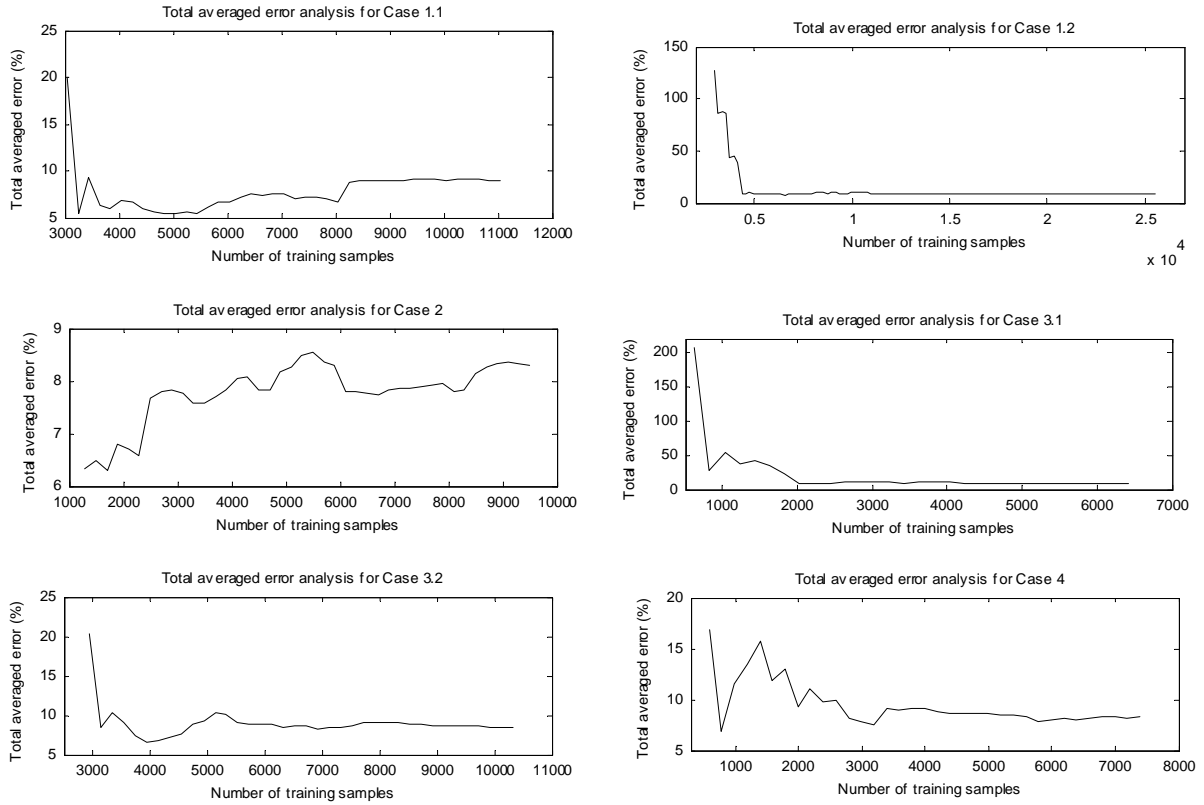
Fig. 83: Evolution of the total averaged error with the number of samples selected for model training

The total averaged error shown in Fig. 83 was calculated as the error of the model predicting the same data set portion used for model training so that the operator can estimate in real time model accuracy in order to perform a reliable long time prediction. One model was built for each one of the cases studied using a limited amount of data acquired after the fault detection according to the error estimations shown in Fig. 83. This model was then used to perform a long time prediction of all the samples available in each data set, so that the model built with data acquired from the early stages of degradation can be used to predict future process performance under different operating conditions. Table 24 shows the normalized estimation error calculated for each variable of each case studied for the long term prediction. The starting point and the number of samples used to train the model are presented in columns 2 and 3 respectively.

Table 24: Average normalized estimation error for each measured variable (%)

Set	Start	Samples	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	Avg
1.1	5688	5000	3.05	3.21	4.47	4.41	4.71	2.87	27.54	1.19	1.82	7.85	41.36	2.21	1.17	-	1.54	1.39	5.25	7.13
1.2	4478	5000	7.16	7.35	11.35	11.11	11.79	4.93	32.68	1.21	1.74	8.32	41.18	2.57	0.46	-	1.96	2.50	6.70	9.56
2	3170	3000	5.72	5.90	7.36	3.71	2.41	15.31	58.94	1.42	1.59	8.06	37.85	1.58	1.25	4.96	1.63	5.34	5.03	9.89
3.1	3392	2000	6.62	6.86	8.78	8.60	3.65	7.47	33.07	1.24	1.65	8.73	44.48	2.36	1.12	8.06	1.15	3.69	4.97	8.97
3.2	3821	5000	6.19	6.38	8.51	8.28	2.71	2.73	31.55	2.34	0.63	8.93	32.67	1.07	1.75	7.28	2.99	1.50	6.69	7.78
4	6187	3000	5.45	6.21	11.20	2.87	2.61	3.08	64.62	1.36	5.88	14.03	77.46	1.08	1.62	5.99	1.87	5.11	6.80	12.78

All the cases showed similar error rates than the estimations obtained from modelling normal operating conditions, showing that if enough data is used to train the model it is possible to represent the process behaviour under different operating conditions. Fig. 84 shows the prediction results under faulty conditions of the most significant variables of each case. As mentioned before, it was not possible to produce results for VC501 in cases 1.1 and 1.2 due to its constant value once the fault was introduced. All the predictions obtained show a similar behaviour than the estimations under normal operating conditions, with some oscillations in the initial estimations but a good accuracy once the transient is extinguished. This result shows that it is possible to perform a reliable long time prediction using a small amount of data acquired during the early stages of degradation, and this prediction will be reliable if the forecast of future inputs is known. This prediction can be used by plant operators to estimate how the faulty system will react to different operating conditions, and determine whether it is safe and appropriate to operate the system in these conditions.
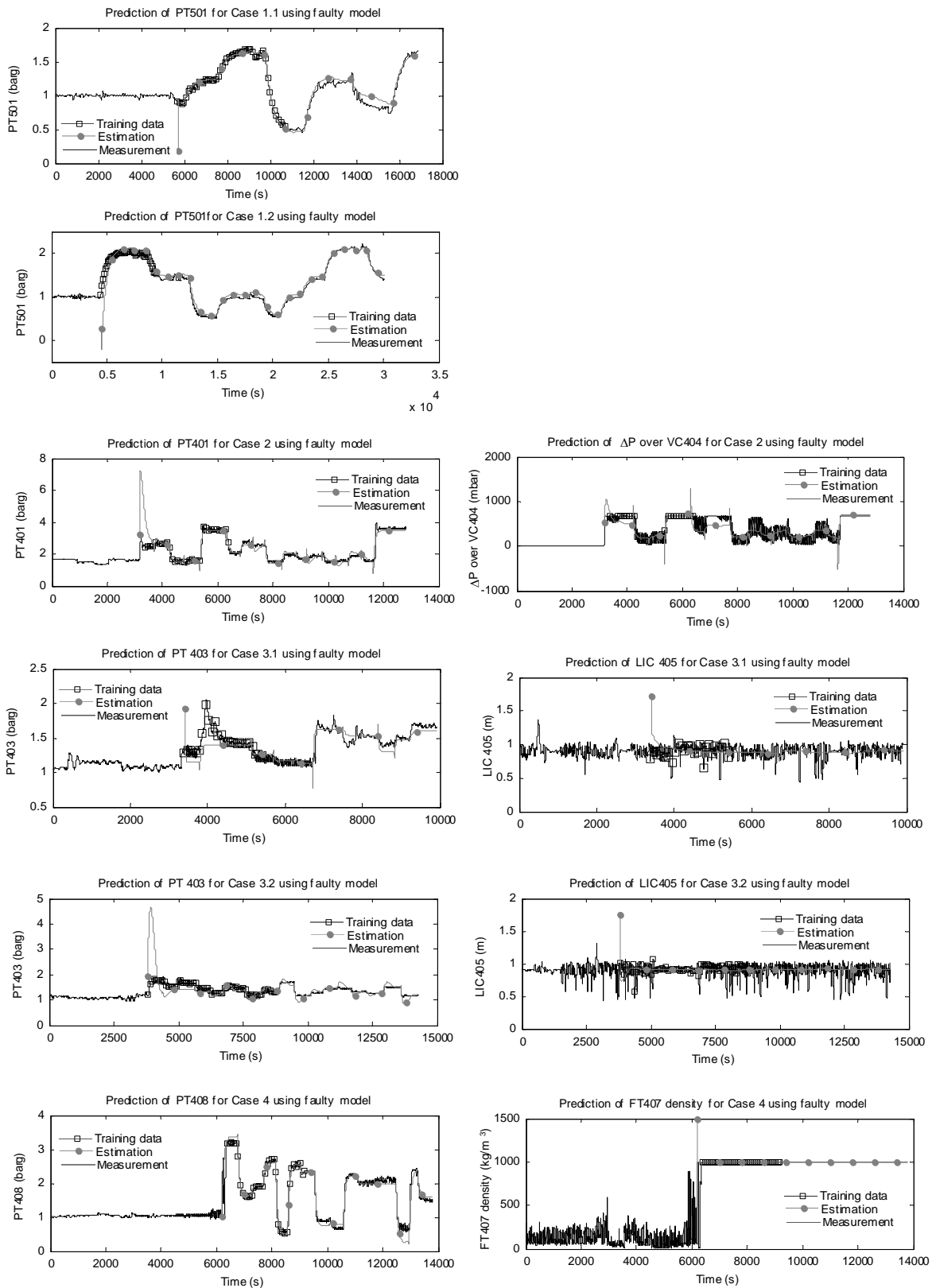
Fig. 84: Summary of prediction results under faulty conditions for the most significant variables in each case

155

5.5 Conclusion

Process data was acquired from an experimental large-scale multiphase flow facility to test the capabilities of CVA to detect and diagnose process faults, as well as to estimate performance degradation and predict the behaviour of the system working under varying operating conditions. This system is non-linear and was operated under changing operational conditions.

The four different faults introduced were successfully detected using the T^2 and Q indicators within a reasonable detection time, low number of false alarms and false negatives. The most significant variables affected by the fault were identified using contributions plots to help in the process of fault diagnosis. Once the faults were detected, CVA was used to build a state-space model that represents the system working under normal operating conditions. This model was used to observe the difference between the process measurements when working under faulty conditions and the model estimations assuming normal operation. The estimations obtained were close to the process measurements before the introduction of the faults, and the differences observed afterwards can be used to quantify the impact of the fault on the different process variables. Secondly, a new model was built for each case studied using data acquired during the early stages of degradation after the fault introduction. This model was used to predict the faulty process performance for different operating conditions.

The system identification procedure based on CVA was performed in a fast and stable manner, due to the numerical and computational benefits of singular value decomposition. The estimation error was always relatively low, allowing an accurate estimation of all process variables under normal and faulty operation. The model was able to represent the process dynamics working under varying operating conditions, however very fast oscillations and noisy measurements were not precisely estimated.

The results obtained from this investigation demonstrate that CVA can be used for system identification using real process data acquired from a large and complex facility. This technique allowed the estimation of performance

156

degradation as the difference between actual measurements and estimations provided by a model trained under normal operation. This information can be used by plant operators to measure the impact of the fault on the process performance and take it into account for scheduling optimal maintenance and production plans that consider the condition of the process. In addition, the behaviour of the faulty process was modelled using data acquired during the early stages of degradation, allowing the operators to predict how the fault will affect the process for future operating conditions. This methodology advances the traditional condition monitoring procedure of fault detection and diagnosis, and provides estimations of the impact of the fault on the system behaviour that can be used to react to the faults with the optimal maintenance and production strategies leading to more efficient, reliable and profitable processes.

6 APPLICATION OF LINEAR PREDICTION, SELF-ADAPTIVE NOISE CANCELLATION AND SPECTRAL KURTOSIS IN IDENTIFYING NATURAL DAMAGE OF A ROLLING ELEMENT BEARING IN A GEARBOX

Abstract

The ability to detect and diagnose faults in rolling element bearings is crucial for modern maintenance schemes. Several techniques have been developed to improve the ability of fault detection in bearings using vibration monitoring, especially in those cases where the vibration signal is contaminated by background noise. Linear Prediction and Self-Adaptive Noise Cancellation are techniques which can substantially improve the signal to noise ratio of the signal, improving the visibility of the important signal components in the frequency spectrum. Spectral Kurtosis has been shown to improve bearing defect identification by focusing on the frequency band with a high level of impulsiveness. In this paper the ability of these three methods to detect a bearing fault is compared using vibration data from a specially designed test rig that allowed fast natural degradation of the bearing. The results obtained show that the Spectral Kurtosis was able to detect an incipient fault in the outer race of the bearing much earlier than any other technique.

6.1 Introduction

Rolling element bearings are important components in rotating machinery. By monitoring the vibration signature of bearings, it is possible to obtain important information about their condition, and use this information to improve the maintenance strategy. Diagnostic techniques based on vibration are mainly concerned with the extraction of defect features in the acquired signal, which can be related to the healthy or defective state of vital parts in a machine. Many different diagnostic methods have been successfully used to identify machine faults, processing the vibration signal in the time or frequency domain, in order to locate and quantify any existing damage. In complex machines the signal acquired is normally inclusive of additive background noise from other machine

components or subsystems, which can make it difficult or sometimes impossible to identify the fault patterns in the signal.

In the case of bearings, the fault is produced typically by the damage of the surface of the inner or outer race of the rolling elements. When a damaged surface contacts another rolling surface a force impulse is generated which excites resonances in the bearing and the machine [88]. The successive impacts generate a vibration signal which often has an impulsive repetitive nature that is easy to identify in the presence of low background noise. In a real machine, the background noise can mask the bearing fault components of the signal, especially in gearboxes because the gear meshing can generate a strong level of vibration [89]. For this reason many different signal processing methodologies have been developed in order to facilitate the detection of defects particularly in bearings.

Some examples of classic techniques used to enhance bearing fault features in vibration signals are Linear Prediction (LP), Self-Adaptive Noise Cancellation (SANC), Cyclostationarity, Hilbert-Huang Transform (HHT) or Wavelet Transform (WT). LP is based on the estimation of the deterministic part of a signal as a linear combination of past inputs and outputs of the system while SANC aims to minimize the noise in the manipulated signal by recursively adapting the filtration parameters [90]. Cyclostationarity studies the periodicities of the different features of machine vibration signals using cyclic autocorrelation function and spectral correlation density [91]. HHT can be used to decompose a non-stationary and nonlinear signal into intrinsic mode functions and obtain instantaneous frequency data [92], and WT can be applied on non-stationary signals to increase the frequency resolution at low frequencies and reduce noise in raw signals[93]. All these techniques have been already applied by various researchers for the detection and diagnosis of bearing and gearbox faults.

In this investigation three diagnostic techniques, LP, SANC and Spectral Kurtosis (SK) were applied in identifying a bearing defect in a gearbox where the bearing degradation happened naturally in a specially designed test rig. LP

and SANC have been successfully used as denoising tools in different applications for many years [94; 95]. Nevertheless, even nowadays many researchers are exploring their capabilities to reduce background noise and enhance the fault features in a signal to improve the fault detection and diagnosis in bearings [96-102]. On the other hand, SK is a relatively new methodology which is able to enhance the fault signature in a signal by focusing in the frequency band with a higher level of impulsiveness [103-105]. This technique has demonstrated to be very effective especially for bearing fault detection, and many researchers have reported its benefits [106-111].

The aim of this chapter is to compare the performance of these methodologies in detecting a bearing fault during the early stages of natural degradation and show the benefits of SK over more stablished denoising techniques. For this purpose, these three methodologies have been applied on a vibration signal acquired from a particular gearbox where the bearings failed much earlier than the theoretical life calculated for certain loading conditions. Analysis of acquired vibration signals associated with different stages of bearing degradation proved to be ideal for this comparative study. This was principally because the bearing defect frequency was only evident at the final stage of degradation. Thus the study presented will explore if these techniques can offer the ability to identify earlier the presence of the defect.

6.2 Theoretical background

6.2.1 Linear Prediction

The estimation of a dynamic system output and its later analysis is one of the most important problems in signal processing. Different techniques have been employed by several researchers in a wide range of applications such as neurophysics, electrocardiography, geophysics and speech communication [94]. One of the most powerful estimation models is based on the assumption that the value of a signal x at the time n can be obtained as a linear combination of past inputs and outputs of the system. Those models which use the information from only the past system outputs are called all-pole or autoregressive models, and were first used by Yule [112] in an investigation of

161

sunspot numbers. LP is one of those methods where the objective is to predict or estimate the future output of a system based on the past output observations. The complete mathematical development and a compilation of the different LP approaches have been presented by Makhoul [94].

In vibration based diagnostics, LP is a method that allows the separation of the deterministic or predictable part of a signal from the random background noise using the information provided by past observations [101; 113]. If it is assumed that the background noise is totally random, applying this method it is possible to eliminate the background noise and thus improve the signal to noise ratio. This technique is based on the principle that the value of the deterministic part of a signal can be predicted as a weighted sum of a series of previous values:

$$\hat{x}(n) = -\sum_{k=1}^{p} a(k) \cdot x(n-k)$$

(6-1)

Where $\hat{x}(n)$ is the predictable part of the nth sample of the signal x, p is the number of past samples considered and $a(k)$ are the weights attached to each past observation. The weighting coefficients can be obtained at each step n, by a linear operation from the autocorrelation function R_τ of the time series $x(n)$, which can be efficiently solved using the Yule-Walker equation [114]:

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_{p-1} \\ R_1 & R_0 & \cdots & R_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ R_{p-1} & R_{p-2} & \cdots & R_0 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} -R_1 \\ -R_2 \\ \vdots \\ -R_p \end{bmatrix}$$

(6-2)

Where:

$$R_\tau = \frac{1}{N} \sum_{t=\tau}^{N} x(t-\tau) \cdot x(t)$$

(6-3)

N is the number of past samples considered at each step, in this case only p past samples were considered for each prediction for computational reasons, but all the available past samples at each time point were used in the calculation of the values R_τ.

The results of the algorithm depend on the number of past observations p considered. Small values of p produce a poor prediction, giving a result of

negligible improvement in the signal to noise ratio, while very high values of p affect negatively to the computational cost, over restrain the prediction and tend to reduce even the main components of the signal. For this particular investigation several analyses were carried out using different numbers of past samples, in order to establish the value p for each test case which optimizes the signal to noise ratio of the output signal.

6.2.2 Self-Adaptive Noise Cancellation

Adaptive Noise Cancelling (ANC) is another technique used to reduce the background noise in a signal and increase the signal to noise ratio, improving the visibility of the different signal components in the frequency spectrum. The first work in ANC was performed by Howells and Applebaum at the General Electric Company between 1957 and 1960. The first ANC system was designed and built at Stanford University in 1965 [95]. Since then, this method has been successfully applied to a number of additional problems including electrocardiography, cancelling noise in speech signals, cancelling antenna sidelobe interferences, etc.[95]

The general ANC concept is shown in Fig. 85, and a basic explanation of the method was given by Chaturvedi et al. [115]: the input $x(n)$ composed by the signal of interest S and additive noise n_0 is received at the primary sensor. A reference noise n_1 (which must be related to the noise n_0 in some unknown way but is not coherent with the signal S) is received at the reference sensor. The reference input is then adaptively filtered to match n_0 as closely as possible, which is then subtracted from the primary input $x(n)=S+n_0$ to produce the system output $e=S+n_0-y$. This output contains the signal plus residual undesirable noise. The adaptive filter acts minimizing, indirectly, the average power of this residual noise at the system output e. The output is fed back to the adaptive filter and the filter weights are adjusted at each calculation step to minimize the total output power of the system. It can be demonstrated that minimizing the total output power minimizes the output noise power or, in other words, maximizes the output signal to noise ratio.[95]
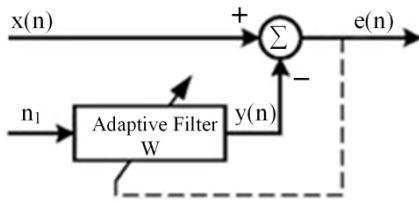
Fig. 85: ANC algorithm
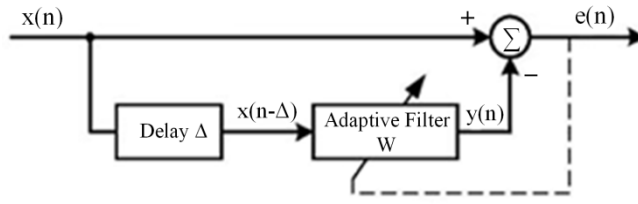
Fig. 86: SANC algorithm

The problem of this method applied to bearing fault detection in real applications is that it is not always easy to identify the source of noise n_1 which is correlated with the noise n_0 (common source) but not with the fault signal. Chaturvedi et al.[115] presented an example where the method was applied to detect an induced bearing fault in a gearbox using two sensors; one was placed in the surroundings of the bearing housing to obtain the main signal and another sensor was placed at a remote location in the casing of the gearbox to obtain the reference signal. To solve this issue, a further development of ANC was formulated using a delayed version of the primary signal [95]. This latter version was named the SANC and the schematic concept is represented in Fig. 86. The time delay Δ which is fixed forces the delayed version of the input signal to become uncorrelated with the primary signal introducing a phase difference. The adaptive filter responds firstly by compensating for the phase shift so that the sinusoidal components cancel each other at the output, and secondly by removing as much noise as possible to minimise the output error [116]. As it happens in the original ANC, the output error is then fed back to the adaptive filter to adjust recursively the filter weights w in order to minimize the total output power and thus, the output noise power. There are many adaptation rules to do this; the most well-known is the least mean square [95]:

$$w_i^{n+1} = w_i^n + \mu \cdot e(n) \cdot x(n - \Delta - i)$$

(6-4)

Where the parameter μ (forgetting factor, strictly positive) controls the stability and rate of convergence of the process and the subscript i differentiates each of the H weighting coefficients of the filter. The recursive weight calculation starts with a random value for each weight w_i. The output of the filter $y(n)$ can be calculated as:

$$y(n) = W^T(n) \cdot X(n - \Delta) \tag{6-5}$$

where W is a vector containing the H weighting coefficients w_i and $X(n\text{-}\Delta)$ is another vector containing the H components of the delayed signal immediately preceding the sample n. The output $e(n)$ is easily obtained from:

$$e(n) = x(n) - y(n) \tag{6-6}$$

As shown by equation (6-4), the performance of the SANC algorithm clearly depends on the choice of three parameters: the time delay Δ, the filter length H and the forgetting factor μ. The influence of these parameters was investigated by Ho et al.[117] who suggested some parameter selection guides. Δ should be large enough to ensure that the delayed signal becomes uncorrelated with the original, and H should be chosen to cancel all the broadband components of the delayed signal. In both cases, if the selected value for the parameter is too large it will lead to computation problems. Ho [117] stated the forgetting factor depends mainly on the filter order H. In this particular investigation the parameters Δ, H and μ were selected after several tests with the aim of optimizing the signal to noise ratio of the output signal. From these tests it was concluded that the selection of μ is crucial for the process performance: very small variations on this parameter can change the output signal, from no noise reduction effect if the selected value is too high, to distortion of the main signal components when it is too low. This influence can be seen in Fig. 87 where the SANC was applied to a representative signal acquired during the tests using different values for the forgetting factor μ. It is always important to check the convergence of the filter weights to ensure optimal performance.
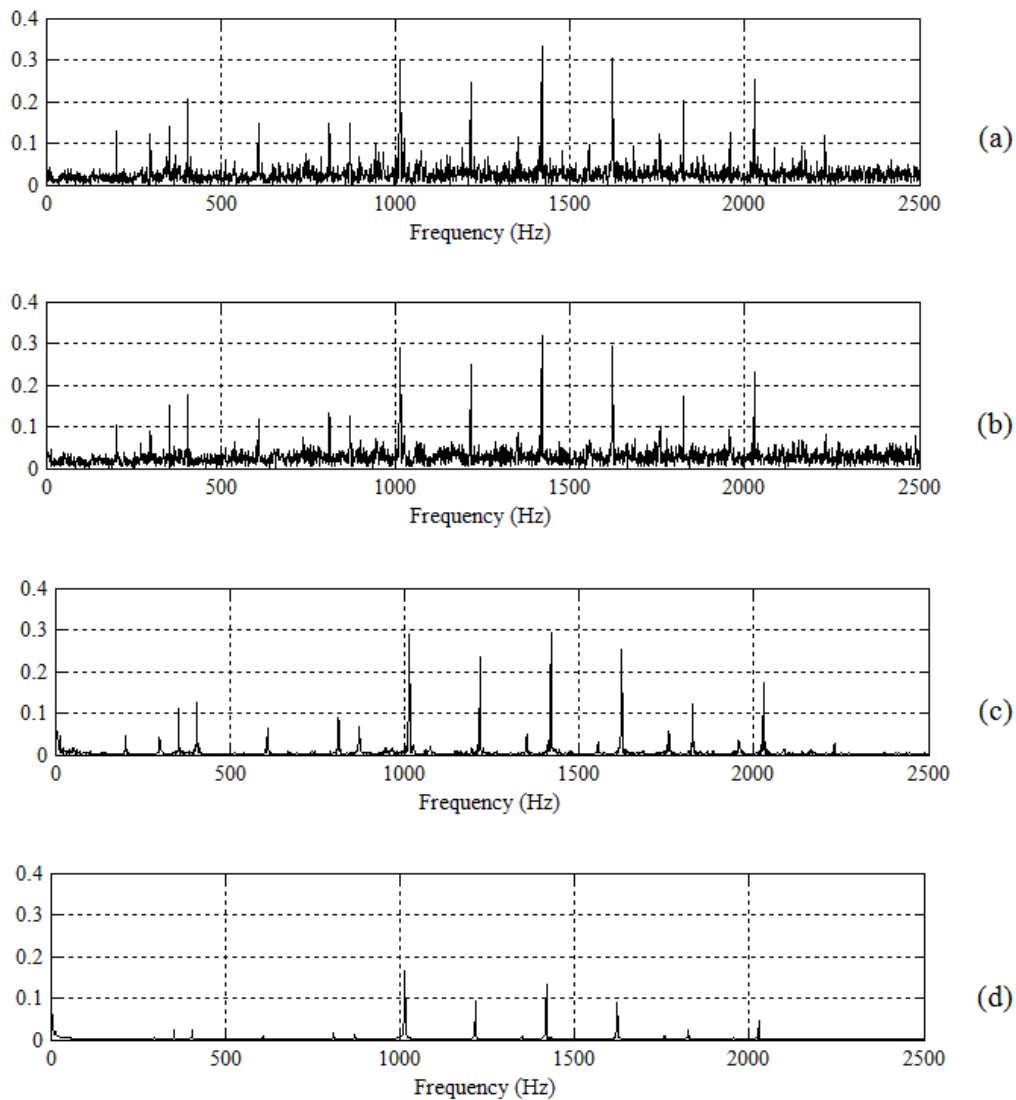
Fig. 87: Effect of the forgetting factor μ on the SANC results. (a):original spectrum; (b) signal processed through SANC using μ=0.0001; (c) signal processed through SANC using μ=0.00001; (d) signal processed through SANC using μ=0.000001

6.2.3 Spectral Kurtosis and Envelope Analysis

Kurtosis is defined as the degree of peakness of a signal with probability density function *p(x)*, and mathematically it is defined as the normalized fourth moment of a probability density function [118]:

$$K = \frac{\int_{-\infty}^{\infty} [x - \mu]^4 p(x)dx}{\sigma^4}$$

(6-7)

Where x is the signal of interest with average μ and standard deviation σ.

As mentioned earlier, in real applications background noise often masks the signal of interest and, as a result, the Kurtosis is unable to capture the peakness of the fault signal, giving usually low Kurtosis values. Therefore, in applications with strong background noise, the Kurtosis as a global indicator is not useful, and it gives better results when it is applied locally in different frequency bands [104]. This technique is named Spectral Kurtosis (SK).

The SK was first introduced by Dwyer [119] as a statistical tool which can locate non-Gaussian components in the frequency domain of a signal. This method is able to indicate the presence of transients in the signal and show their locations in the frequency domain. It has demonstrated to be effective even in the presence of strong additive noise [104]. The basic principle of this method is to calculate the Kurtosis at different frequency bands in order to identify non stationarities in the signal and determine where they are located in the frequency domain. Obviously the results obtained strongly depend on the width of the frequency bands Δf in which the analysis is performed and its influence was analysed by Antoni [105].

The Kurtogram is basically a representation of the calculated values of the SK as a function of f and Δf [120]. However, the exploration of the whole plane ($f, \Delta f$) is a complicated computation task difficult to deal with, though Antoni [105] suggested a methodology for the fast computation of the SK. In this approach, at each bandwidth level the number of filtered sequences is increased by a factor 2, and the Kurtogram is finally estimated by computing the Kurtosis of all sequences.

The importance of the Kurtogram relies on the fact that it allows the identification of the frequency band where the SK is maximum, and this information can be used to design a filter which extracts the part of the signal with the highest level of impulsiveness. Antoni et al.[104] demonstrated how the optimum filter which maximizes the signal to noise ratio is a narrowband filter at the maximum value of SK. Therefore the optimal central frequency and bandwidth of the band-pass filter are found as the values of f and Δf which

maximise the Kurtogram. The filtrated signal can be finally used to perform an envelope analysis, which is a widely used technique for identification of modulating frequencies related with bearing faults. In this investigation the SK computation and the subsequent signal filtration and envelope analysis was performed using original Matlab code programmed by Jérôme Antoni [121].

This investigation assesses the merits of these three techniques in identifying a natural degraded bearing under conditions of relatively large background noise.

6.3 Experimental set up

The vibrational data used in this investigation was obtained from a specially designed gearbox test rig. The gearbox type employed is a part of the transmission driveline on the actuation mechanism of secondary control surfaces in civil aircrafts. The bearing of this gearbox failed in an endurance test at around 30% of its total expected life (around 3000 hours), making it an ideal candidate for this investigation where fast natural degradation of the bearing was needed. The rig was built originally to identify the origin of premature failure in order to modify the gearbox design. The acquired vibrational signal was used in this investigation to find traces of the fault during the early stages of degradation, which is an obvious advantage from a maintenance point of view.

This gearbox, whose basic cross section is shown in Fig. 88, has two bevel gears with 17 teeth on each gear, generating a transmission ratio of 1:1. Each gear is supported by two angular contact bearings with 12 balls each and a contact angle of 40°, mounted in a back-to-back configuration. The main dimensions of the bearing and the attached bearing defect frequencies can be seen in Table 25 and Table 26 respectively. The test rig was built trying to emulate the actual transmission system used in the aircraft, and it is schematically represented in Fig. 89. The transmission is driven by an electric motor with a nominal speed of 710 r.p.m. An electric load motor placed at the opposite side of the transmission line was used to apply different loads used during the experiment. In order to simulate the actual loading conditions expected during the life of the gearboxes, the test rig was subjected to a mixture of seven different types of flight load cycles derived from the actual flight data

and loads. These load cycles include the simulation of takeoff and landing with different flap positions, ground maintenance, etc. The expected bearing life for these loading conditions was around 3000 hours. Fig. 90 shows a typical type 3 load profile, which was chosen as an illustrative example because it contains several speed changes and the highest torque is applied in this particular load cycle. The loading conditions of each cycle type applied are explained in Table 27, which specifies the number of times each cycle was applied during the experiment for the expected bearing life, the duration of each cycle and the maximum toque applied in each case.
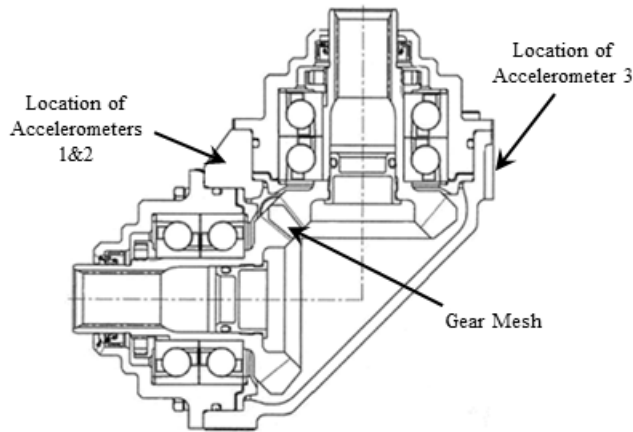
Fig. 88: Gearbox Section

Table 25: Bearing Main Dimensions

No. of rolling elements	12
Ball Diameter (B_d)	0.4063"
Contact Angle (Φ)	40°
Pitch Diameter (P_d)	1.811"
Input Shaft Speed (RPM)	710 rpm
Gear Teeth	17

Table 26: Main Defect Frequencies and Harmonics (Hz)

Harmonic	1X	2X	3X	4X	5X	6X
Shaft speed frequency (SS)	11.8	23.7	35.5	47.3	59.2	71
Gear mesh frequency (GM)	201.2	402	604	805	1006	1207
Inner race defect frequency(IRD)	83.2	166	250	333	416	499
Outer race defect frequency (ORD)	58.8	118	176	235	294	353
Cage defect frequency	4.9	9.8	14.7	19.6	24.5	29.4
Ball spin frequency	25.6	51.2	76.8	102	128	154
Rolling element defect frequency	51.2	102	154	205	256	307
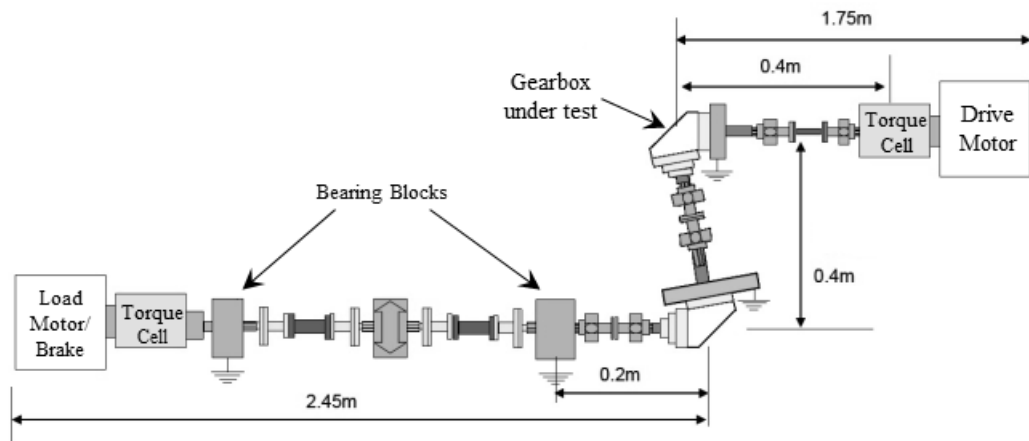
Fig. 89: Layout of the test rig
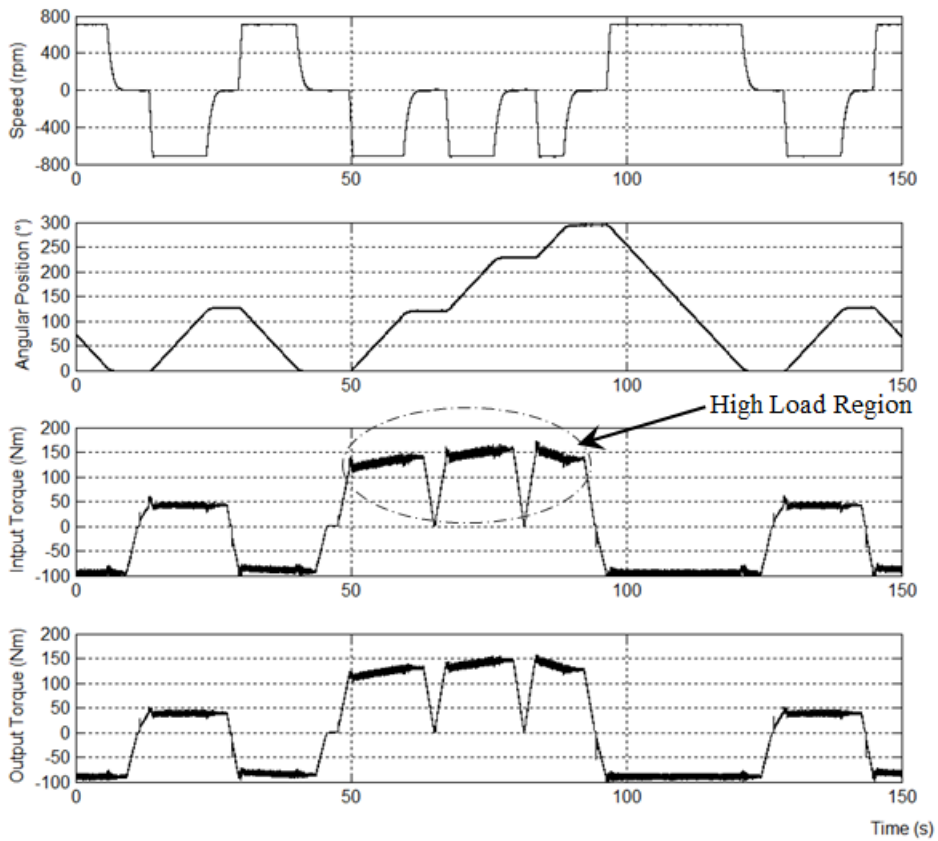
Fig. 90: Type 3 load cycle profile

Table 27: Load cycles characteristics summary

Cycle type	1	2	3	4	5	6	7	8	9
Number of repetitions during bearing life	18296	22869	4574	462	462	2200	6600	4620	41580
Duration (sec)	131	131	131	350	42	71	268	52	52
Torque max. (Nm)	126.1	126.1	158.6	126.1	126.1	42.8	42.8	12.4	97.7

The experiment ran continuously for 24 hours a day over a duration of 36 days, but at certain points during the test run the rig was stopped for visual inspection for damage in the bearings. The gearbox was always then reassembled and the sequence continued. Fig. 91 shows a detail of the bearing outer race during a visual inspection undertaken one month after the experiment started, covering 24% of the estimated bearing life.

Fig. 91: Detail of bearing outer race after one month

Three accelerometers were mounted in the gearbox at locations identified in Fig. 88, two of them placed on the top of the gearbox measuring acceleration in the vertical plane and a third one placed on the casing measuring the acceleration in the horizontal plane. The selected accelerometers (Omni Instruments model RYD81D) had an operating frequency range of 10Hz to 10 kHz. These accelerometers were connected to signal conditioners (model Endevco 2775A) which were attached to a NI USB 6009 data acquisition device. This digital data was filtered, windowed and stored in the computer using Dasylab version 10.0, and finally it was exported for its final manipulation in Matlab R2010A. Other than the vibration data, various parameters were

monitored and stored at the same time and with the same sampling frequency: angular position of the input shaft, input and output torque and shaft speed.

The experiment started running on the 19/07/2010 and the vibration measurements were taken on the 19/08/2010, 22/08/2010 and finally on the 24/08/2010. For each measurement case a total 1048569 points were acquired at a sampling rate of 5 kHz which resulted in a measurement length of approximately 3.5 minutes; sufficiently long to cover a whole loading cycle. The stored data was then analysed, selecting groups of 8192 data samples in the region of constant speed where the load applied was maximum (Fig. 90). After a preliminary data analysis it was decided to always use the signal acquired by the third channel in the next steps of the analysis. This signal comes from the accelerometer which measures acceleration in the horizontal plane, and the characteristics found in the signal spectrum were representative of what was observed in the other channels.

The visibility of main signal components is usually measured using the signal to noise ratio (SNR). This concept is widely used in electronics to evaluate the performance of different electronic devices such as amplifiers or radio receivers because it gives a measure of the signal quality. In those applications the signal to noise ratio is calculated as the ratio between the power of the signal and the power of the background noise. Another definition of SNR is the ratio between the average amplitude of the main signal components μ, and the standard deviation of the background noise σ, which is equivalent to the reciprocal of the coefficient of variation [122]. This alternative definition is used in those applications where it is difficult to differentiate between the main signal and the background noise such as image processing and this definition was employed in this investigation. The average amplitude of the main signal components was calculated in each case as the average amplitude of the visible peaks associated with the characteristic defect frequencies (Table 26), while the rest of components with significant lower amplitude were considered as background noise. In order to estimate the average amplitude of the main signal components, the amplitude attached to each characteristic frequencies of

the rig (Table 26) and its harmonics, was calculated for each spectrum. Obviously, because not all the possible defects were present at all times, it was necessary to determine whether there is a visible peak at each of those defect frequencies or not for each measurement. The assumption made was to consider main signal components only those peaks whose amplitude in the spectrum is at least 3 times the average amplitude across the whole frequency range. This average was calculated excluding the amplitudes related with the defect frequencies. Using this procedure it was possible to separate the main peaks in the spectrum attached to the known defect frequencies and the rest of the components in the spectrum, considered as background noise. According to this definition, in each case studied the improvement in the signal to noise ratio was measured as a percentage comparing the SNR of the manipulated signal against the SNR of the raw signal.

6.4 Results

Once the experiment was carried out, the data acquired was processed using the methodologies mentioned in section 6.2. The results obtained for each measurement are plotted in this section with the following format:

a) Amplitude spectrum of the original signal
b) Amplitude spectrum of the signal obtained by LP
c) Amplitude spectrum of the signal obtained by SANC
d) Magnitude of the squared envelope of the signal obtained by filtration in the frequency band of maximum SK

The spectrums of the original signal, and the signals obtained by LP and SANC are represented twice. The left plot corresponds to the spectrum covering a frequency range of 0-2500Hz which contains the gear mesh components and its harmonics. The right plot covers the region of 0-500Hz, where it is easier to identify the typical defect frequencies. The available frequency range of the squared envelope of the signal obtained by filtration after the Kurtosis analysis depends on the filter parameters, different for each analysis. The Kurtograms of the different observations and the main information extracted taken can be seen in 6.4.4

6.4.1 First Observation (19/08/2010)
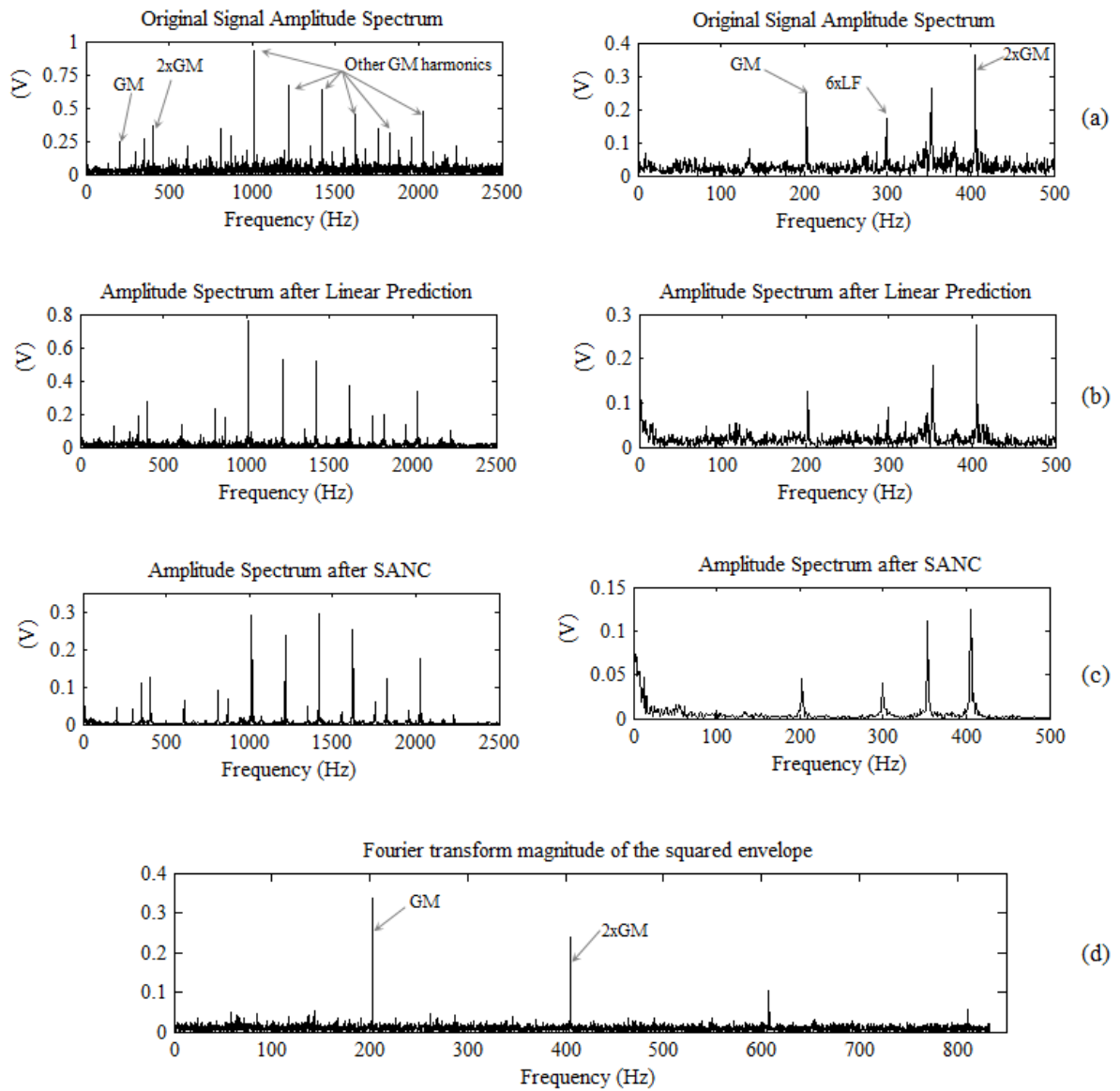
Fig. 92: Results obtained from the first observation (19/08/10)

For this observation, the LP analysis (Fig. 92 b) was performed using 200 past samples for each prediction, and the parameters selected for the self-adaptive filter (Fig. 92 c) were: delay Δ=100 samples, filter order H=1000 and forgetting factor μ=0.00001. The maximum Kurtosis found was 2.4, at a frequency band centred in 2083.33Hz and a bandwidth of 833.3Hz.

6.4.2 Second Observation (22/08/2010)
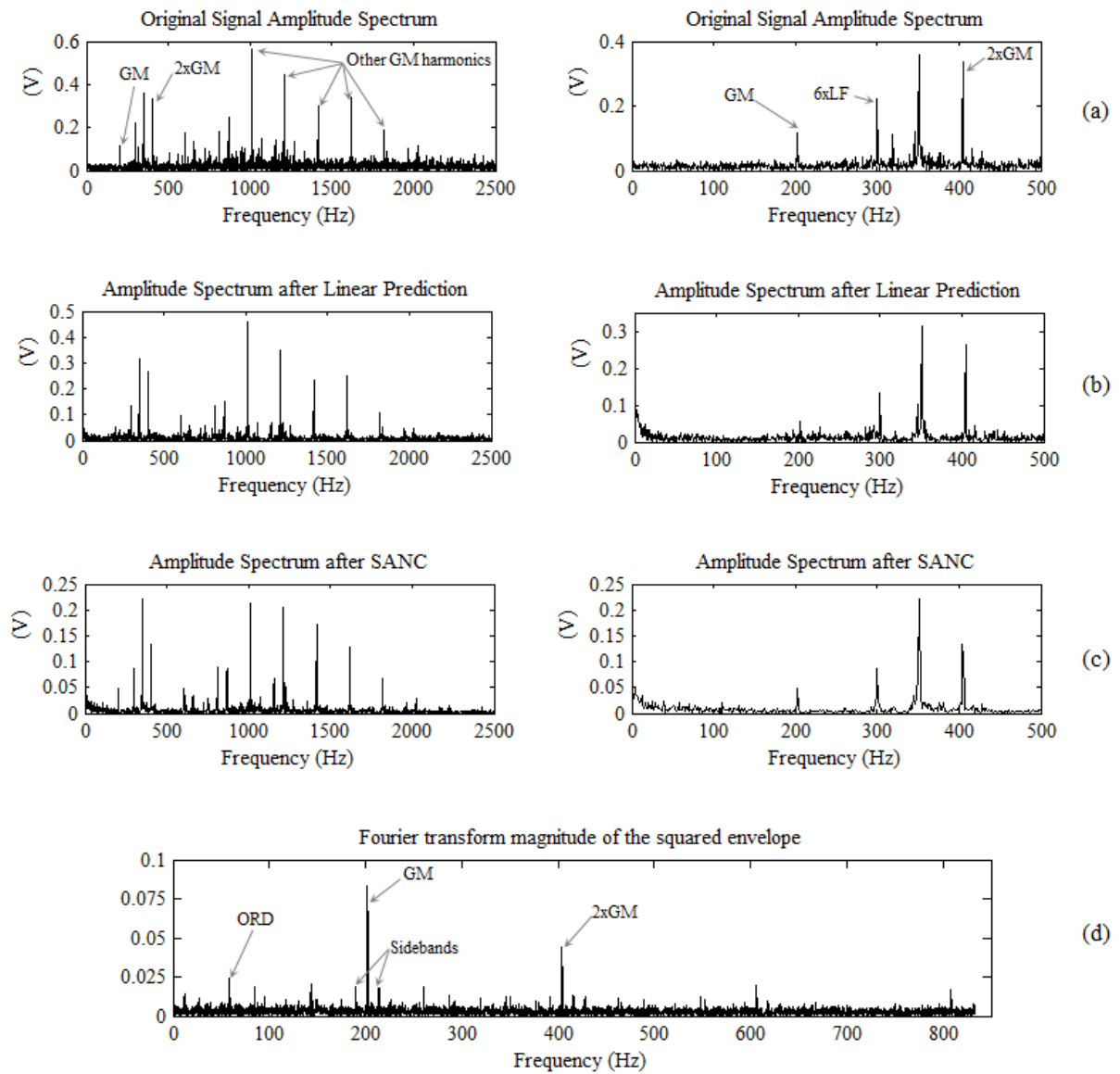
Fig. 93: Results obtained from the second observation (22/08/10)

For the second observation, the LP analysis (Fig. 93 b) was performed using 200 past samples for each prediction, and the parameters selected for the self-adaptive filter (Fig. 93 c) were: delay Δ=500 samples, filter order H=1000 and forgetting factor μ=0.00005. The maximum Kurtosis found was 2.4, at a frequency band centred in 2083.33Hz and a bandwidth of 833.3Hz.

6.4.3 Third Observation (24/08/2010)
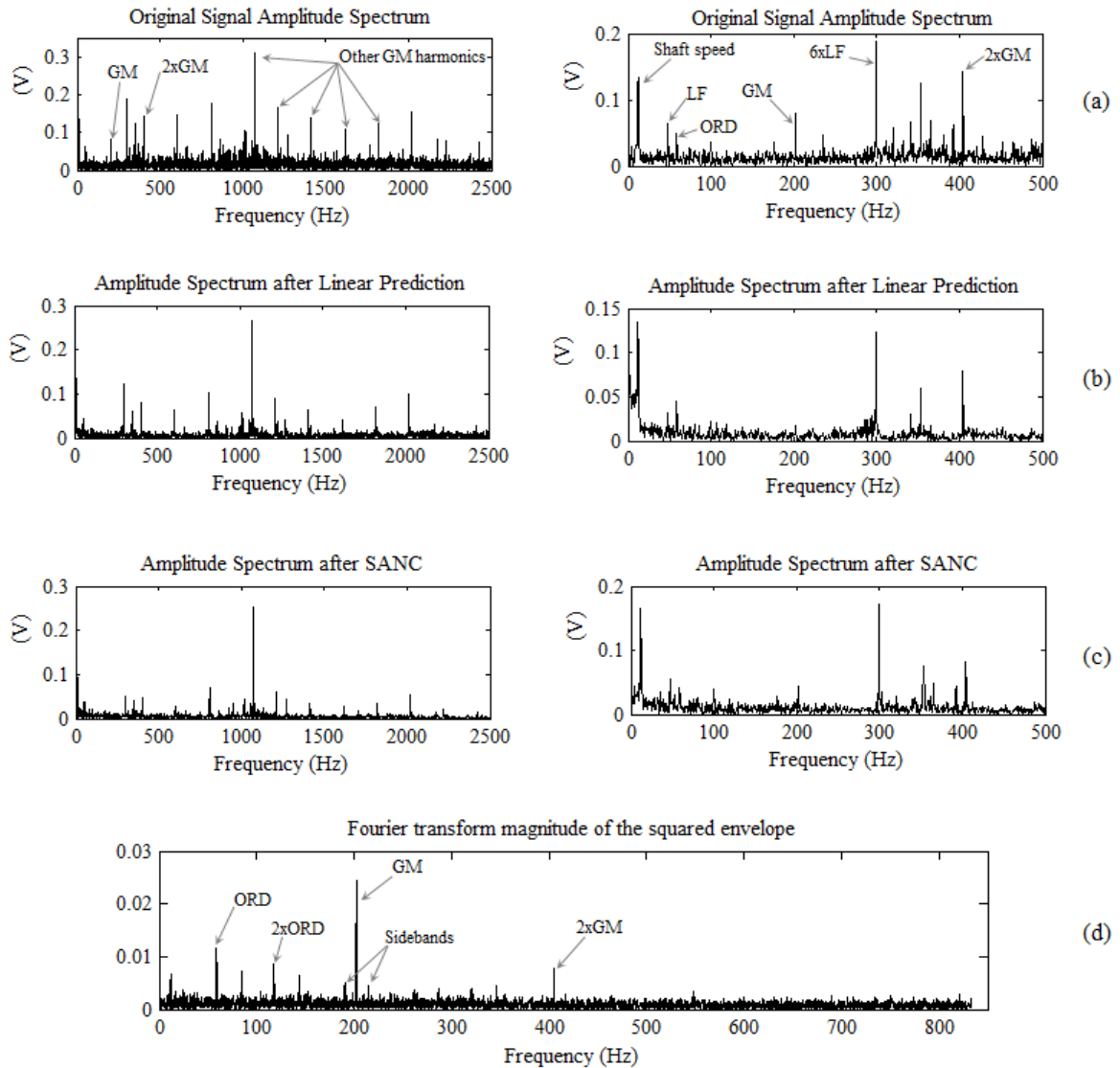
Fig. 94: Results obtained from the third observation (24/08/10)

For the third observation, the LP analysis (Fig. 94 b) was performed using 200 past samples for each prediction, and the parameters selected for the self-adaptive filter (Fig. 94 c) were: delay Δ=500 samples, filter order H=1000 and forgetting factor μ=0.0001. The maximum Kurtosis found was 1.7, at a frequency band centred in 2083.33Hz and a bandwidth of 833.3Hz.

6.4.4 Summary of results

First Observation

Second Observation
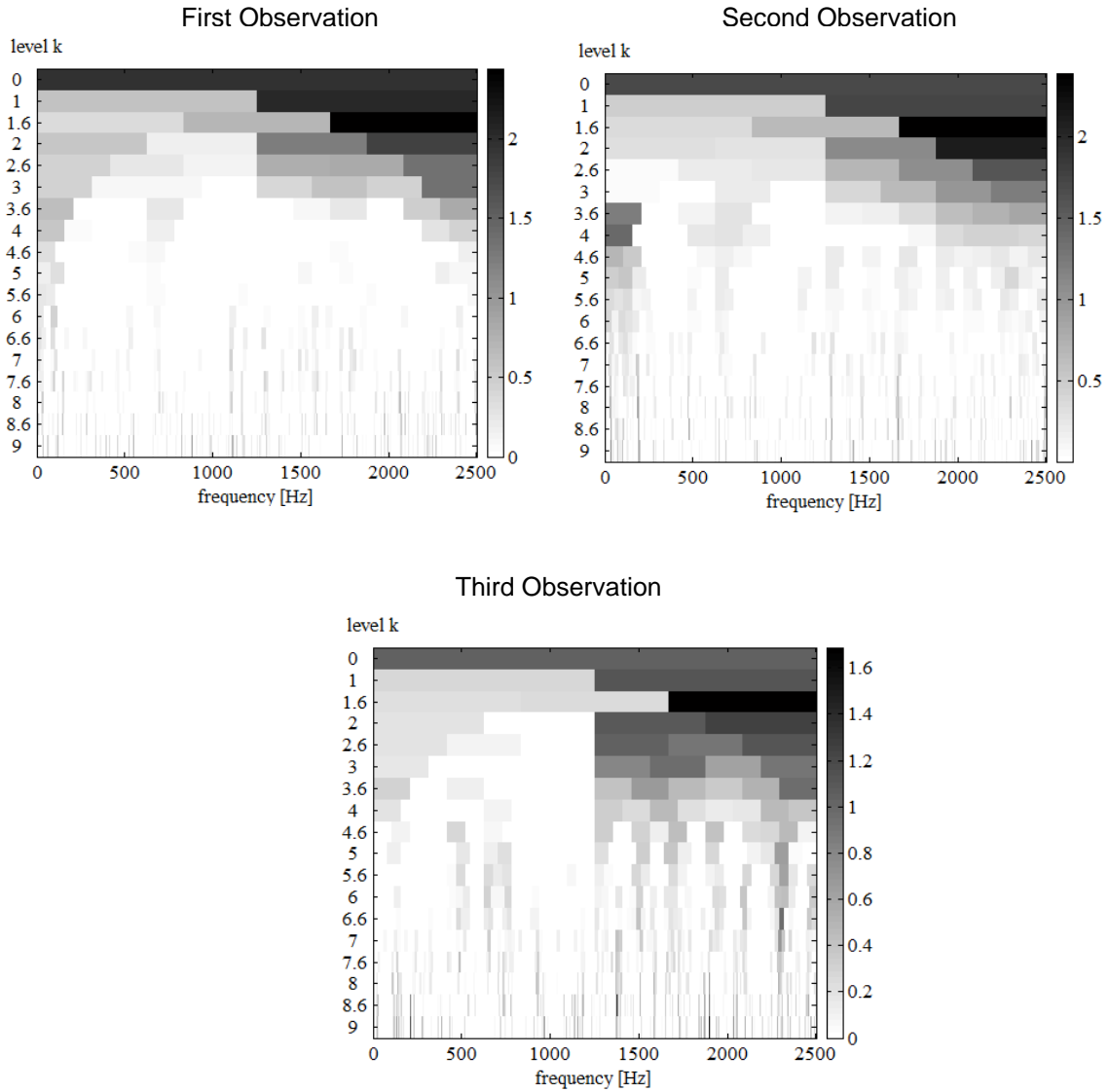
Third Observation

Fig. 95: Kurtograms of the different observations

Table 28: Maximum Kurtosis location

Observation	F_c(Hz)	Δf(Hz)	K_{max}	Frequency Band(Hz)
1	2083.33	833.3	2.4	1666.7-2500
2	2083.33	833.3	2.4	1666.7-2500
3	2083.33	833.3	1.7	1666.7-2500

177

Table 29: Summary of signal-to-noise ratio results

Observation	Method	μ	σ	SNR
1	Original	0.2001	0.0190	10.5315
	LP	0.1004	0.0093	10.8054
	SANC	0.0612	0.0041	15.0496
	SK	0.0570	0.0054	10.6368
2	Original	0.1304	0.0120	10.8667
	LP	0.0701	0.0056	12.4314
	SANC	0.0548	0.0043	12.6596
	SK	0.0158	0.0014	11.1926
3	Original	0.0886	0.0087	10.1839
	LP	0.0487	0.0040	12.1750
	SANC	0.0353	0.0023	15.4082
	SK	0.0081	0.0008	9.6747

6.5 Results discussion

The first measurement (19/08/10, Fig. 92) was acquired one month after the start of the experiment, which corresponds approximately to a 24% of the expected bearing life. The spectrum of the original signal is clearly dominated by the gear mesh frequency (~202Hz) and its harmonics. However, looking closely to the lower frequencies (right column) it is possible to see a peak around 352Hz which is close to the 6th harmonic of the outer race defect frequency (ORDF), but any defect in the outer race at this point was ruled out by visual inspection (see Fig. 91). The presence of this peak is attributed to a natural frequency of the structure or a consequence of deformation due to the three point clamping during grinding the outer ring and it will be present in all the stages of the experiment. There is also a component at 300Hz which corresponds to the 6th harmonic of the 50Hz line frequency (LF). This was corroborated by the fact that this peak appears even for those analysis carried out using data from the load cycle region where the motor speed and load was 0, indicating that this is a parasite component coming from the electrical grid.

From analysis and observations of Fig. 92, background noise was reduced by LP and especially by SANC, increasing clearly the signal to noise ratio compared with the original signal (2.6% and 42.9% respectively, see Table 29) and facilitating the identification of the different signal components. The amplitude of the main peaks in the frequency spectrum was also reduced in magnitude, but this is not significant in terms of component identification, because the signal contains the same main components at a better signal to noise ratio. No new peaks masked by the background noise were identified. The envelope obtained after the signal filtration at the maximum Kurtosis frequency band shows clearly that the signal is dominated by the gear mesh frequency, but at this moment it does not provide any information of an incipient fault in the system.

The original spectrum of the second measurement taken 3 days after the first one (see Fig. 93) shows more or less the same components noted in the first observation, with the difference being that there is a reduction in the amplitude of the peaks and the background noise is slightly lower. No new signal components are identified by LP or SANC, despite of the fact that the background noise reduction is considerable, with an improvement of the 14.4% and 16.5% respectively in the signal to noise ratio in comparison with the original signal (see Table 29). The most interesting result of this analysis is the signal envelope obtained after the filtration at the maximum Kurtosis band: apart of the typical gear mesh frequency and harmonics, it is possible to identify a new peak at the frequency of 58.4Hz, indicating an incipient fault in the outer race of the bearing, in addition to sidebands around the gear mesh frequency at 190.2Hz and 214Hz. The distance between them and the gear mesh frequency is approximately 12Hz, the shaft speed. It is important to emphasize in the fact that the chart d in Fig. 92, Fig. 93 and Fig. 94 represents the spectrum of the squared envelope of the filtered signal, not the spectrum of the filtered signal itself.

On 24/08/10 the last data capture was performed (see Fig. 94). The first observation to note is that the amplitude of the different components is lower in

179

this case. This is due to the fact that this measurement was done during a loading cycle type where the maximum transmitted torque was lower (42.8 Nm) than in the previous measurements (125 Nm). Even under this low torque conditions and despite the reduction in amplitude, all previously noted peaks where evident in the original spectrum, in addition to a clear peak at 58.8HZ, indicating the defect in the outer race of the bearing. Moreover, several sidebands around the harmonics of the gear mesh frequency, separated by the shaft frequency were noted, showing that probably the bearing failure was causing a shaft misalignment which was affecting to the gear mesh. The spectrum of the squared envelope showed the peak at 58.8 Hz and a second harmonic of it at 117.9Hz, indicating the fault in the outer race that was confirmed after by visual inspection of the component (Fig. 96).
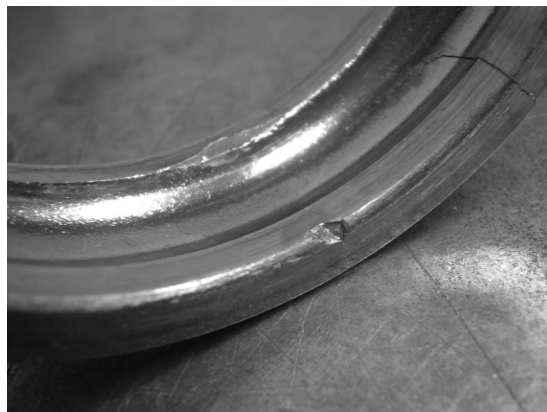
Fig. 96: Bearing outer race degradation after 36 days

6.6 Conclusion

This investigation shows the results of the application of three different vibration based analysis methodologies for bearing diagnosis: LP, SANC and SK together with envelope analysis. These techniques are typically used for applications where strong background noise masks the mechanical signature of a machine making the identification of the fault source challenging. This is the case of the experiment presented in this investigation. LP and specially the SANC showed its capability to reduce the background noise and facilitate the identification of the different components in the signal spectrum, but in this

specific application they did not identify the defect on a bearing earlier than the SK.

The latter technique demonstrated the ability to identify the defect earlier than all other methods. This method is thus a very powerful tool for early detection of faults in bearings, even for those applications where strong background noise from other sources in the machine masks the characteristic fault components in the frequency domain.

7 USE OF SPECTRAL KURTOSIS FOR IMPROVING SIGNAL TO NOISE RATIO OF ACOUSTIC EMISSION SIGNAL FROM DEFECTIVE BEARINGS

Abstract

The use of Acoustic Emission (AE) to monitor the condition of roller bearings in rotating machinery is growing in popularity. This investigation is centred on the application of Spectral Kurtosis (SK) as a denoising tool able to enhance the bearing fault features from an AE signal. This methodology was applied to AE signals acquired from an experimental investigation where different size defects were seeded on a roller bearing. The results suggest that the signal to noise ratio can be significantly improved using SK.

7.1 Introduction

As it was mentioned in chapter 0, vibration based monitoring is one of the most popular methods for condition monitoring and there are plenty of signal processing techniques available to enhance fault features in the acquired signal. Despite of the success of vibration-based methodologies, over the last decades bearing condition monitoring techniques based on AE have become very popular [123-142]. AE has demonstrated to be a very powerful tool for fault detection and diagnosis particularly in bearings, and some recent studies [124; 143; 144] reported that AE can be more sensitive in detecting incipient faults in bearings over other methodologies based on vibration data. Matthews J.R. [145] defined AE as transient elastic waves generated from a rapid release of energy, caused by a deformation or damage within or on the surface of a material. In the particular case of bearings, AE is generated by the interaction of two surfaces that are in relative movement to each other. The interaction of the rolling elements with a defective surface will generate an AE burst superimposed onto continuous AE background noise, as seen in Fig. 97

Theoretically, it is possible to identify the defect source just by measuring the interval of time between each burst. This interval is related with a particular defect frequency that can be easily calculated from the bearing geometry,

indicating the origin of the bursts. This technique has been successfully applied by different researchers [123; 144] to identify outer race defects, but there have been reported difficulties in identifying inner race defects using AE [123; 131; 146].

In addition, Al-Dossary, S. et al. [123] reported that there is a strong correlation between the AE burst duration (see Fig. 97) and the actual size of the defect; in an investigation where defects of different size were seeded on a roller bearing. In that experiment, the measured burst duration was close to the theoretical time duration of the roller passing over the defect, irrespective of the speed and load condition. This fact demonstrated that AE can be used not only to detect and diagnose bearing faults, but also to obtain information about the defect severity. The main problem faced by the authors was the relatively low signal to noise ratio found in the signal, especially in those tests where the defect size was small. This issue made it difficult to identify the instant when an AE burst overcame background noise level; essentially the point where the burst goes over background noise level indicating its start or end.

Eftekharnejad et al. [144] was the first to report that it is also possible to effectively increase the signal to noise ratio and enhance the bearing fault features in an AE signal using SK and the Kurtogram. The main aim of this study is to complement the work of Al-Dossary et al. [123] by post-processing the AE signal acquired during the tests in order to increase the signal to noise ratio, reinforcing the novel idea of applying SK in AE signals as Eftekharnejad et al. [144] postulated recently.
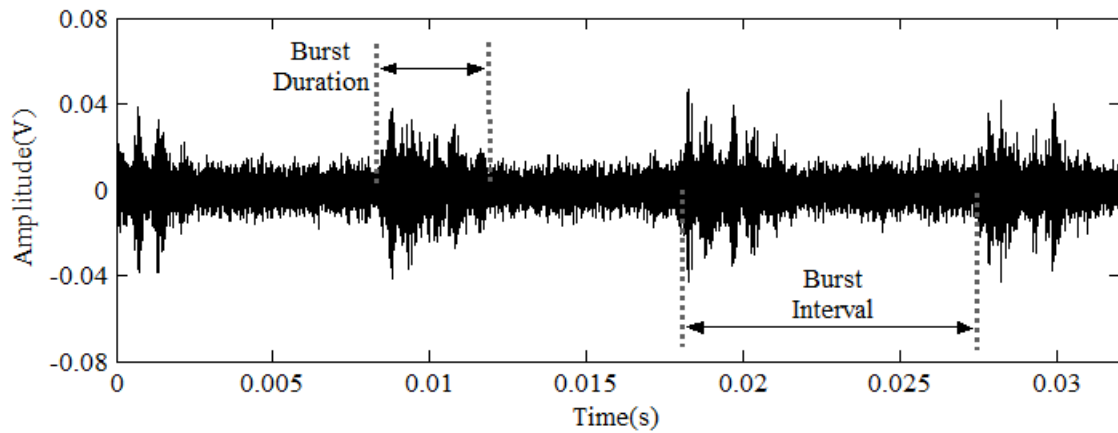
Fig. 97: Typical AE bursts associated to an outer race defect

7.2 Experimental methodology

7.2.1 Test rig set-up

The data used for this investigation is the same used by Al-Dossary et al. in [123]. In the test-rig, the test bearing is mounted on a shaft driven by an electric motor. The shaft is supported by two large slave bearings and the radial load is applied to the test bearing using a hydraulic cylinder (see Fig. 98).
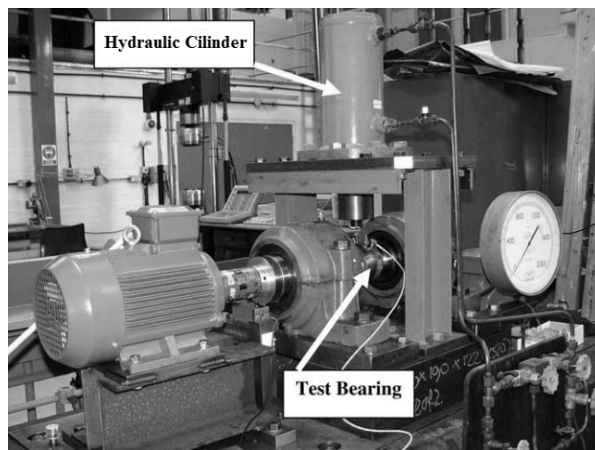
Fig. 98: Layout of experimental test-rig

Table 30: Main bearing dimensions

No. of rolling elements	10
Roller Diameter	12 mm
External Diameter	84 mm
Pitch Diameter	68 mm

185

The test bearing chosen for the experiment was a Cooper cylindrical roller type 01B40MEX, and its main dimensions are summarized in Table 30. The selected sensor was a piezoelectric "Physical Acoustic Corporation type WD" with an operating frequency rate of 100-1000kHz, which was mounted on the top half of the bearing housing. The acquired signal was sampled at 8MHz for a rotating speed of 1500 rpm, and then amplified at 40 dB. Nearly 250,000 points where acquired for each measurement, which corresponds approximately to a signal of 30ms.

In the original experiment [123], 4 different bearings were tested under different load and speed conditions. Defects of different sizes were seeded in such a manner as to understand how the defect size influences the AE waveform. For that purpose, different bearings with seeded faults were tested at 300, 1500 and 3000 rpm under loads of 2.7kN, 5.3kN and 8kN. For this study, only the AE signal acquired from one of those test bearings is analysed, where the load applied to the bearing was 5.3 kN and the rotational speed of the motor was 1500 rpm. The defects on the outer race were made using an electric engraver with a carbide tip. A total of 9 different defect sizes were employed: see Table 31 where L represents the length measured in the circumferential direction and W is the width across the bearing race. After each test, the bearing was disassembled and the next seeded bearing assembled. The rig was run for 15 minutes prior to the data acquisition in order to bring it up to thermal equilibrium.

Table 31: Incremental defect sizes (outer race)

Defect	Size (L x W)
D1	Circle D=0.5mm
D2	0.9 x 2.5
D3	0.9 x 4
D4	0.9 x 8
D5	0.9 x 12
D6	3 x 12
D7	5 x 12
D8	7 x 12
D9	9 x 12

7.2.2 Methodology

As mentioned in the introduction section, the aim of this study is to improve the signal to noise ratio of AE signals in order to enhance the bearing fault features in the acquired signal. The main benefit of using SK is that the frequency band where the signal is band-pass filtered to enhance the bearing defect signature is selected automatically depending on the original signal characteristics. The Kurtogram is used in the first instance to locate the frequency region where the Kurtosis is maximum, indicating a higher level of impulsiveness. For this investigation, the algorithm proposed by Antoni, J. [105; 121] for the fast computation of the Kurtogram was used, and the Kurtograms obtained for the different tests performed are shown in Annex C.

The information provided by the Kurtogram is used to design a band-pass filter for the purpose of enhancing the impulsiveness of the original signal. In this work FIR filters were employed to filter the signal, and the filters were designed using Matlab Filter Design and Analysis Tool. The characteristics of the filters (according to Fig. 99) were: A_{pass}= 1dB, A_{stop1}= 60dB, A_{stop2}=80dB, and finally F_{stop1} and F_{stop2} were always selected 2kHz below and above the frequency band limits. The information provided by the Kurtogram for the different signals analysed (central frequency F_c, frequency resolution Δf and maximum Kurtosis (K_{max}) is summarized in Table 32. K_{max} represents the maximum Kurtosis value found in all the frequency regions inspected for all the levels analysed.
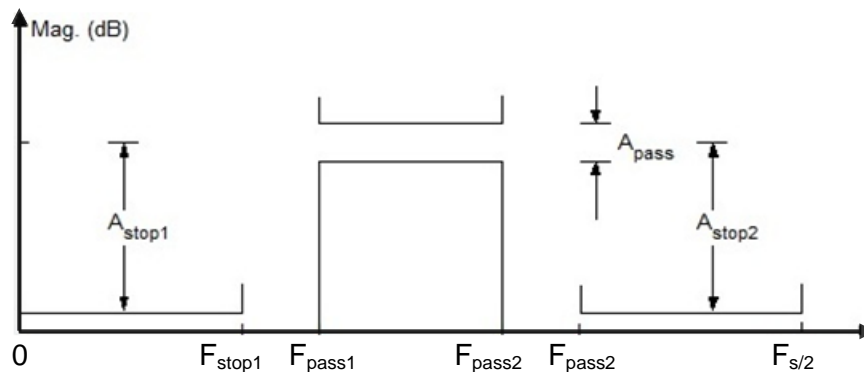
Fig. 99: FIR filter characteristics

Table 32: Information extracted from Kurtogram and filter design toolbox

Defect	F_c(Hz)	Δf(Hz)	K_{max}	Frequency Band(Hz)
D1	1,000,000	2,000,000	1.2	0 – 2,000,000
D2	531,250	20,833.33	5.4	521,000 – 542,000
D3	46,875	31,250	16.9	31,250 – 62,500
D4	52,083.33	20,833.33	34.2	41,660 – 62,500
D5	70,312.5	15,625	18.2	62,500 – 78,120
D6	78,125	31,250	17.7	62,500 – 93,750
D7	31,250	62,500	45	0 – 62,500
D8	54,687.5	15,625	5	46,880 – 62,500
D9	72,916.67	20,833.33	6.8	62,500 – 8,330

In addition, the effectiveness of the denoising effect of SK on AE signals has been quantified by comparing the SNR of the original and filtered signals. The signal to noise ratio has been estimated using the Crest Factor (CF), which is defined as the ratio between the maximum absolute value and the signal r.m.s., giving an indication of the peak-to-average ratio [144]. The CF has already been successfully used as an indicator of damage severity in other investigations about AE applied to bearing fault detection [123; 129; 144; 147].

7.3 Results and discussions

7.3.1 Time domain

Fig. 100 shows the original signals acquired for the 9 defects tested in the left column and the corresponding filtered signals in the right column (the Y axis represents the amplitude in Volts). Each waveform represents 30 ms of acquired signal, which corresponds approximately to ¾ of one revolution. The outer race defect frequency is 4.1 times the shaft speed; therefore 3 or 4 burst spaced at 9.75 ms are expected to be visible in the captured time waveform. The Kurtograms for the different signals are attached in Annex C, and details about the information provided by SK can be found in Table 32.

Fig. 100: Signal amplitude (Volts) of the original (left) and filtered signals (right)
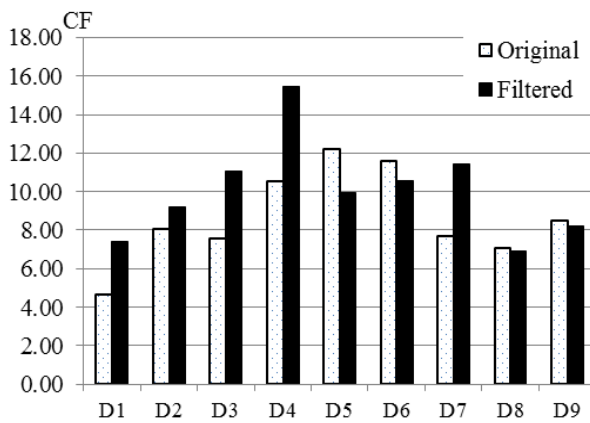
Fig. 101: Crest factor for the original and filtered signals

Table 33: CF for the original and filtered signals

	Original	Filtered	Improvement %
D1	4.65	7.43	59.71
D2	8.01	9.22	15.07
D3	7.53	11.07	47.13
D4	10.50	15.45	47.07
D5	12.18	9.95	-18.29
D6	11.54	10.54	-8.66
D7	7.67	11.46	49.30
D8	7.01	6.90	-1.66
D9	8.48	8.20	-3.31

189

Intuitively, it is possible to see in Fig. 100 that level of background noise in the filtered signals has been reduced when compared with the original signals. The same bursts are visible in the filtered signal though with lower amplitude due to the different voltage ranges. In the filtered signals the signal to noise ratio is higher, especially in the smaller defects tested, showing the denoising capability of SK. This fact can be observed for example in D2 (Fig. 100), where the bursts happening approximately 0.0025, 0.0125 and 0.0225 seconds after the beginning of the measurement are much more visible over the background noise in the filtered signal than in the original signal. Specifically the SNR in this observation has improved by 15%. The same effect can be observed in tests D3 to D9, where the peaks in the signal generated by the rolling element passing over the defects are more visible over the background noise compared with the original signals. This improvement can facilitate the task of identifying those peaks during the early stages of degradation when the peak intensity is lower, or in applications where the background noise is significantly high. In the smallest defect tested (D1) there are no visible bursts in the original signal and SK was not able to extract the impulsive part of the signal. Consequently it is not possible to identify any bursts in the filtered or the original signal for the first defect.

Fig. 101 shows the crest factor in the original and filtered signals for each defect tested. The same information is detailed in Table 33, including the percentage of improvement achieved in the filtered signal. The improvement in the signal to noise ratio is relatively high in the smaller defects tested (D1 to D4) and in D7, but it is insignificant in the rest of defects and much worse in the case of D5. The average percentage of improvement in the CF excluding D1 is 15.8%. A possible explanation for this phenomenon is the change in the probability density function of the signal as the defect size grows. These higher intensity impulses occur as the defect size increases and can sometimes excite structural natural frequencies, harmonics of other rotating parts, etc. which span a large frequency range and mask the original defective signal [104]. As a result, the Kurtosis value decreases, hindering the fault detection process. In Table 32 it can be seen that the value of the maximum Kurtosis found by the

Kurtogram decreases after D4, exactly at the same point when the SNR starts to decline.

The improvement in the crest factor for the smaller defects is especially interesting, due to the importance of detecting incipient faults during the early stages of degradation. In the original signals acquired for D2, D3 and D4 it is difficult to differentiate between peaks generated as a consequence of the bearing ball rolling over the defect from other signal components produced mainly by background noise. Nevertheless, in the signal processed using SK those peaks are much more visible over the background noise, which allows an easy detection and identification of the defects during the first stages of the degradation process. The bursts observed in D5 to D9 are already clearly visible in the original signals. As a consequence, the slight reduction in the CF for those cases (except D7) does not affect the detectability of the defect, which is still clearly visible in the filtered signals as well. In D7 the maximum Kurtosis found is curiously high compared with the values found in D5, D6, D8 and D9, and for that particular case the SNR was significantly improved in the filtered signal.

7.3.2 Frequency domain

In Fig. 102 are represented the envelopes of the original (left) and filtered (right) signals for the 9 different defects tested. The same spikes remain after the filtration process and again it is evident that the background noise level has been reduced in the filtered signal envelopes. The CF represented in Fig. 103 and detailed in Table 34 shows similar trends as showed in Fig. 101 and Table 33, with a clear improvement of SNR in the smaller defects (D1 to D4) and D7, but worse results were obtained for the defects D5, D6, D8 and D9. No new peaks indicative of burst presence are visible in the envelope of the filtered signal for the defect D1 despite of an increment in the CF of 38.31%. This result indicates that D1 is too small to be detected either in the time or the frequency domain even using SK.
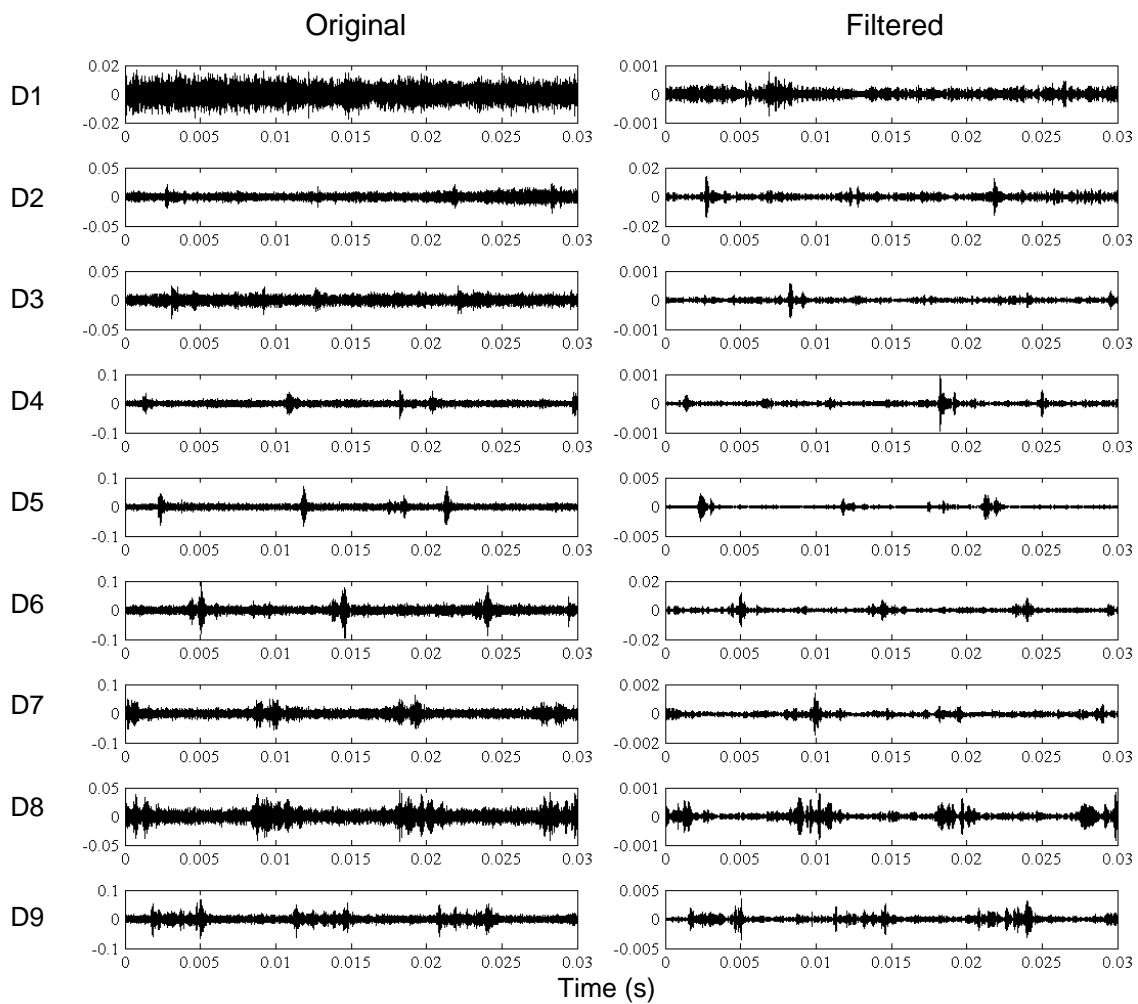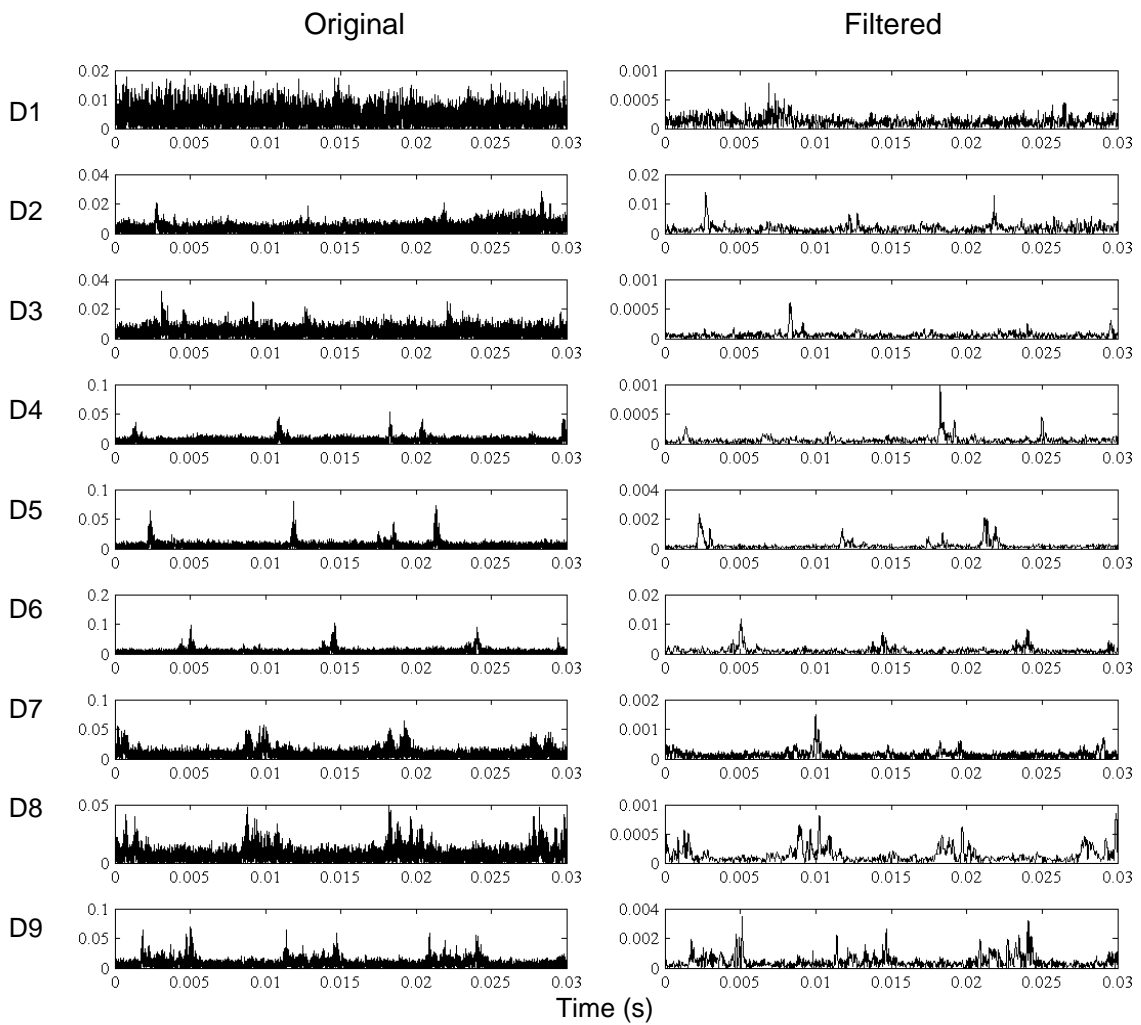
Fig. 102: Squared envelope of the original (left) and filtered signals (right)
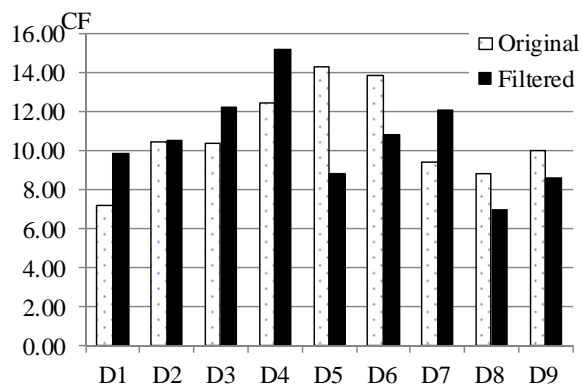
Fig. 103: Crest factor for the original and filtered envelopes

Table 34: CF for the original and filtered envelopes

	Original	Filtered	Improvement %
D1	7.12	9.85	38.31
D2	10.43	10.54	0.99
D3	10.35	12.22	18.07
D4	12.40	15.16	22.28
D5	14.24	8.80	-38.19
D6	13.83	10.78	-22.05
D7	9.36	12.07	28.86
D8	8.81	6.96	-20.95
D9	9.98	8.60	-13.86

192

7.4 Conclusion

The capability of SK to enhance the signature of a bearing fault is well known when applied to vibration signals as shown in chapter 6, but only the work of Eftekharnejad, B. et al. [144] reported a successful application of SK to detect bearing faults using Acoustic Emission. The sensitivity of SK when applied to AE signals has been tested in this work for an artificially damaged bearing.

From the results obtained in this work it is possible to conclude that SK can be a very useful tool to reduce the background noise and improve the burst visibility in bearing AE signals obtained from defective bearings with different defect sizes. The improvement in the signal to noise ratio achieved was dependant on the defect size, obtaining a higher percentage of improvement in the smaller defects, precisely where the AE bursts where less visible in the original signal. This affirmation reinforces the recent findings of Eftekharnejad, B. et al. [144], confirming that SK can be helpful for the early identification of bearing defects using AE, especially during the early stages of degradation. The application of SK to AE signals can thus improve the detection rate of bearing faults and reduce the detection time once the degradation process starts. Nevertheless the results showed that it is not always possible to improve the signal to noise ratio for AE signals associated with an advanced stage of bearing damage.

8 COMBINATION OF PROCESS AND VIBRATION DATA FOR IMPROVED CONDITION MONITORING OF INDUSTRIAL SYSTEMS WORKING UNDER VARIABLE OPERATING CONDITIONS

Abstract

The detection and diagnosis of faults in industrial systems is a very active field of research due to the reduction in maintenance costs achieved by the implementation of improved condition monitoring methods. In particular it is easy to find in literature examples of successful application of data driven methods for process monitoring such as PCA, PLS or more recently CVA in real or experimental cases. However it is difficult to use these methods to detect incipient mechanical faults that can occur in process assets like gear faults, bearing faults, shaft misalignment or rotor unbalance if only process data is included in the analysis. In the early stages of degradation these faults have a minor effect on the process performance, but can cause extensive damage to the machines if no maintenance action is taken. Typically the condition of rotating machinery is monitored separately using vibration analysis or other specific techniques. Conventional vibration-based condition monitoring techniques are based on the tracking of key features observed in the measured signal. Typically steady-state loading conditions are required to ensure consistency between measurements.

In this study, a technique for merging process and vibration data in order to improve the detection of mechanical faults in industrial systems working under variable operating conditions is proposed. This allows the fusion of process and vibration data into a single data matrix for further analysis. The dynamic characteristics of CVA allow the detection and diagnosis of faults in systems working under variable operating conditions. This approach was tested using experimental data acquired from a compressor test rig where different process faults were introduced. In addition, mechanical faults were simulated and seeded in the data to prove the validity of the approach. Results suggest that the combination of process and vibration data can effectively improve the

detectability of mechanical faults in systems working under variable operating conditions.

8.1 Introduction

The detection and diagnosis of faults in industrial systems is a very active field of research due to the reduction in maintenance costs achieved though the implementation of improved condition monitoring methods. A reduction of the number of unplanned shutdowns, improvement of system availability, capacity to pre-order spare parts as needed, increased safety in plant operations and the increase of the process efficiency are some of the main benefits of condition based maintenance. Modern industrial facilities are heavily automated and instrumented; consequently there is a lot of process data available which can be used to monitor the condition of the system. The difficulties attached to the development of accurate and reliable first-principle models of large and complex industrial facilities has led to the success of data driven methods for condition monitoring such as PCA, PLS or CVA [2]. Literature gives examples of extensive application of these methods for detection and diagnosis of faults using computer simulated data [23-30] or real data obtained from industrial facilities or experimental test rigs [63-74]. Despite their success, PCA and PLS (and their corresponding dynamic approaches known as Dynamic PCA and Dynamic PLS [15; 16]) have been reported not to be as efficient as other state-space based methodologies such as CVA, and in chapter 4 it was demonstrated its superior performance using experimental data. The benefits of CVA are especially relevant when applied to systems working under variable loading conditions, principally due to the representation of the system dynamics [17-19].

Despite the success of the aforementioned methods in the detection and diagnosis of process faults, these methodologies can be insensitive to incipient mechanical faults if only process variables are analysed, as typical faults such as misalignment or unbalance have a minor effect on the performance of the machine during the early stages of degradation. However, this kind of fault can have a disastrous effect on the machines, causing catastrophic failures if the

malfunction is not corrected due to the dynamic effects of the additional loads generated by the faulty condition. Vibration-based condition monitoring is probably the most common method for detection and diagnosis of mechanical faults in rotating machinery and it has several advantages against other methods. Typically, analysis of the vibration frequency spectrum can point directly to the source of the fault and there are plenty of signal processing techniques available to help the user to undertake diagnosis in conditions of high background noise as it was shown in chapter 6.

The simplest and most used method to detect the presence of faults with vibration analysis involves a comparison of different signal features (such as RMS value, Kurtosis or peak amplitude) in the measured signal against a machine working under healthy conditions. Assuming that the initial status of the machine was healthy, any changes in the measured feature response are caused by the deterioration of the machine condition. However, this assumption is only valid if all the measurements are taken under the same loading conditions, as different levels of load will generate different vibration levels [53]. It is possible to find in literature some examples of techniques used to monitor the condition of machines working under variable loading conditions using vibration data. McFadden [59; 60] proposed a method based on band pass filtered time-domain synchronous averaging (TSA) and Hilbert transform where Kurtosis was used as an indicator of fault severity. However there are some disadvantages in this technique due to the user involvement in the election of the bandwidth for the band-pass filter [148]. Other methods are based on the examination of time-frequency maps where the instantaneous power spectrum is represented [149], but this method does not produce a single indicator of the machine condition that can be tracked in time. Parker Jr. et al. [58] proposed a method based on change detection in the bispectral domain which produced severity indicators independent of the loading conditions, but requires long computational times. The work presented by Zhan et al. [57] proposes a technique based on motion residuals, which are calculated as the difference between the TSA of a signal and the average vibration observed in the healthy state under different loading conditions. This area has gained importance in the

last years and Braun [56] reviewed the state of the art of vibration diagnostics using TSA in 2011. Other methodologies presented recently are based on capturing the correlation between features extracted from the vibration signal and the operating conditions. This kind of approach has been applied successfully for diagnostics of planetary gearboxes in a bucket wheel excavator [55] and wind turbine bearing diagnostics [54].

There are several examples in industry where machines are working under severe changes in loading conditions, such as the mining industry or wind turbines. Industrial needs are evolving towards more flexible production schemes in order to promote efficiency and maintain their competitiveness in a market where the demand, the price of raw materials and even the price of the energy can be very volatile. That is why it is important to develop condition monitoring tools that can detect and diagnose faults in industrial systems working under variable operating conditions. These improved methods should be able to deal with both process and mechanical faults in order to ensure the quality of the product and the safe and economical operation of the plant.

In this study it is proposed a method to combine process data and key features obtained from vibration measurements in order to provide a more robust and reliable condition monitoring tool. This method takes advantage of the relation between the load and the energy of the measured vibration signal and the capacity of CVA to capture correlations between the different variables measured. In order to overcome the challenge of the different sampling rates in process and vibration data, the vibration signal is split into sections with the same duration as the time between measurements of consecutive process variables. It is assumed that if the sampling rate of process variables is high enough compared with the changes in operational conditions, the changes in the vibration signal during each section analysed will be negligible, and consequently no order tracking or resampling is needed. Then, the desired features are extracted from each of the vibration signal sections and treated as new variables that can be combined in a data matrix together with process variables. This technique can be used not only for tracking feature changes in

vibration data but in any kind of signal with an oscillatory nature such as AC current or sound.

In order to prove the validity of the method, it was tested using experimental data acquired from a compressor test rig. This rig is highly instrumented and specially designed to study the behaviour of a centrifugal compressor driven by an electric motor under normal and abnormal conditions. The operating point of the machine can be modified by manipulating the rotational speed of the motor and the valves situated in the compressor inlet and outlet lines. Different process faults such as compressor surge or pipe blockage were seeded while the system was working under different operating conditions. The associated data sets acquired were used to study the capabilities of the proposed method in terms of fault detection and diagnosis. Additionally, different mechanical faults were simulated and seeded in the experimental data for this investigation. An additional study of compressor surge was carried out in order to demonstrate that it is possible to improve the detection capability by adding vibration data into a multivariate condition monitoring algorithm using acquired vibration data (not computer simulated data). The results suggest that it is possible to improve the performance of the CVA method in a real system by adding vibration data in the analysis.

8.2 Methodology

8.2.1 Combination of process and vibration data for CVA application

As mentioned in section 8.1, most vibration-based monitoring techniques involve the observation of changes in characteristic features of the vibration signal which can point to mechanical faults. Each mechanical fault has a determined "signature" in the vibration signal that can be observed in the time or frequency domain, such as changes in the RMS or Kurtosis values, presence of peaks at particular frequencies in the signal spectrum, appearance of new peaks or sidebands, etc. These changes are used to detect the fault, locate its source, identify the type of fault and measure its severity.

The evolution of these changes in the vibration signal can be tracked in time, using either continuous or intermittent observations. An example of this procedure can be seen in [150], where the historic trend of vibration amplitude and phase at 1X and 2X in the frequency spectrum is monitored over time to observe the development of a flexible coupling failure.

The proposed approach consists in the extraction of the desired features from the vibration signal (or other type of signal such as current, pressure…) at the same rate as the rest of process variables are sampled. Normally vibration signals, as well as alternating current measurements, require a much higher sampling rate than conventional process measurements such as pressure, flow rate or temperature, due to the fast dynamics of vibration and current measurements. The concept represented in Fig. 104 consists of splitting the vibration signal into segments, obtaining one segment for each process measurement acquired. The length L of these segments can vary depending on the requirements (time window represented, resolution of the frequency spectrum…) and can overlap each other depending on the selected length. For every process measurement acquired at time t_i, the L samples acquired in the vibration signal immediately before t_i are extracted to configure the corresponding signal section, obtaining one section for each process measurement. Then, for each one of the sections it is possible to extract the desired features in the time domain (RMS, Kurtosis…) or in the frequency domain after windowing the corresponding section and performing a short Fourier transform (SFFT).
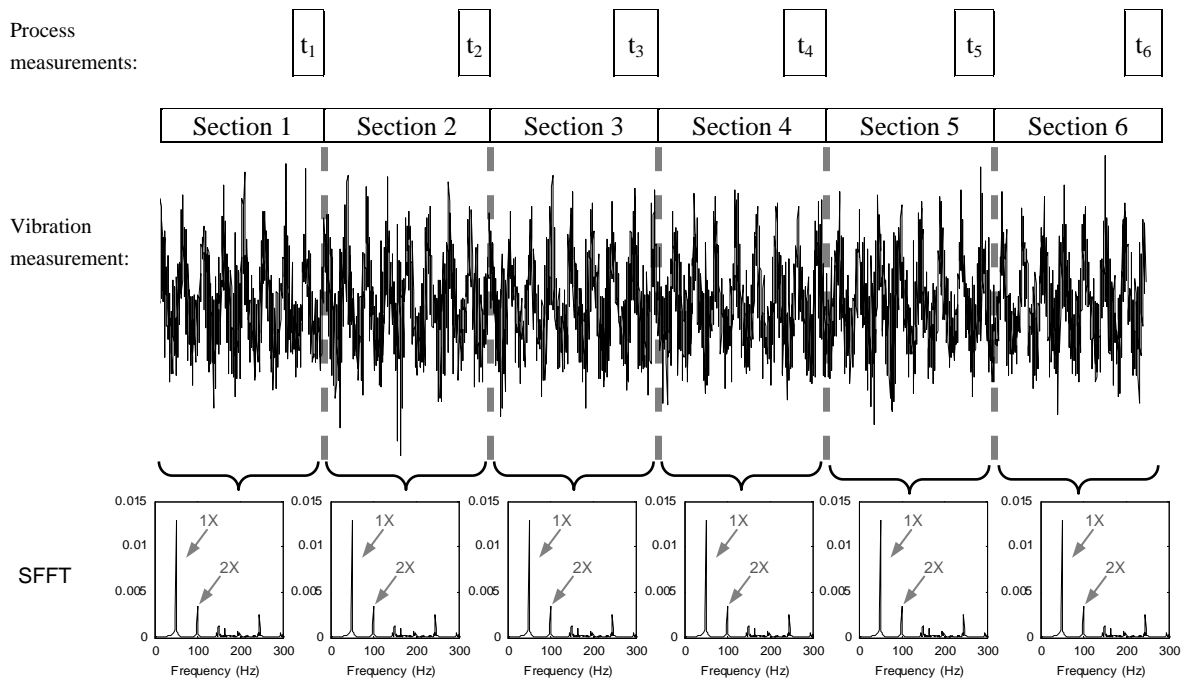
Fig. 104: Vibration signal feature extraction

Once the desired features have been extracted from the vibration signal, the values obtained can be combined with the process measurements in a single data matrix (see Fig. 105), which can be used to perform a CVA analysis using the methodology explained in section 4.2.1

	Process measurements				Vibration measurements		
time	Outlet pressure (bar)	Temperature (°C)	Motor speed (rpm)	Flow rate (m^3/h)	Vibration RMS (g)	1X Vibration amplitude (g)	2X Vibration amplitude (g)
t_1	1.203	12.61	1999.65	501.6	0.062	0.0131	0.0046
t_2	1.189	12.60	1999.72	502.2	0.074	0.0136	0.0041
t_3	1.926	12.63	1999.70	502.8	0.081	0.0142	0.0049
…	…	…	…	…	…	…	…
t_n	2.468	11.56	3999.12	628.3	0.126	0.0216	0.0071

Fig. 105: Example of combined data matrix containing process and vibration measurements

The use of multivariate algorithms to monitor characteristic features of vibration measurements is not new. As an example, Ahmed et al. [151] used PCA to successfully detect and diagnose faults in a reciprocal compressor extracting parameters like peak factor, RMS, entropy, crest factor, etc. from the vibration signals acquired. The methodology proposed here takes advantage of the correlation between process and vibration measurements and the capacity of

CVA to capture this correlation even under varying operating conditions to provide a more robust monitoring tool able to detect and diagnose both process and mechanical faults. After the extraction of the desired features from the vibration signal the rest of the signal containing non-relevant information is dismissed, minimizing the amount of space required to record the data which is one of the main drawbacks of permanent vibration monitoring systems.

8.2.2 Simulation of mechanical faults in vibration data

In order to assess the performance of the proposed methodology it is necessary to obtain process and vibration data which has been acquired simultaneously from a system working under changing operational conditions in the presence of faults. The data acquisition system installed on the rig described in section 8.2.3 was ideal for that purpose, and several data sets were acquired from the system working under different induced faults. However, for safety reasons it was not possible to introduce mechanical faults in the rig whilst ensuring the mechanical integrity of all its components. That is why for this analysis the vibration signature of different mechanical faults was simulated and then seeded in the acquired data. In research it is normal practice to use simulated vibration data for the study of different mechanical faults, being rotor unbalance, shaft misalignment and bearing fault the most common examples [152-160].

Rotor unbalance is caused by the displacement of the rotor centre of mass away from its rotation centre. The centrifugal force generated $\Delta F(t)$ has an amplitude proportional to the rotor mass m_r, the eccentricity e and the square of the rotational speed ω, and has a phase angle δ:

$$\Delta F(t) = \omega^2 m_r e \cdot sin(\omega t + \delta) \qquad \textbf{(8-1)}$$

This relation was used by Sekhar et al. [153; 160] in order to model the effect of the additional forces generated by rotor unbalance in the vibration of a system, which is basically an increase in the amplitude of the 1X peak in the frequency spectrum. The same procedure was used here to reproduce the effects of rotor unbalance.

Vibrational response due to misalignment in shafts connected by flexible couplings has been studied in several investigations [153-159]. All of them agree that the most common effect of misalignment in the vibration signal is an increase in the amplitude of the peak at 2X due to the change in the assembly stiffness twice per revolution, although 1X and other harmonics can also be affected depending on the running speed and the type of coupling. An example of the forces generated by a 1.5° angular or 1.5mm parallel misalignment are presented in [154], where the oscillating part of the force signal has a typical peak to peak amplitude of around 100N. The evolution of the 2X peak amplitude for an experimental case of increasing misalignment can be seen in [157]. In this last reference it is stated that the load has a great influence on the vibration levels and consequently the measurements had to be taken under the same loading conditions. This means that a method that can take into account changes in operating conditions like CVA can be a great improvement. The forces generated by the misalignment are a function of the coupling stiffness [153; 156] which is unknown for the coupling installed in the rig used in this investigation.

In [152] bearing faults are simulated in a simplified manner as impacts periodically repeating at the characteristic defect frequency. An improved simulation method is proposed where the duration of the impact corresponds to the time that it takes for the rolling element to pass over the defect. The peak amplitude of the acceleration response generated is in the order of $1m/s^2$. The same approach was used in this investigation to simulate a bearing fault in the drive-end bearing of the electric motor.

Once the forces generated by the fault are known its effect on the system can be represented as a residual load $\Delta F(t)$ which acts on the undamaged system adding this new force to the forces already existing [160]. Consequently, the motion observed in the damaged system is a combination of the motion caused by the excitation forces in the undamaged system and the motion caused by the virtual damage forces. The problem can be represented in a simplified manner

as a one degree of freedom system where a mass m is connected to the foundation by an elastic element with stiffness K and a viscous damper with damping D which during normal operation is subjected to a force $F_0(t)$. The motion $u_0(t)$ of the undamaged system can be obtained solving the equation:

$$m \cdot \ddot{u}_0(t) + D \cdot \dot{u}_0(t) + K \cdot u_0(t) = F_0(t)$$

(8-2)

When a fault affects the system the additional force $\Delta F(t)$ changes its vibrational behaviours whose motion $u(t)$ can now be obtained from:

$$m \cdot \ddot{u}(t) + D \cdot \dot{u}(t) + K \cdot u(t) = F_0(t) + \Delta F(t)$$

(8-3)

The residual vibrations $\Delta u(t)$ induced by the fault represent the difference between the vibrations produced in the damaged and undamaged system:

$$\Delta u(t) = u(t) - u_0(t)$$

(8-4)

Consequently the motion equation for the residual vibration generated by the fault is given by:

$$m \cdot \Delta\ddot{u}(t) + D \cdot \Delta\dot{u}(t) + K \cdot \Delta u(t) = \Delta F(t)$$

(8-5)

This equation was used to simulate the vibration response to unbalance, misalignment and bearing fault with the assumptions mentioned above. The solution of this equation can be obtained as the convolution of the force signal and the impulse response function $h(t)$ of the transmission path from the source to the measurement point if the force function is represented as a sequence of impacts:

$$\Delta u(t) = conv(\Delta F(t), h(t))$$

(8-6)

The solution of this equation can be calculated more efficiently in the frequency domain, as the convolution operation turns into a simple multiplication in the frequency domain:

$$\Delta u(f) = \Delta F(f) \cdot h(f)$$

(8-7)

As an example, Fig. 106 represents the results obtained in terms of displacement in a system with characteristics m=5kg, D=150Ns/m and K=10000N/m when an impulse force of increasing amplitude is applied:
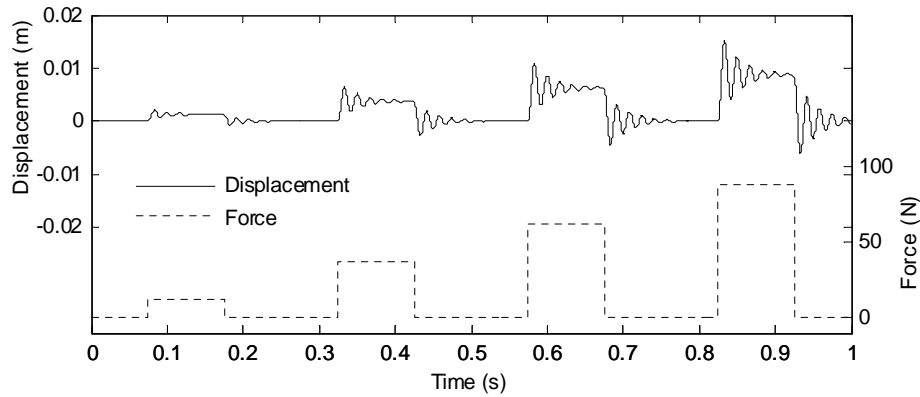
Fig. 106: Example of response to impact force

Similarly Fig. 107 shows an example of the response simulated in the same system working under a sinusoidal load with a period of 0.05s and increasing amplitude:
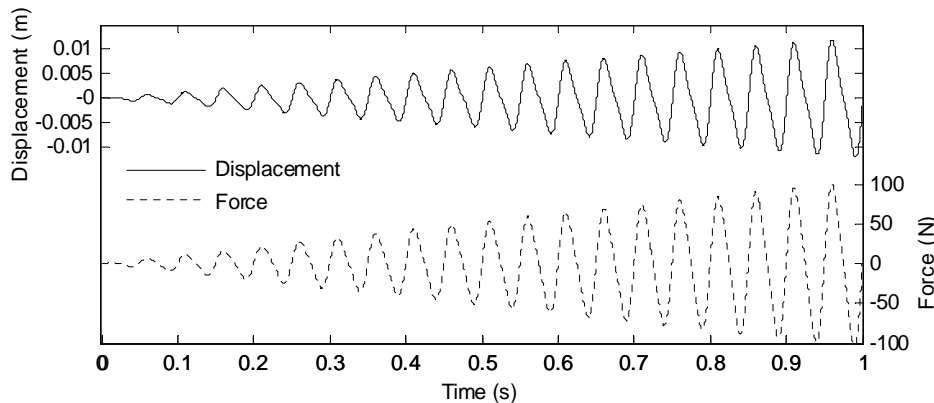
Fig. 107: Example of response to sinusoidal force

The impulse response function depends on the parameters m, D and K which define the dynamic properties of the transmission path between the source of the excitation fault (the corresponding bearing) and the measurement point. In order to simulate accurately the response generated by the fault it is important to select these parameters appropriately. Unfortunately the damping and stiffness in the motor installed in the rig used for this investigation were unknown, and its estimation via experimental tests such as the hammer test were complicated due to the difficulties related with the force application point

and the fact that there is a rolling element bearing in the transmission path. However the exact estimation of vibrational response is not required in this work, as its main aim is to demonstrate that it is possible to track changes in the vibration signal due to typical mechanical faults using CVA and not to simulate these faults with perfect accuracy. For this reason, the unknown parameters were selected according to typical values used in literature [153; 160] making sure that the results obtained are coherent with the vibration measurements taken from the system.

In order to simulate a fault that develops over time, and subsequently assess the sensitivity of the proposed method to different fault severities, throughout testing the amplitude of the residual loads simulated was increased linearly with time. The residual motion obtained from the simulation was differentiated twice to convert it to acceleration and added to the measured vibration signal in the appropriate units. The resulting combined signal was then processed using the methodology proposed in this section and analysed using CVA as explained in 4.2.1.

8.2.3 Experimental set up

The data sets used in this investigation were acquired from a laboratory-scale compressor test rig, designed to be able to function over a wide range of operating conditions through the control of the motor rotational speed and the opening of valves situated in the compressor inlet and outlet lines. The rig is shown schematically in Fig. 108, which also includes details of the different components of the installation and some of the measured signals. The list of tags of the variables included in the analysis is indicated in Table 35.
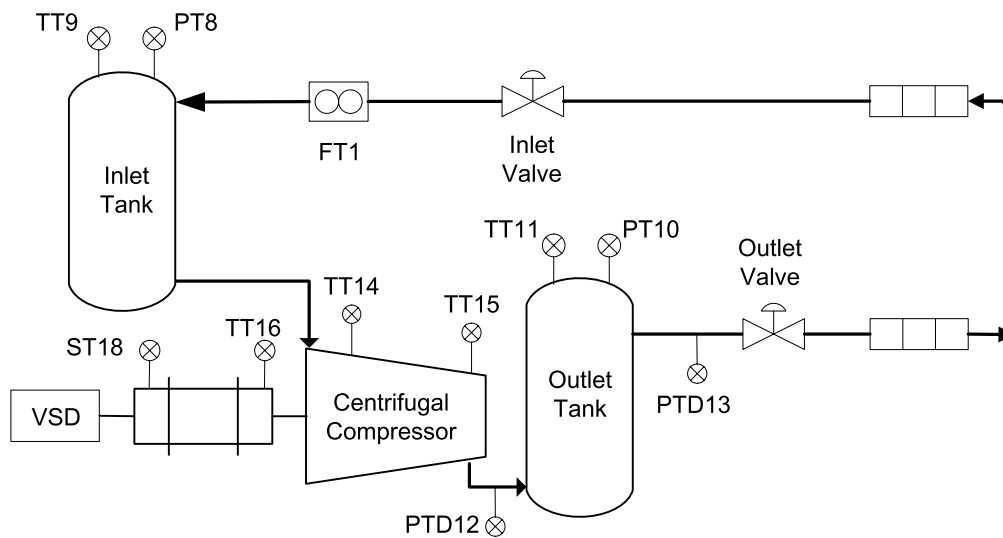
Fig. 108: Schematic representation of the gas compression experimental rig

Basically the rig is composed of a 5 stage centrifugal compressor driven by an induction motor the speed of which can be modified using a variable speed drive (VSD) connected to the grid through a transformer. The inlet line of the compressor is composed by a 90 mm diameter PVC pipe line, a silencer, an inlet valve operated by an electric motor and a 0.5m^3 tank. Similarly the outlet line is constituted by a 90 mm diameter PVC pipe line, a 0.5m^3 tank, a pneumatically operated valve and a silencer. Air is collected from and exhausted to the atmosphere. The motor-compressor arrangement can be seen in Fig. 109.

Fig. 109: Arrangement Motor-Compressor

The following list contains additional details about the main components of the rig:

-Variable Speed Drive: Model ABB ACS800, 11kW rated power. This drive can be configured to use scalar and vector control (Direct Torque Control) strategies, and admits motor speed or torque as reference inputs. Voltage and current measurement signals are processed and recorded by the high-level controller, as well as measurements of motor speed and torque estimated by the drive.

-Electric motor: Model ABB M3AA 160 MLB2, two pole 3-phase induction motor, 15kW rated power, speed range 0 – 6000 rpm. The motor is directly coupled to the compressor by a John Crane TSKS-0013 coupling. The motor speed is measured by an optical encoder on the shaft (sensor ST18 in Fig. 108); the temperature on the motor windings is also measured (sensor TT16 in Fig. 108).

-Compressor: Model Continental Industrie 020.05, five-stage centrifugal compressor, nominal flow 1000 m^3/h, pressure rise 300 mbar, 22kW rated power, maximum speed 4700 rpm. The temperatures of the bearings on the driven and non-driven ends are measured by TT14 and TT15.

-Inlet and outlet tanks: Each tank has a volume of 0.5 m^3. The tanks are used to model the volume of a pipeline downstream and upstream of the compressor in a larger installation. Together with the pipe length between them and the compressor inlet/outlet, the compressor surge frequency is set at 2.4 Hz for this particular installation. The tanks are also responsible for damping flow disturbances. The pressure and temperature in the tanks are measured respectively in PT8 and TT9 in the inlet tank and PT10 and TT11 in the outlet tank.

-Process valves: Two valves restrict the air flow in the installation at the inlet and outlet lines. The inlet valve is operated by an electric motor while the outlet valve is operated pneumatically.

-Main controller: The ABB AC800 PEC controller ensures a high level control over the drive and valve controls. The controller also collects and synchronizes

the signals from the drive and sensors. The controller is accessed via a desktop computer.

Table 35 summarizes the process variables acquired during the experiments:

Table 35: List of process variables

Variable nr	Location	Measured Magnitude	Units
1	Estimated	Motor torque	Nm
2	FT1	Air Flow	Sm^3/h
3	PT8	Inlet Tank Pressure	bar
4	TT16	Motor Phase Temperature	°C
5	TT9	Inlet Tank Temperature	°C
6	TT14	Left Bearing Temperature	°C
7	PDT12	Differential Pressure 1 (orifice plate)	bar
8	TT15	Right Bearing Temperature	°C
9	TT11	Outlet Tank Temperature	°C
10	Drive	Drive Current	A
11	ST18	Speed (Encoder)	rpm
12	ZT4	Outlet Valve Feedback	%
13	PT10	Outlet Tank Pressure	bar
14	PDT13	Differential Pressure 2 (orifice plate)	bar

In addition to the process variables listed in Table 35, 6 accelerometers and 3 current sensors were installed in the rig in order to monitor the vibration levels and motor current during the experiments. On the motor the accelerometers were placed in the horizontal and vertical direction in the drive end of the motor casing (Fig. 110), in addition to a third accelerometer placed in one of the motor feet (Fig. 111). On the compressor 2 accelerometers were placed in the casing of the drive end bearing (vertical and horizontal, Fig. 112) and a third one was placed in the axial direction in the non-drive end (Fig. 113). The accelerometers used were SKF CMSS2110 (frequency range 0.8Hz-10kHz) in all cases except in the motor foot, where the model IMI 608A11 (frequency range 0.5Hz-10kHz) was selected due to its smaller size. The vibration data was acquired at a sampling rate of 5120Hz using a NI 9234 data acquisition card which included an antialiasing filter. The current drawn by the motor was measured in each one

of the phases using an ABB EL55P2 sensor. The current signal was acquired at a sampling rate of 5kHz using a NI9203 data acquisition card. The measurements of vibration and current were synchronised with process measurements by the use of the timestamps produced by the data acquisition system, making sure that the start and end time for all the data sets was exactly the same in each experiment.

Fig. 110: Motor casing accelerometers
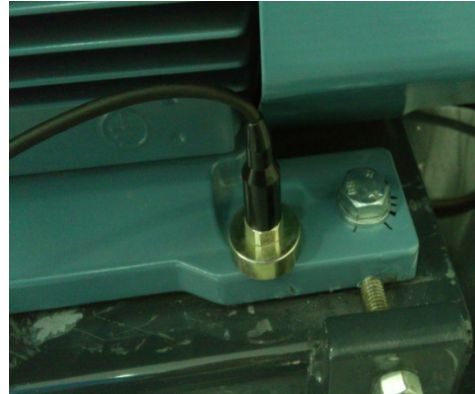
Fig. 111: Motor foot accelerometers

Fig. 112: Compressor driven end

Fig. 113: Compressor non-driven end

8.2.4 Acquisition of data sets

During the various experiments conducted, the process variables listed in Table 35 were originally sampled at 1 kHz. This sampling rate is much higher than that which is typically used in industry to monitor process variables such as pressure, flow rate or temperature. This relatively high sampling rate was

selected to allow process, electrical and vibration variables to be compared more easily, allowing various multivariate approaches to be investigated. In this investigation the acquired process variable data sets were down sampled to 1 Hz, which is more in line with sampling rates typically seen in industry. During the tests, the motor speed and the position of the outlet valve were given different control set points in order to obtain data from the system working under variable operating conditions. Throughout testing the inlet valve was set to be fully open. As well as reducing the number of potential operating points to be considered as normal operation, this approach allowed the valve to be used to simulate a blockage in the inlet pipeline.

In addition to the process variables listed in Table 35, 15 additional variables extracted from the current, vibration and pressure sensors were included in the analysis following the methodology presented in section 8.2.1. The vibration signal observed in the accelerometers placed in the horizontal direction in the motor and the compressor was very similar to the signal acquired from the sensors in the vertical direction. In order to avoid redundant information and reduce computation time, the signals measured in the horizontal direction were rejected for the CVA analysis of the data. Similarly only one of the three current measurements was included due to the similarities found in the signals in all three phases. The features extracted from the vibration, current and pressure signals were selected to give the CVA method a high sensitivity to typical faults such as unbalance, misalignment or compressor surge. The features selected for the CVA analysis were:

-The amplitude of the peaks at the rotational speed (1X) and its second harmonic (2X) in the spectrum of current signal, vertical vibration measured in the compressor, vertical vibration measured in the motor casing and vertical vibration measured in the motor foot.

-RMS value of the current signal, vertical vibration measured in the compressor, vertical vibration measured in the motor casing and vertical vibration measured in the motor foot.

-Amplitude of the peak at 2.4Hz (Surge frequency) and its second harmonic (4.8Hz) in Differential Pressure 1 signal (PDT12) and amplitude of the peak at 2.4Hz in the outlet tank pressure signal (PT10). Obviously these features were extracted from the original signals acquired at 1 kHz before the data was down sampled.

In order to obtain a reasonable resolution in the frequency spectrum of all the signals studied, the length L of the window analysed was 1s for current and vibration measurements and 5 seconds for pressure and differential pressure measurements. Table 36 summarizes the variables constructed by the extraction of these characteristic features in the time and frequency domain which were included in the CVA analysis:

Table 36: List of process variables

Variable nr	Origin	Measured Magnitude	Units
15	Current	Amplitude of peak at 1X	A
16	Current	Amplitude of peak at 2X	A
17	Current	RMS	A
18	Compressor vibration	Amplitude of peak at 1X	g
19	Compressor vibration	Amplitude of peak at 2X	g
20	Compressor vibration	RMS	g
21	Motor casing vibration	Amplitude of peak at 1X	g
22	Motor casing vibration	Amplitude of peak at 2X	g
23	Motor casing vibration	RMS	g
24	Motor foot vibration	Amplitude of peak at 1X	g
25	Motor foot vibration	Amplitude of peak at 2X	g
26	Motor foot vibration	RMS	g
27	Differential pressure1 (DPT12)	Amplitude of peak at 2.4Hz	bar
28	Differential pressure1 (DPT12)	Amplitude of peak at 4.8Hz	bar
29	Outlet tank pressure (PT10)	Amplitude of peak at 2.4Hz	bar

Three data sets (T1, T2 and T3) were acquired from the system working under normal operating conditions for training purposes. To ensure that the conditions tested during normal operation were representative 16 different combinations of motor speed and outlet valve position were chosen to be tested in each one of the training data sets (see Table 37). Note that the 16 points were not present in

every data set, and some operating points were repeated during a test period. The operating points were selected to cover a wide range of operating conditions within the safe operating region in the compressor map, trying to avoid the surge and choke regions. In each training data set, the operational conditions were changed, though not identically, in order to obtain a good variety of large, small, long and slow process changes happening in different directions (increment or decrement). The objective of this variety in the operational conditions is to ensure that the dynamics of the system are captured in all circumstances. As an example, Fig. 114 represents the set points for the operating conditions tested in data set T1.

Table 37: Typical set point values for motor speed and outlet valve position

Motor speed (rpm)	1000	2000	3000	4000
Outlet valve position (%)	40	60	80	100
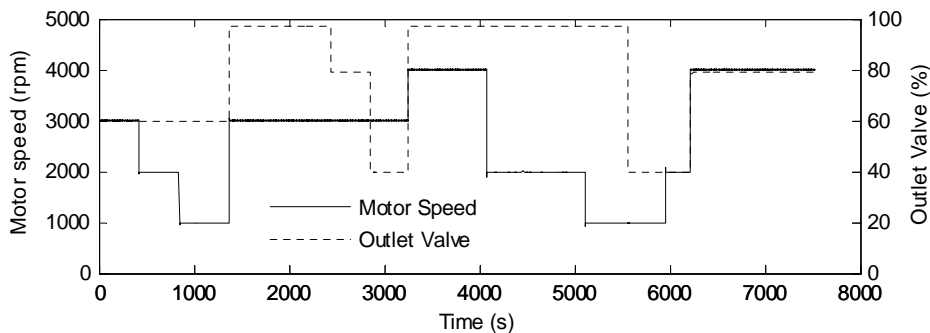
Fig. 114: Operating conditions for T1

In addition to the training data sets, different sets of data were acquired from the system working under artificially induced faulty conditions where 4 different process faults were seeded into the system. These faults simulate typical malfunctions that could be experienced in a real system such as pipeline blockages, electric supply perturbations or abnormal operating conditions (stall and surge). These faults were introduced gradually when possible, in order to observe how the severity of the fault affects the indicator. After reaching a certain level of severity the fault condition was removed, returning the system to

normal conditions. The case of surge was also studied from the vibration point of view by removing the variables related with pressure measurements from the analysis. This allowed the detection of the fault to be based on vibration measurements in order to prove the validity of the proposed method using real (not simulated) vibration data. In addition to these process faults, 3 mechanical faults were simulated following the procedure proposed in 8.2.2, including unbalance, misalignment and bearing fault.

The faulty condition was introduced after a certain period of normal operation when possible in order to observe changes in the health indicators. During these tests the flow rate conditions were changed in a similar way as was undertaken for normal operation. The following list describes each one of the faulty conditions studied:

-Inlet line blockage: This fault was introduced by closing gradually the inlet valve, which was fully opened during the training period (normal operation). The objective of this test is to simulate a pipeline blockage that can happen in a real installation due to dirt accumulating in critical points of the line, deteriorated filters, corrosion in the pipe lines, valve malfunctions, etc. This fault causes an increment in the pressure drop of the pipe line, affecting the performance of the whole system.

-Speed set point perturbation: The objective of studying this fault is to observe how a sinusoidal perturbation in the speed set point can affect the performance of the system. This fault can be caused in a real system by communication errors or incorrectly tuned controller gains, and has a similar effect as a perturbation in the frequency of the current supplied to the motor.

-Compressor stall: Consists in a localised airfoil stall which does not destabilise the compressor completely but causes a decrement in the compressor effectiveness. It is normally caused by the operation of the machine in the boundary of the design limits. This fault was simulated by closing gradually the outlet valve beyond the limits tested during the training period in order to increase the pressure ratio and force the stall to appear.

-Compressor surge: Complete disruption of the air compression and reversal of the air flow inside the compressor, caused by the inability of the machine to work against the already compressed air when the pressure ratio is too high. After this disruption, the compressor will reach stable flow once the pressure ratio is reduced to a normal level and, if the operating conditions that caused the surge remain, the cycle will restart producing a new flow reversal. The appearance of this phenomenon can be critical as it can cause catastrophic damage in a compressor, but this particular rig is designed to be able to cope with surge conditions for a certain amount of time. In order to study this abnormal operating condition, surge will be induced by moving the operating point of the machine to the surge region, in the left hand side of the compressor map.

The mechanical faults simulated and added to the measured data were explained in section 8.2.2 (rotor unbalance, shaft misalignment and bearing fault).

8.3 Results and discussion

The results provided by the application of CVA for the detection and diagnosis of the faulty conditions introduced above are presented in this section. The first step for the application of CVA for monitoring purposes is the calculation of the transformation matrices and the threshold for the T^2 and Q indicators using data acquired under normal operating conditions.

The optimal number of past and future lags considered in the analysis (p and f) can be calculated computing the autocorrelation function of the summed squares of all measurements [17]. The autocorrelation function measures the cross correlation between a signal and a delayed version of itself at different lags. The objective of this analysis is to take into consideration a number of lags which contains relevant information in terms of autocorrelation. Fig. 115 shows an example of autocorrelation function for the training data set T1 against a confidence bound of ±5% (dashed lines).
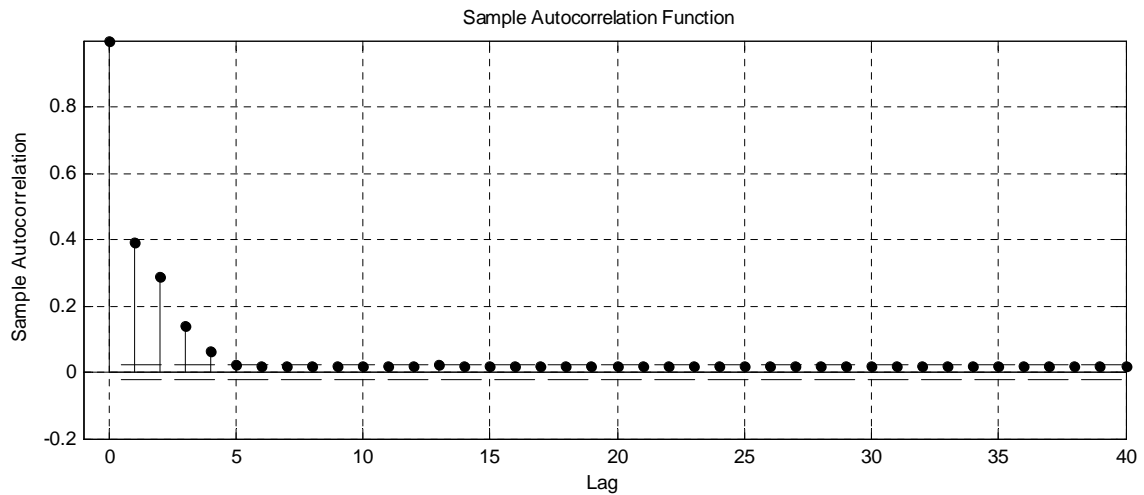
Fig. 115: Sample autocorrelation function against ±5% confidence bound (dashed lines).

For this study p and f were set to 5 according to the results obtained from the analysis of the autocorrelation function of the three training data sets. The most common methodologies used for the calculation of the optimal number of states retained r are the analysis of the dominant singular values in the matrix D [84] and methodologies based on the Akaike Information Criterion (AIC) [2]. Fig. 116 shows an example of normalized singular values obtained from (4-11) applying CVA to the training data set T1. In this particular case, setting the number of retained states based on the dominant singular values will result in an unrealistic high order model because the singular values decrease slowly [17]. In addition, the number of states retained is not especially relevant for this study because both statistical indicators (T^2 and Q) are used simultaneously for fault detection. As a result, the system variations not captured in the retained space (represented by T^2) will be captured by the residual space (Q) and vice versa.

It is essential for the analysis to use a training set that covers the entire spectrum of possible operational conditions. It is possible to combine different training data sets in order to enrich the variety of the data included in the analysis. In this investigation the combination was done by calculating the past and future matrices individually for each training data set according to (4-4) and (4-5) and then merging the matrices obtained. The original length of the data sets was 7527 s for T1, 12653 s for T2 and 9663 s for T3.
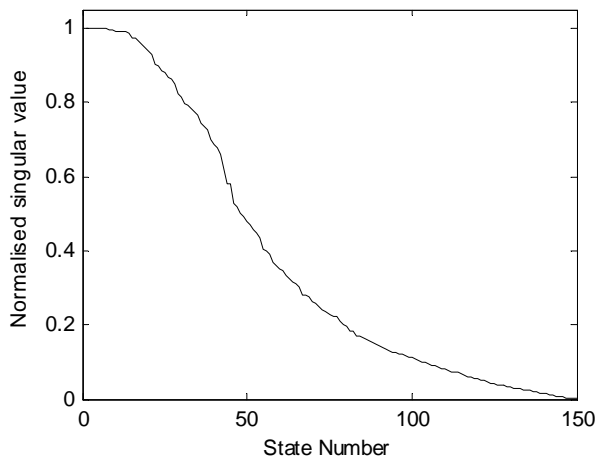
Fig. 116: Normalized singular values for T1

Following a similar procedure as in 4.3.1 the three data sets were mixed in pairs generating 3 different combined sets that were used to check its capacity to represent the system dynamics accurately producing a low number of false alarms. Again, the objective of analysing these three combined data sets is to select the data set combination which produces a lower false alarm rate when the remaining data set is used for the monitoring period. CVA was performed for each one of the three combined training data sets using a range of values for the number of states retained r in order to select the optimum value for this parameter. For low values of r the number of false alarms produced is high because the retained space is not able to accurately represent the states of the system and consequently the number of the T^2 threshold violations increases. On the other hand if the state order selected is too high it results in the model overfitting the data [2], increasing again the false alarm rate.
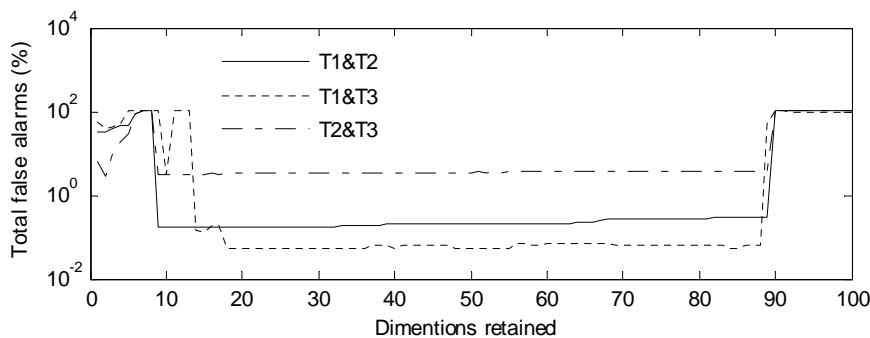

Fig. 117: Analysis of the influence of the number of states retained

The data set combination which produced lower false alarm rates was T1 and T3. The number of states retained was set to 30 (r=30) for the data analyses in order to minimize the false alarm rate in normal conditions (see Fig. 117). The UCL calculated using KDE for this configuration using 99% confidence bound was 5481.42 and 7038.46 for the T^2 and Q indicators respectively.

8.3.1 Process Faults

3.1.1 Inlet line blockage

This fault simulates a blockage in the inlet line achieved closing gradually the inlet valve which was fully open under normal operating conditions. Fig. 118 represents the operating conditions (motor speed and outlet valve position) during the test (a) and the fault evolution (inlet valve position) (b).
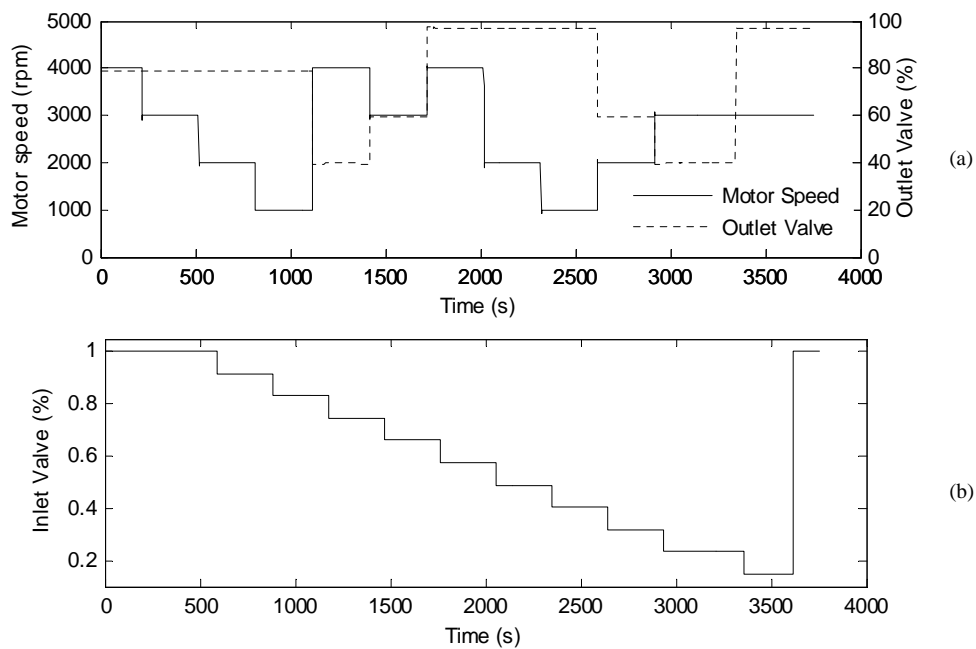

Fig. 118: Operational conditions (a) and fault evolution (b) for inlet line blockage

The test duration was 3752 seconds. The valve was closed at a rate of 8.5% every 5 minutes starting at sample 594. The valve was reopened at sample 3616 to return to normal operating conditions. Fig. 119 represents the results obtained in terms of fault detection (a) and diagnosis (b).
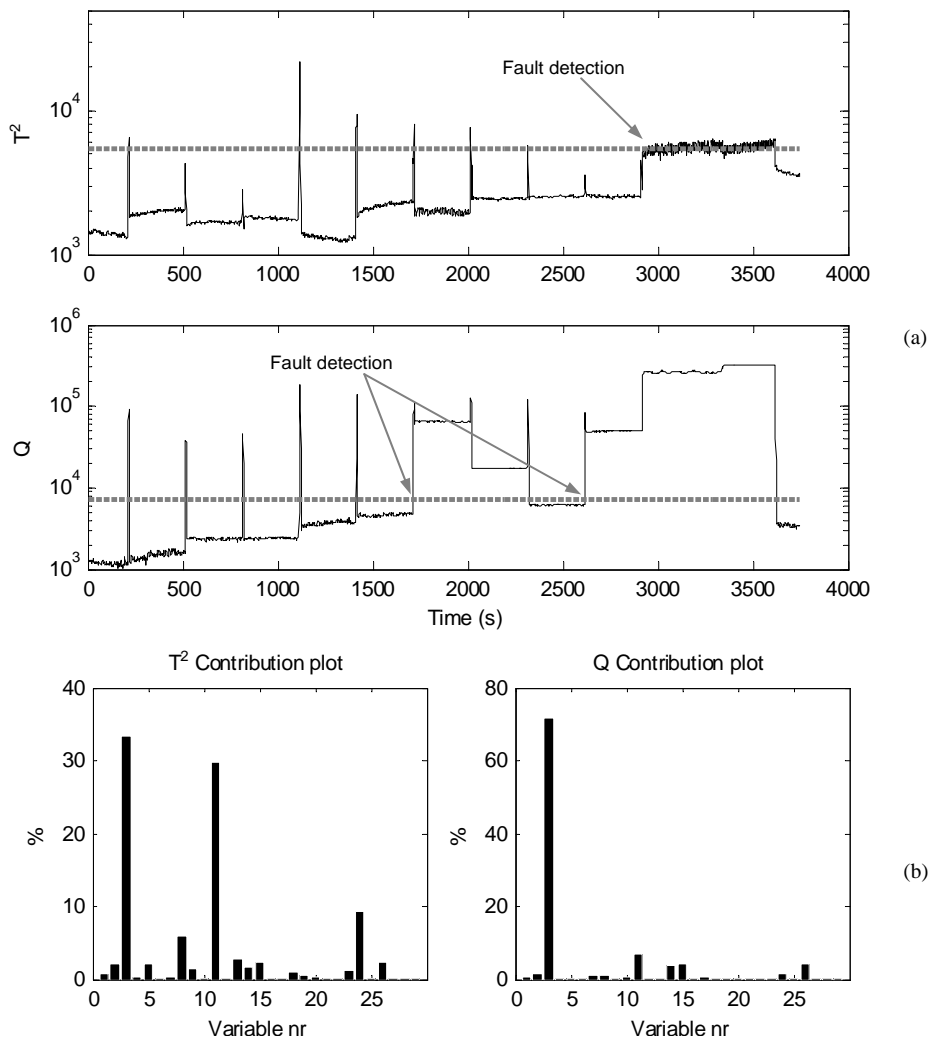
Fig. 119: Results for inlet line blockage: T^2 and Q indicators (a) and contribution plots at sample 1714 (b)

For this data set the first fault detection occurred at sample 2939 after 6 short false alarms for the T^2 statistical indicator and at samples 1714 and again in 2615 for the Q indicator after 5 short false alarms. These samples correspond to a valve opening of 23.5%, 66% and 40.5% respectively. Both indicators fall below the UCL when the fault is removed opening the inlet valve completely. In this particular case the Q indicator shows a much better performance than T^2 in terms of detection time. Both indicators produce short false alarms corresponding with step changes in the motor speed. In the case of Q, after the first fault detection the indicator value falls below the threshold around sample 2300 despite the fact the fault was not removed. This change in the indicator

value is attributed to the change of the motor speed at that point; the reduction in the motor speed and the consequent drop in the air flow reduce the pressure losses at the input valve, which minimizes the fault impact. The contribution plots at the fault detection time (sample 1714) are represented in Fig. 119(b). The variable contributing more to the final value of both indicators is the inlet tank pressure (variable 3), which indicates that the faulty condition is causing some kind of conflict with the pressure in the compressor inlet. For the T^2 indicator the contribution of the motor speed is also significant (variable 11), but at this point the fault has not been detected yet by T^2.

Speed set point perturbation

This fault simulates a perturbation in the motor speed set point for the system controller. Fig. 120 represents the operating conditions (motor speed and outlet valve position) during the test (a) and the fault evolution (set point perturbation) (b).
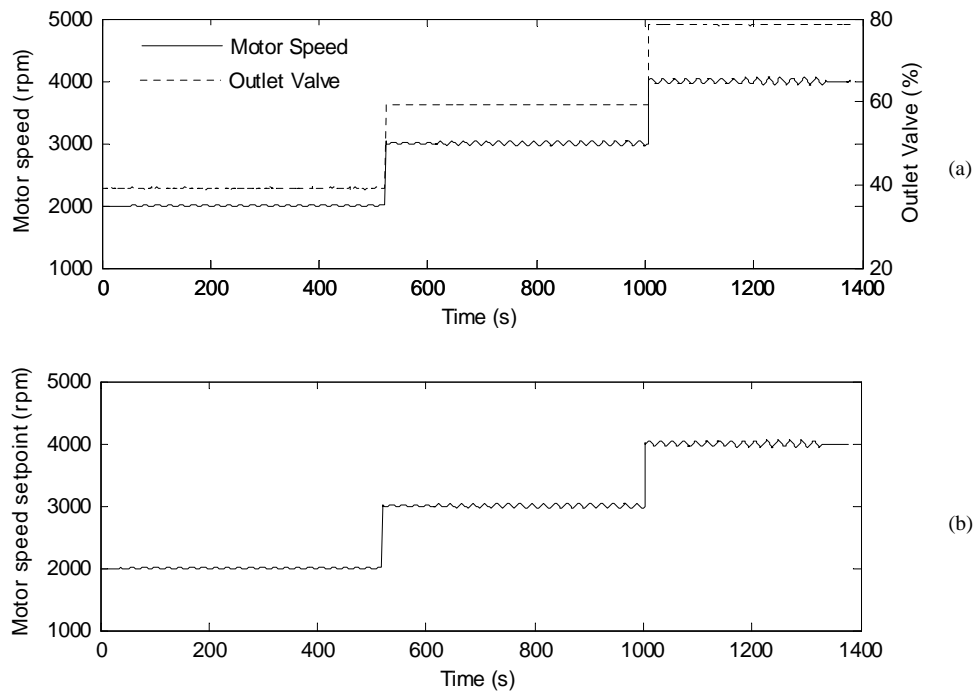
Fig. 120: Operational conditions (a) and fault evolution (b) for speed set point perturbation

The test duration was 1377 seconds. The speed set point was altered from the controller software adding a sinusoidal function with an oscillating period of 20

seconds and increasing amplitude of ±60rpm in the last stages of the test. The perturbation was introduced at sample 33 and removed at sample 1330. Fig. 121 represents the results obtained in terms of fault detection (a) and diagnosis (b).
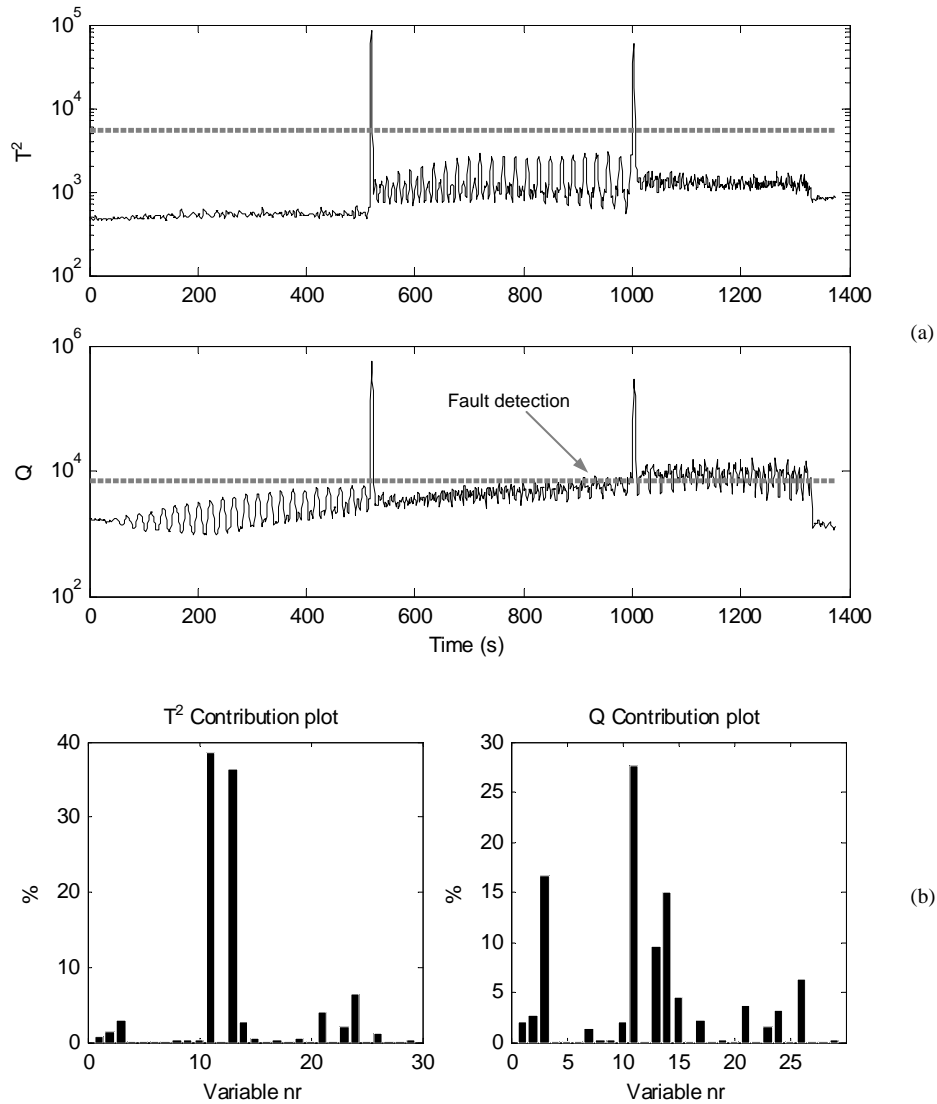
Fig. 121: Results for speed set point perturbation: T^2 and Q indicators (a) and contribution plots at sample 885(b)

For this data set the fault was not effectively detected by T^2, which only produced two false alarms related to motor speed changes, but it was detected at sample 885 by the Q indicator after one short false alarm. Sample 885 corresponds with a perturbation amplitude of ±40rpm. The Q indicator falls below the UCL when the perturbation is removed. It is important to notice that in

this particular case both indicators show an oscillating behaviour which corresponds with the oscillating nature of the defect seeded. Despite the lack of precision of the T^2 indicator in detecting the fault, looking at the trend of the indicator it is obvious that it is being affected by the fault, and it would have been possible to detect it if the UCL was lower. The contribution plots at the fault detection time (sample 885) are represented in Fig. 121(b). The variable contributing most to the final value of both indicators is the motor speed (variable 11), which indicates that the faulty condition is indeed affecting the motor speed. It is also noticeable the high contribution of the outlet tank pressure (variable 14) in both indicators and the flow rate (variable 3) and differential pressure 2 (variable 13) particularly in the Q indicator. This gives an indication that the fault introduced is not only affecting the rotating speed of the machine but also the performance of the system in terms of air supply. In addition there is a significant contribution of the variables measuring changes in vibration (variables 21, 24 and 26), indicating that the non-steady speed is also affecting the mechanical response of the system.

Compressor stall

Compressor stall was simulated closing gradually the outlet valve to increase the pressure ratio. Fig. 122 represents the operating conditions (motor speed and outlet valve position) during the test (a), and the effect of the fault in the measured outlet pressure (b). The fault evolution corresponds with the position of the outlet valve represented in (a)
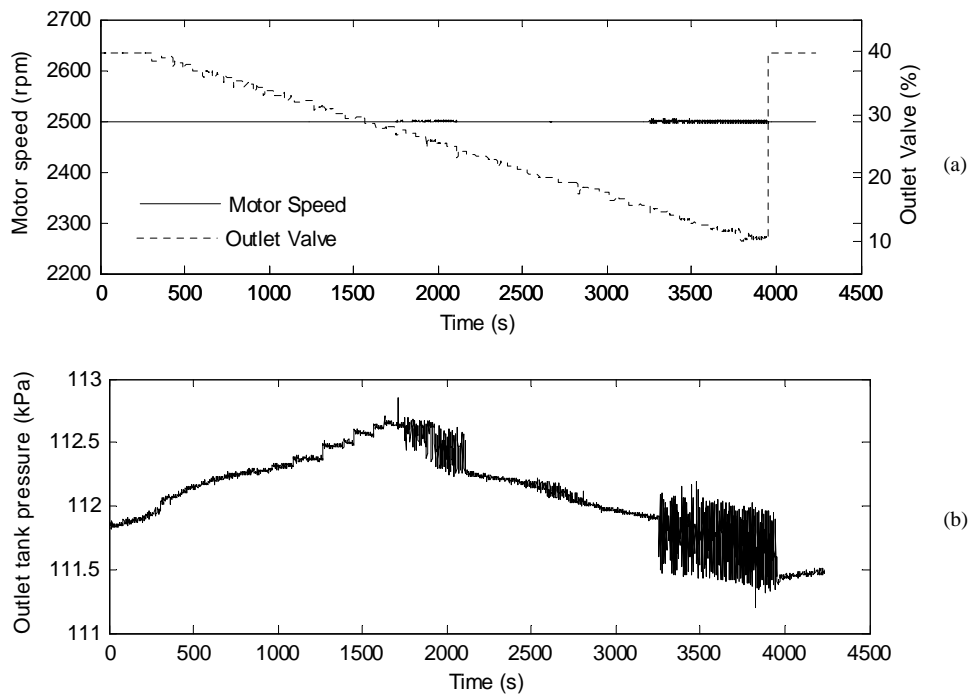
Fig. 122: Operational conditions (a) and fault evolution (b) for stall

The test duration was 4224 seconds. The outlet valve was closed gradually at a rate of 0.5%/min starting from normal operating conditions (40%) at sample 307 to a minimum of 10% at sample 3978. The valve was returned to 40% opening at sample 3947. It can be seen in Fig. 122(b) how initially the outlet pressure increases as the outlet valve is closed as expected, but around sample 1600 this trend changes. After a period of low amplitude oscillations (probably generated by mild surge) the outlet pressure decreases despite the outlet valve being gradually closed. Then around sample 3500 the behaviour changes dramatically with high amplitude oscillations at 2.4Hz which evidence the presence of surge. It is important to notice that the motor speed set point during this experiment was 2500rpm, which was not one of the selected set points during the training period. Fig. 123 represents the results obtained in terms of fault detection (a) and diagnosis (b).

223
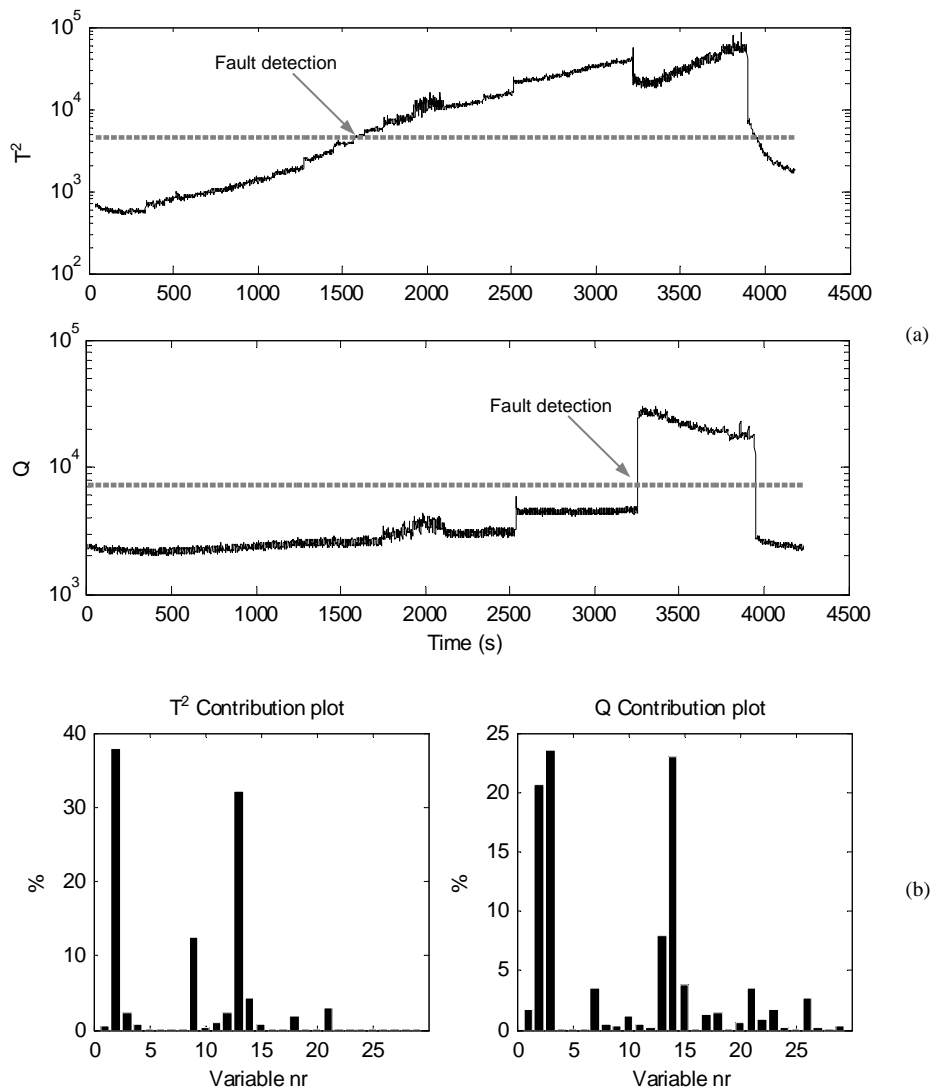
Fig. 123: Results for stall: T^2 and Q indicators (a) and contribution plots at sample 1572(b)

For this data set the first fault detection happened at sample 1572 for the T^2 statistical indicator and at sample 3254 for the Q indicator without short false alarms in any of them. These samples correspond to an outlet valve opening of 28.5% and 14.5% respectively. Both indicators fall below the UCL when the fault is removed by opening the outlet valve to 40%. In this particular case the T^2 indicator shows a much better performance than Q in terms of detection time. T^2 seems to be able to detect the effects of stall from the early stages of degradation, but in Q the fault detection only happens when the compressor is actually suffering surge, in the last stages of the experiment. The contribution

plots at the fault detection time (sample 1572) are given in Fig. 123(b). The variables contributing most to the final value of the T^2 indicator are the air flow rate (variable 2) and the outlet tank pressure (variable 13), pointing to a conflict with the capacity to deliver pressurized air. In the case of Q the most significant variables are the inlet tank pressure (variable 3), differential pressure 2 (variable 14) and air flow rate (variable 2). If the contribution plots are analysed at the point where the outlet pressure surge suffers large oscillations (sample 3254), the result is completely different (see Fig. 124).
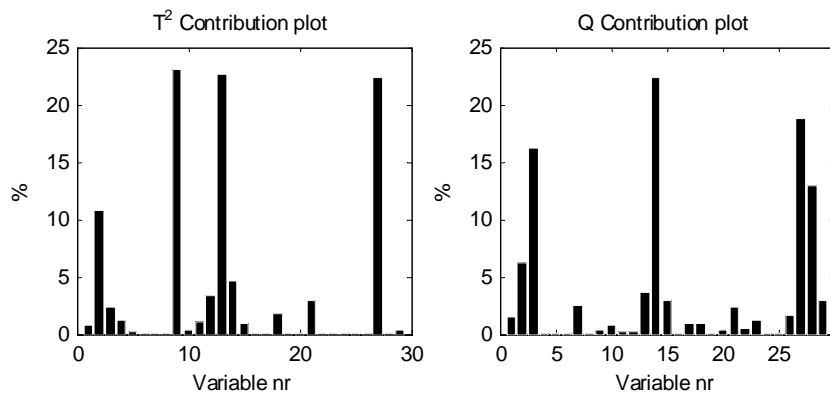
Fig. 124: Contribution plot at sample 3254 for compressor stall

The main difference is the increment of the contribution for variables 27, 28 and 29, which are related with the amplitude of the peaks at 2.4Hz and its second harmonic in the signal "Differential pressure 1" and the peak amplitude at 2.4Hz observed in the spectrum of the outlet tank pressure. The presence of these peaks reveals the appearance of compressor surge when the conditions were appropriate for it.

Compressor surge

Compressor surge was introduced moving the compressor operating point to regimes where surge is expected. Fig. 125 represents the operating conditions (motor speed and outlet valve position) during the test (a), and the effect of the fault in the measured outlet pressure (b). The operating conditions for which surge is expected are shaded in grey.
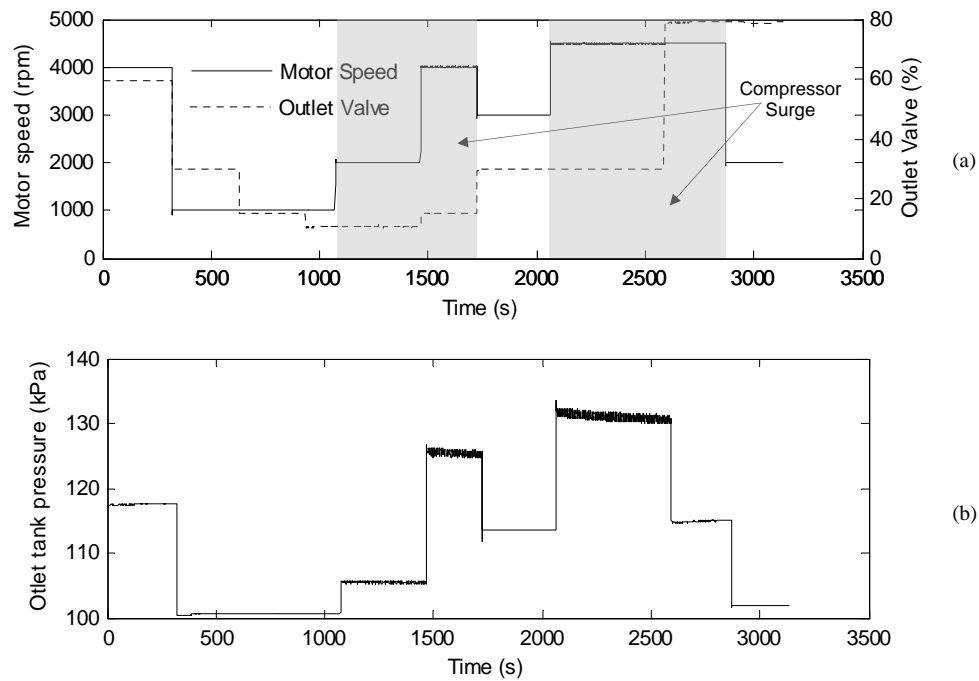
Fig. 125: Operational conditions (a) and fault evolution (b) for surge

The test duration was 3131 seconds. The surge conditions were introduced from sample 1071 to 1725 and 2062 to 2868. Normal operation is expected for the rest of samples in the test. It can be seen in Fig. 125(b) how the faulty condition introduced caused oscillations in the outlet tank pressure. Fig. 126 represents the results obtained in terms of fault detection (a) and diagnosis (b).
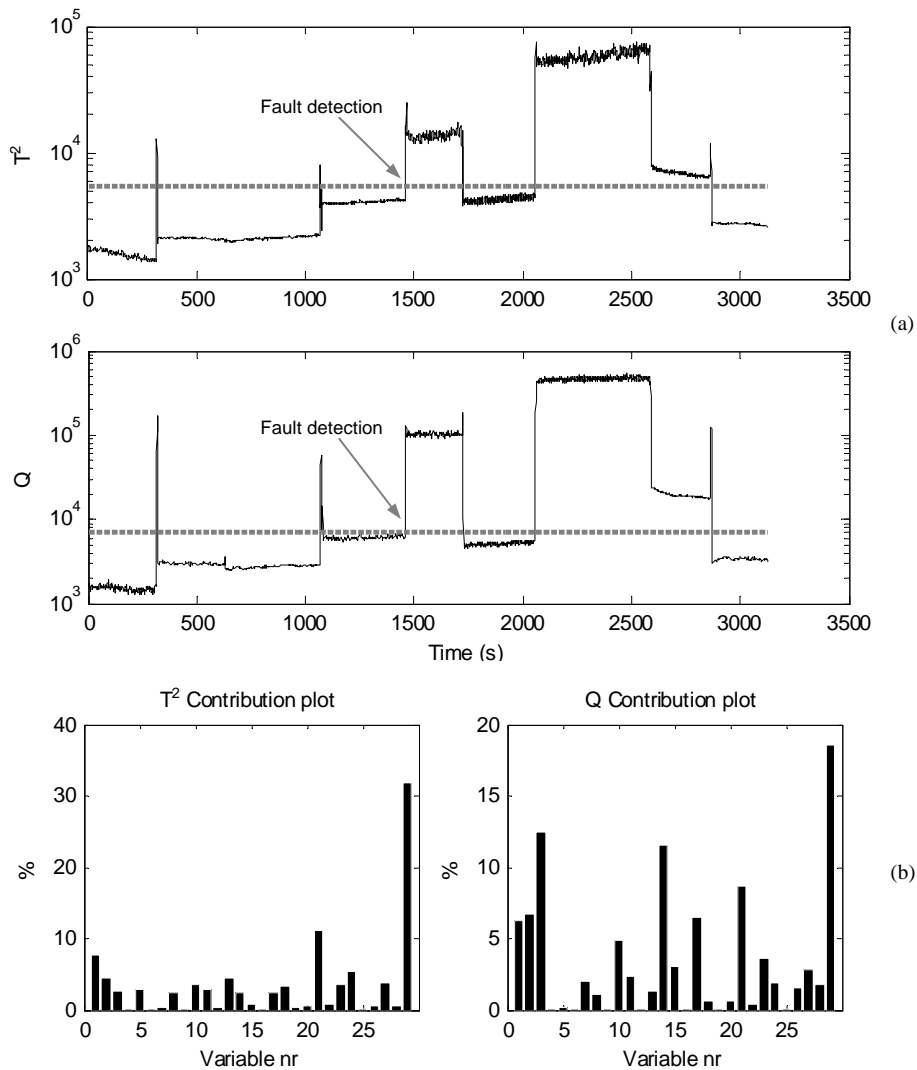
Fig. 126: Results for surge: T^2 **and** Q **indicators (a) and contribution plots at sample 1459(b)**

For this data set the first fault detection happened at sample 1459 for both indicators after two short false alarms which corresponded to step changes in the motor speed. The surge conditions induced from sample 1071 to 1457 were not detected by the algorithm, although the value of both indicators is high for that region. The reason for the lack of detection in that region is attributed to the low severity of surge for that particular conditions (2000 rpm and 10% valve opening), but the fault was detected almost immediately as surge was introduced at higher rotational speeds. Both indicators fall below the UCL when the fault is removed. The contribution plots at the first fault detection time (sample 1459) are represented in Fig. 126(b). The variable contributing most to

227

the final value of both indicators is the amplitude of the peak at 2.4Hz in the frequency spectrum of the outlet tank pressure signal (variable 29). This evidences the fact that the fault is effectively caused by compressor surge and reveals the importance of being able to track features in the frequency spectrum to increase the sensibility of CVA to certain types of faults. Other significant variables contributing to the indicators are amplitude of the 1X peak in the motor vibration (variable 21), and the torque, air flow rate, inlet pressure, differential pressure 2 and current RMS value in the case of Q indicator (variables 1,2,3,14 and 17). Fig. 127 shows the evolution of the amplitude of the 2.4 Hz peak in the outlet tank pressure frequency spectrum (variable 29) (a) and a detail of the outlet tank pressure signal showing the oscillating nature of compressor surge (b). It can be seen how effectively the period of oscillation is around 0.42 seconds (2.4Hz).
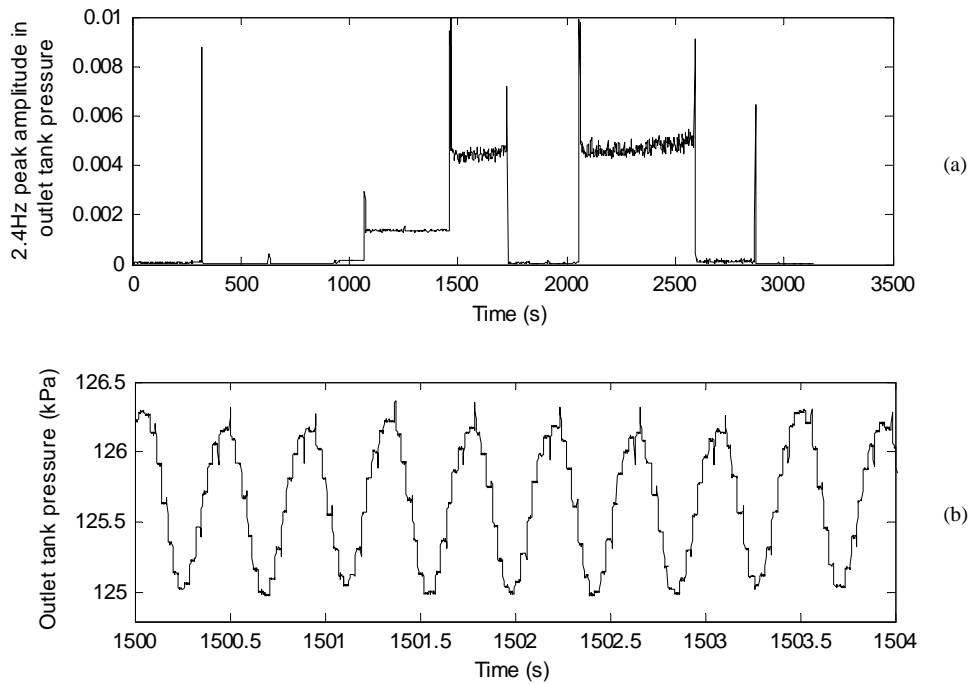
Fig. 127: Evolution of the 2.4Hz peak amplitude (a) and detail of surge pressure oscillations (b)

Another interesting fact observed in this test was the propagation of the fault through the system affecting variables of different nature: process, mechanical and electrical. One evident consequence of surge is the increase in the vibrations, observed in the contribution plots as significant contributions coming

from the variables measuring vibration features. Another effect of the oscillating pressure and reversal flow is the oscillation in the torque demanded by the machine (see Fig. 128(a)) and the consequent modulation in the current feeding the motor (see Fig. 128 (b)) which generates sidebands in the current frequency spectrum separated 2.4Hz from the mean peak (Fig. 128 (c).
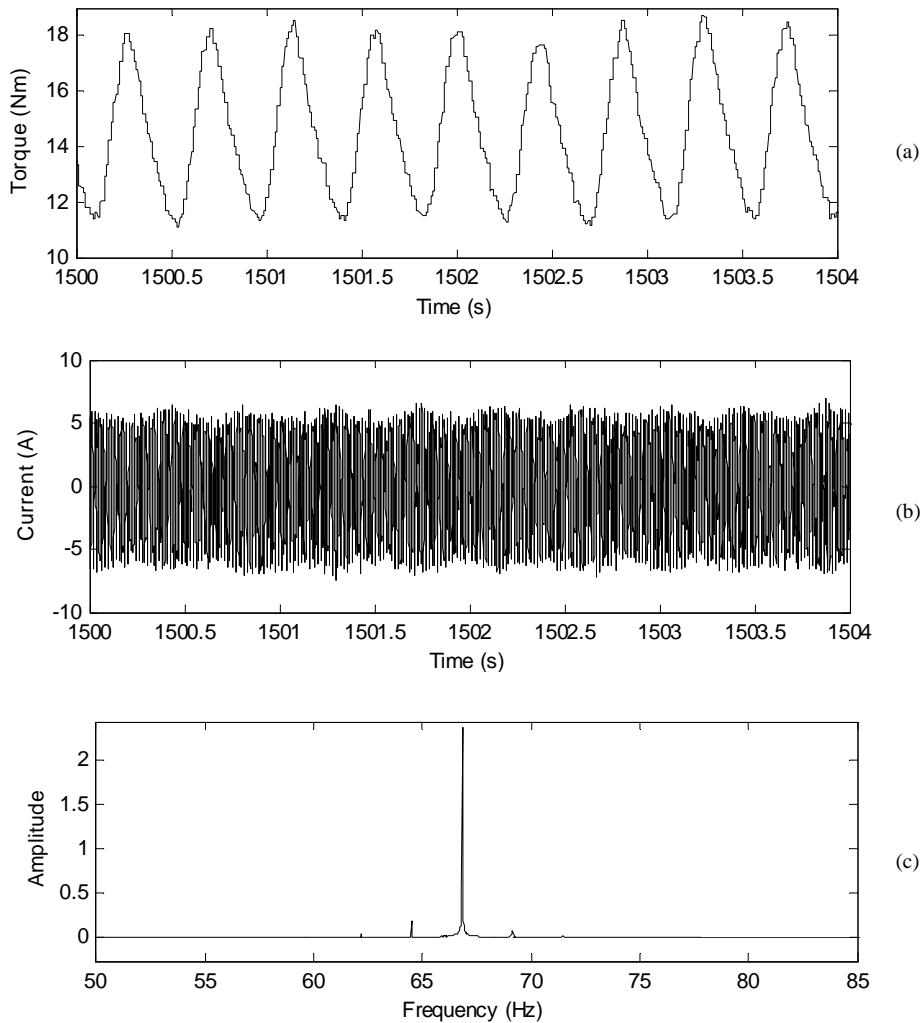
Fig. 128: Effects of surge on the torque (a), current (b) and current spectrum (c)

Surge detection through vibration data

In the previous section it was shown how compressor surge was effectively detected using CVA, being the variables related with features in the pressure signal spectrum and pressure measurements themselves the most significant variables in terms of fault diagnosis (Fig. 126 (b)). Nevertheless, if those variables are removed from the analysis as well as variables related with

features in the vibration and current signal spectrum, the fault is not detected by CVA with only the remaining variables. Fig. 129 represents the results obtained from CVA analysing the same surge data set presented in the previous section but including only variables 1,2,4,5,6,8,9,10,11 and 12 (related with measurements of torque, flow rate, temperatures, current, speed and valve position) in the analysis:
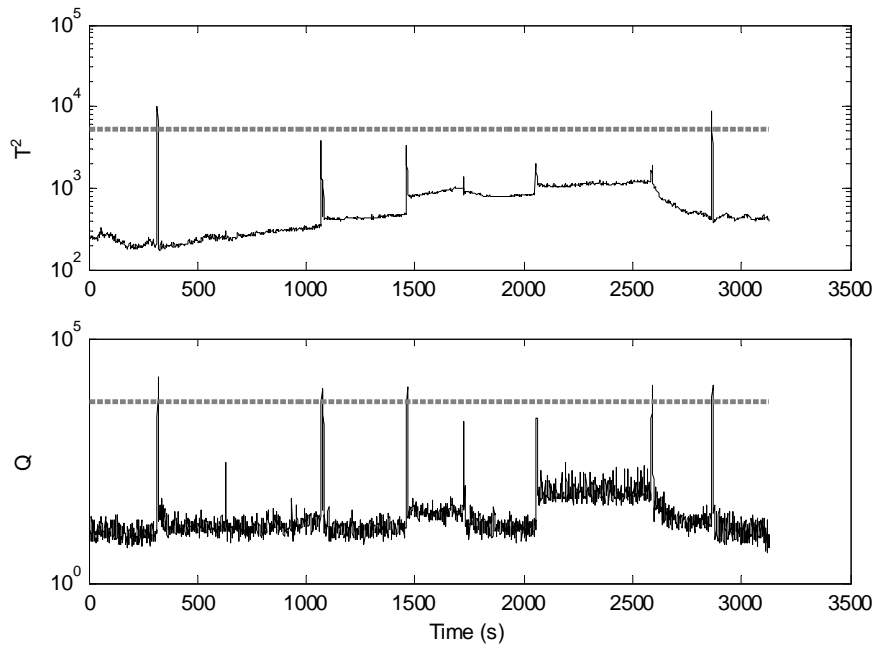
Fig. 129: T^2 **and** Q **indicators for surge detection with limited variables**

In this case the fault is invisible for the algorithm for all operating conditions. However, if the variables related with current and vibration features are added to the analysis (variables 15 to 26) the results obtained in terms of fault detection and diagnosis are very different:
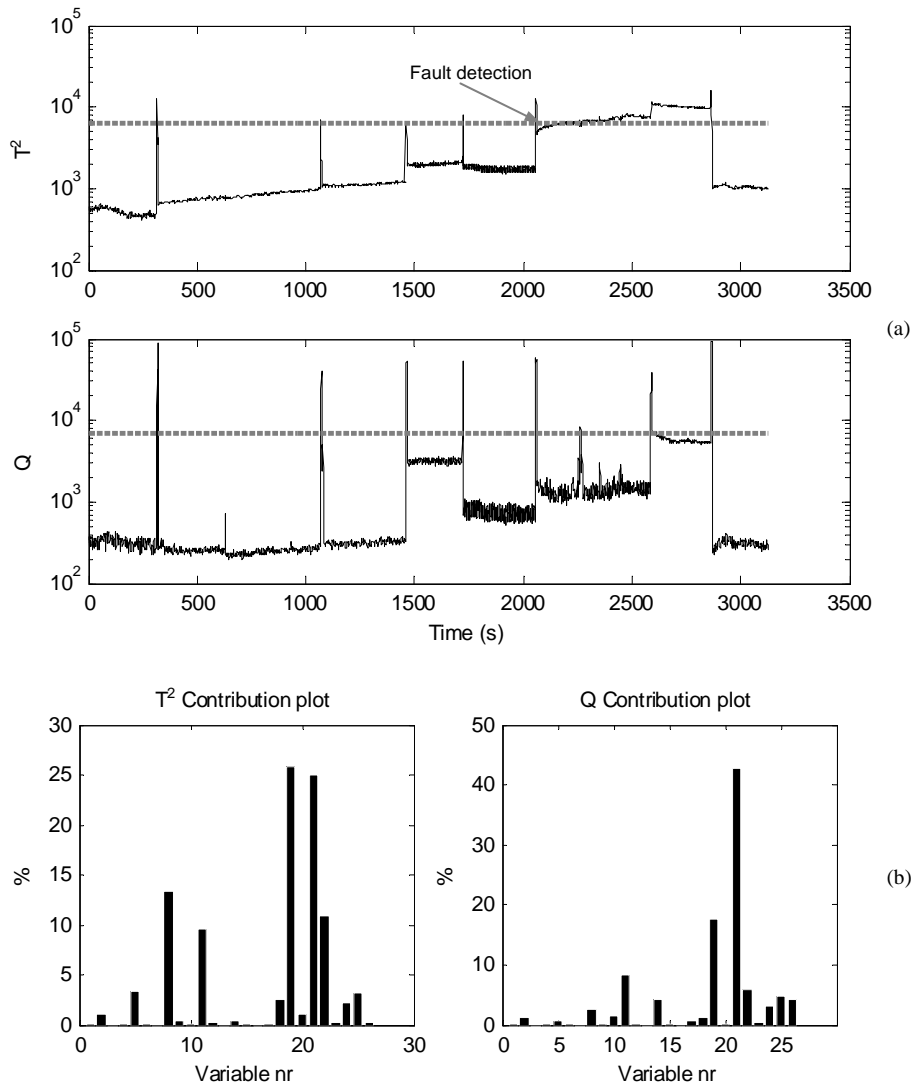
Fig. 130: Results for bearing fault: T^2 and Q indicators (a) and contribution plots at sample 2174(b)

In terms of fault detection, the results are not as good as in Fig. 126(a) but the addition of vibration data allowed the detection of the fault, demonstrating the validity of the method using real (not simulated) vibration data. The fault was detected only by the T^2 statistic in the second introduction of surge at sample 2174 after some false alarms. The contribution plots at sample 2174 reveal the importance of the vibration features in the analysis, the amplitude of the 1X peak and the RMS value of the compressor signal being the most significant variables.

231

8.3.2 Mechanical faults

The vibrational signature of 3 typical mechanical faults (unbalance, misalignment and bearing fault) was simulated using the procedures introduced in section 8.2.2 The vibration signal obtained from the fault simulation was added to the vibration signal of a specific data set acquired from the rig under normal operating conditions. The operating points selected during the data sets are represented in Fig. 131.
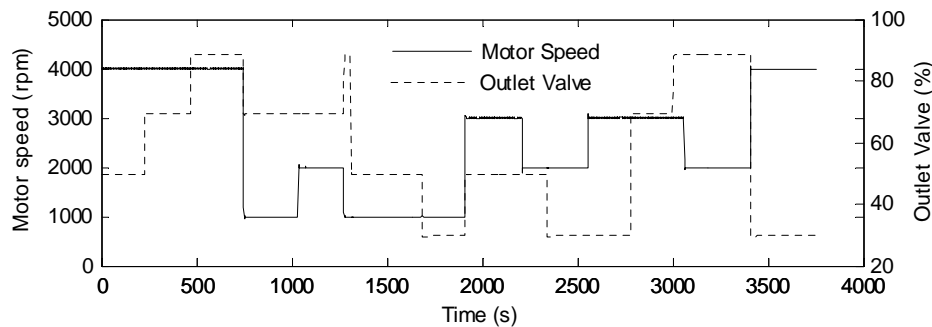
Fig. 131: Operating conditions for simulation of mechanical faults

For the simulation the vibrations caused by mechanical faults were modelled as the response of a one degree of freedom system where a mass m is connected to the foundation by an elastic element with stiffness K and a viscous damper with damping D. The mass m represents the mass of the electric motor or the compressor (depending for which element the vibrations are being simulated) and its value is 105kg and 140kg respectively. The parameters K and D are difficult to obtain unless a specific study is carried out. For this investigation the value of these parameters was tuned used typical values based on informed literature [153; 160]. The selected values for K and D were 10^8 N/m and 1000Ns/m respectively.

Rotor unbalance

The objective of this test is to simulate the vibrational response generated by rotor unbalance in the motor where the eccentricity grows linearly from perfect condition to a maximum of 0.5 mm. The residual force $\Delta F(t)$ generated by the simulated fault was estimated using (8-1). The obtained system response in terms of acceleration was added to the vibration measurement taken from the

232

motor in the vertical direction and then processed for CVA analysis. The parameters used for the simulation are represented in Table 38:

Table 38: Parameters for simulation of rotor unbalance

Parameter	Value
K	10^8 N/m
C	1000 Ns/m
m	105 Kg
m_r	25 Kg
e	0-0.5 mm

Fig. 132 shows the force simulated during the test Fig. 132 (a) and a detail of its sinusoidal nature Fig. 132 (b).
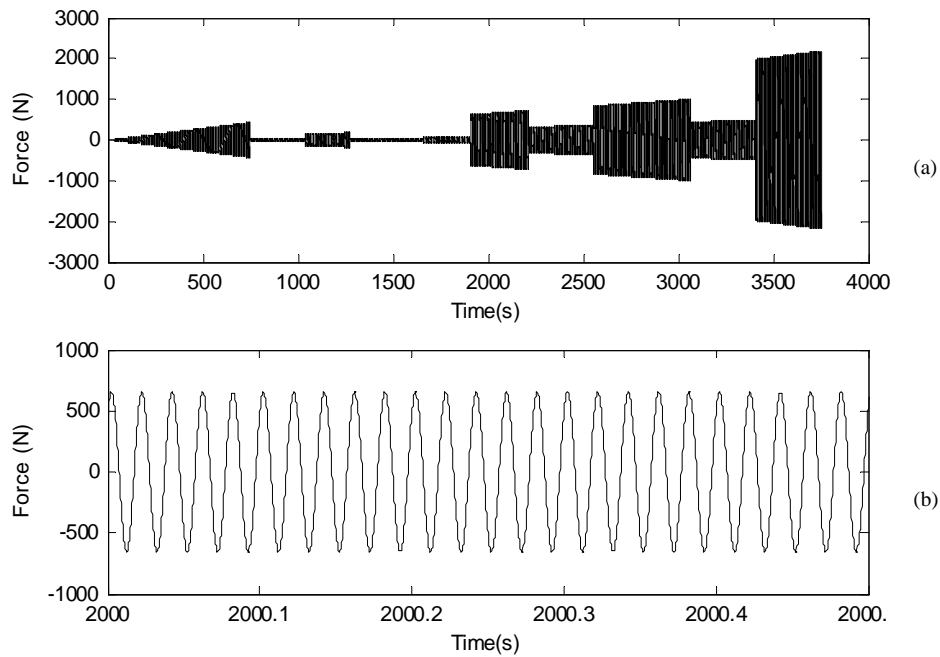
Fig. 132: Unbalance force (a) and detail (b)

Fig. 133 represents the generated system response Fig. 133 (a) and a detail of the signal waveform Fig. 133 (b). It can be observed that there are transients of very high amplitude corresponding to motor speed step changes.
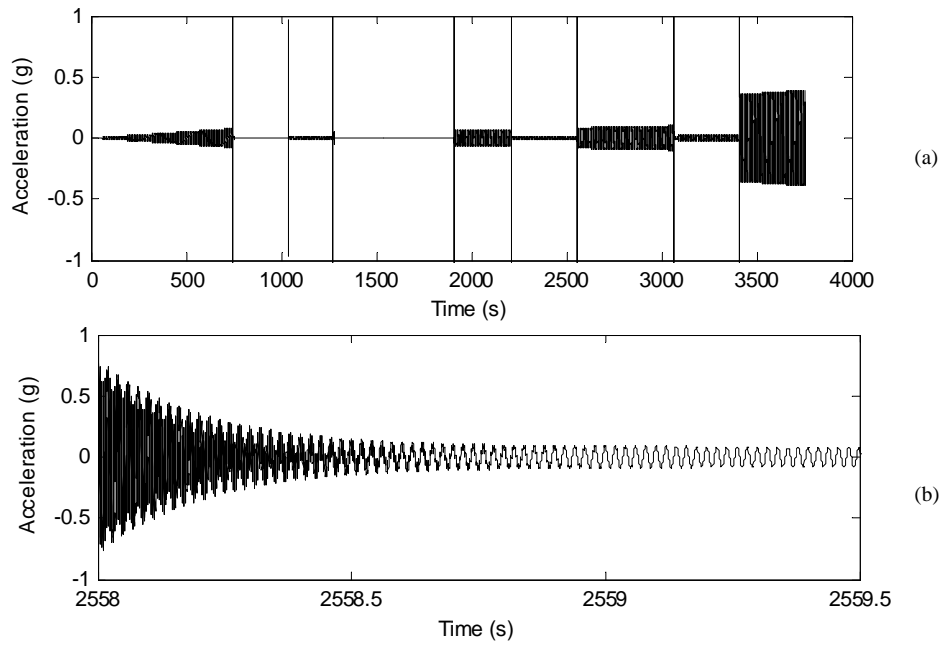
Fig. 133: System response to unbalance (a) and waveform detail (b)

Fig. 134 represents the original vibration signal measured from the motor in vertical direction (a) and the combined signal after adding the simulated fault (b):
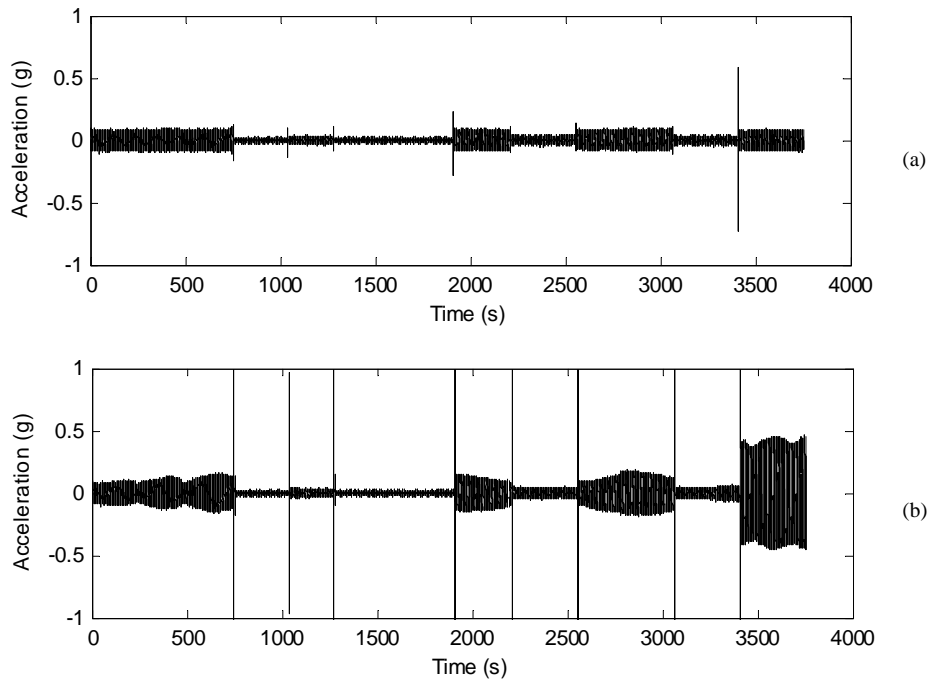
Fig. 134: Original motor vibration signal (a) and signal including seeded fault (b)

The effect of the seeded fault in the frequency spectrum of the vibration signal in the last stages of degradation (maximum unbalance) can be seen in Fig. 135.
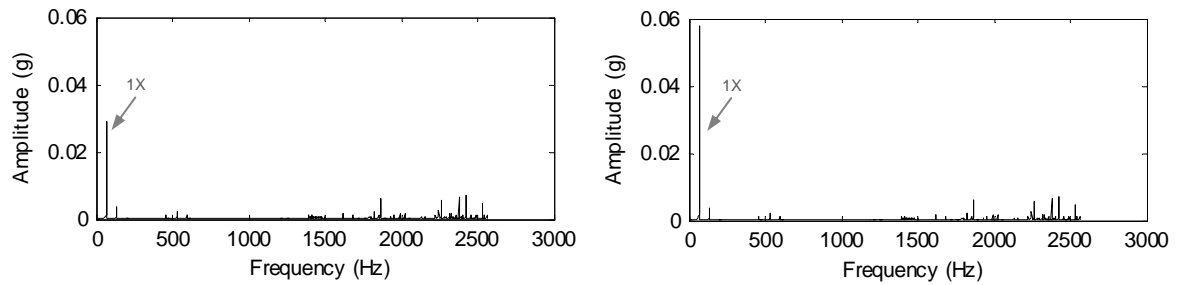
Fig. 135: Spectrum of the vibration signal before (left) and after (right) the introduction of unbalance

Fig. 136 represents the results obtained in terms of fault detection (a) and diagnosis (b) after the application of CVA to the data set. For this data set the first fault detection happened at sample 305 for the T^2 statistical indicator after one short false alarm and at sample 75 for the Q indicator. These samples correspond to an eccentricity of 0.04mm and 0.01mm respectively. The value of both indicators fluctuates above and below the UCL depending on the operational conditions. The cause of this is the dependency of the fault severity with the motor speed, as can be seen by the shape of the simulated residual load in Fig. 132 (a). The curved shape of both indicators is attributed to differences in phase between the simulated vibration and the main component at 1X in the real vibration, caused by slight variations in the real rotating speed. This change in phase causes an effect of increase or attenuation of the vibration depending on whether or not the signals are in phase. This oscillation can also be observed in Fig. 134 (b). The variables contributing more to the final value of both indicators are the amplitude of the peak at 1X in the frequency spectrum of the vibration signal measured in the motor and the RMS value of the same signal, which shows the sensitivity of the method to changes in the vibration frequency spectrum. Fig. 137 shows the evolution of the amplitude of the 1X peak in the spectrum of the motor vibration. It can be seen how the differences are minimum for low motor speeds (sample 762 to 1926) and significantly higher as the motor speed increases. The effect of

235

increment/reduction in the modified signal due to the differences in phase is also visible.

(a)

(b)
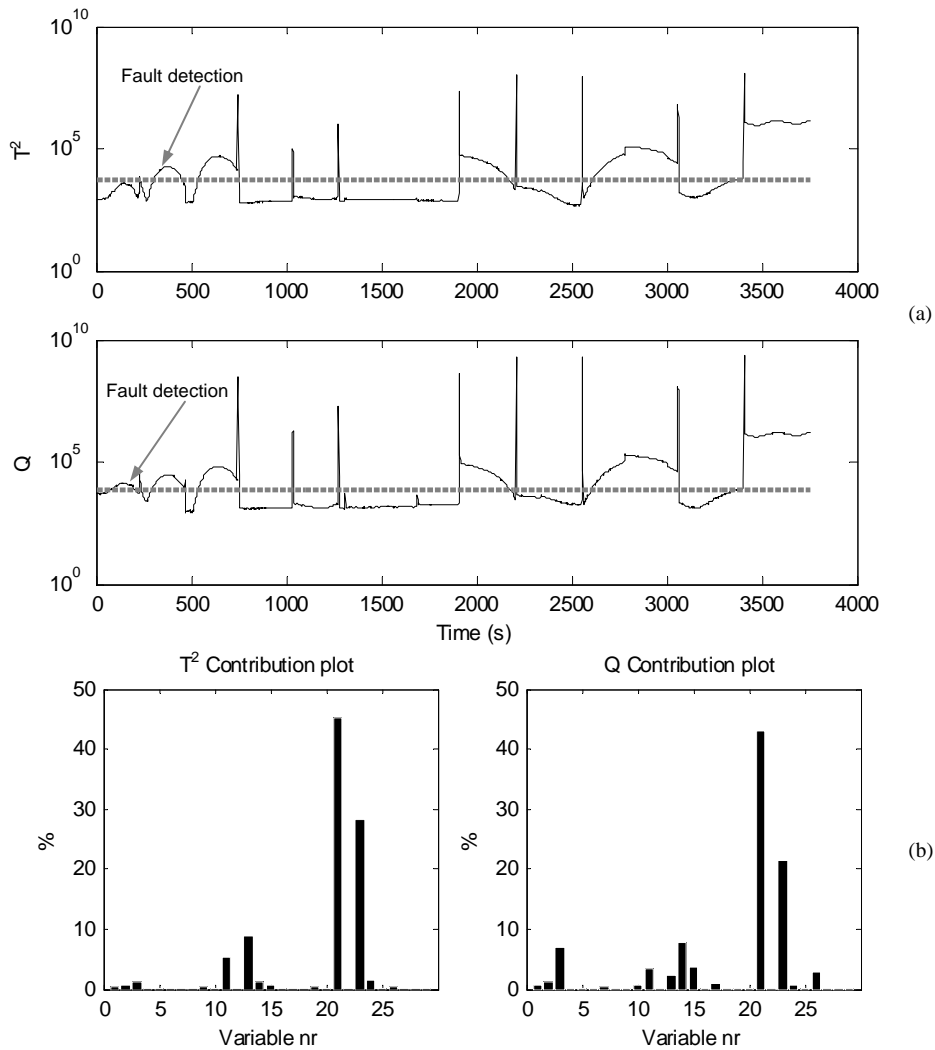
Fig. 136: Results for unbalance: T^2 and Q indicators (a) and contribution plots at sample 75(b)

Fig. 137: Evolution of 1X peak amplitude in motor vibration spectrum

Shaft misalignment

The objective of this test is to simulate the vibrational response generated by misalignment between the compressor and motor shafts. The residual force $\Delta F(t)$ was simplified for the analysis as a sinusoidal force with an oscillating frequency of twice the rotating speed (2X). The amplitude of the sinusoidal force grew linearly during the test from 0 to 100N at the end of the experiment simulating a linear increment in the misalignment. This approach neglects the dependence of the force generated by misalignment with the load but covers a wide spectrum of misalignment forces. The obtained system response in terms of acceleration was simulated for these loading conditions in the motor and the compressor, and added to the corresponding vibration measurement to be processed using CVA analysis. The parameters used for the simulation are represented in Table 39:

Table 39: Parameters for simulation of misalignment

Parameter	Value
K	10^8 N/m
C	1000 Ns/m
m *(motor)*	105 Kg
m *(compressor)*	140Kg
ΔF	0-100N

Fig. 138 shows the force simulated during the test Fig. 138 (a) and a detail of its sinusoidal nature Fig. 138 (b).
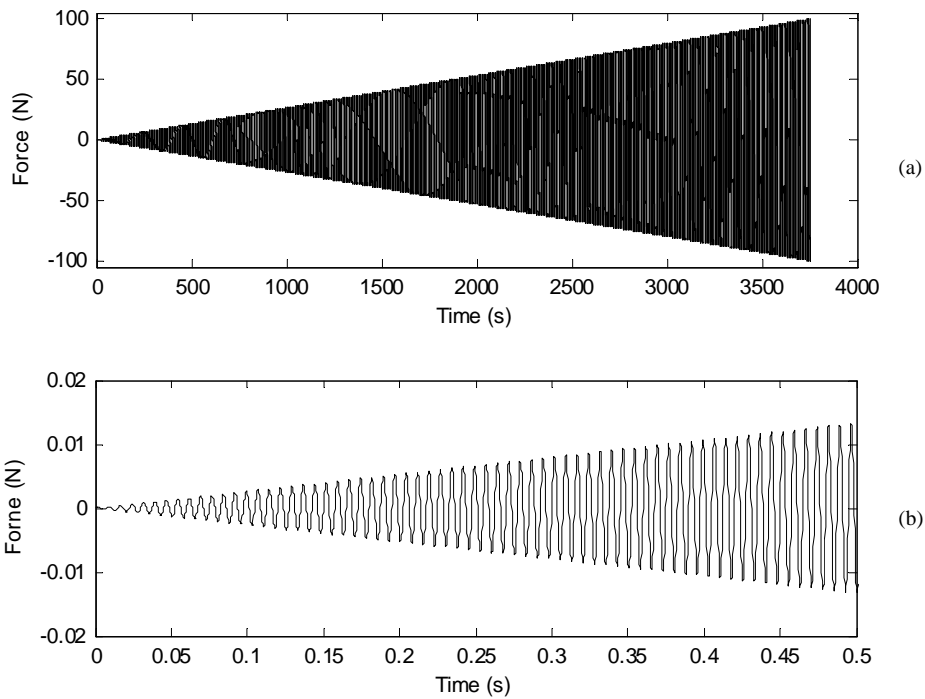
Fig. 138: Misalignment force (a) and detail (b)

Fig. 139 represents the generated system response in the motor Fig. 139 (a) and the compressor Fig. 139 (b). It can be observed that there are transients of very high amplitude corresponding to motor speed step changes.
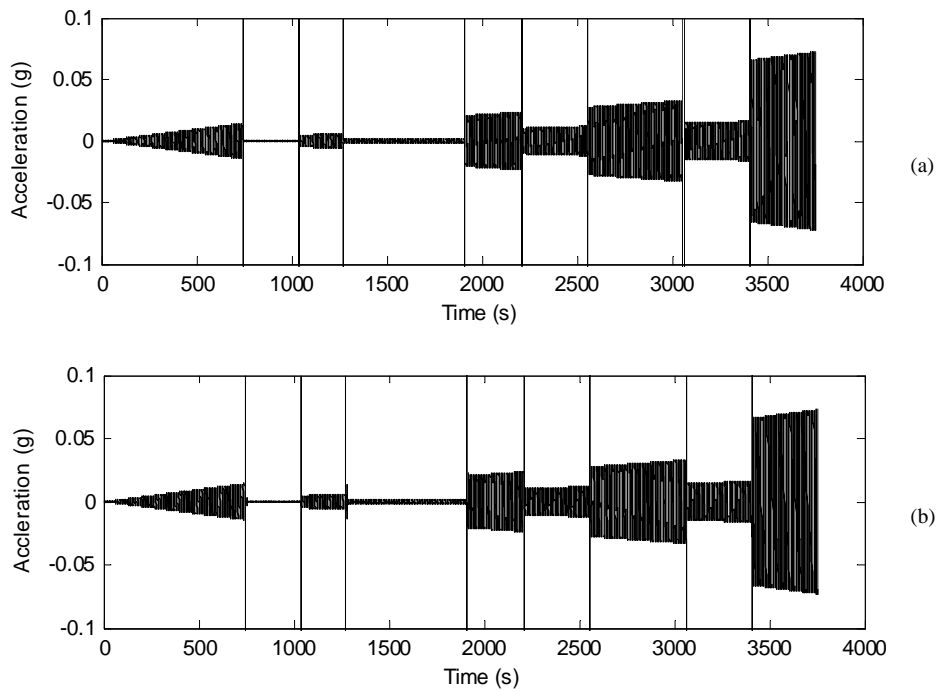
Fig. 139: System response to misalignment (a) and waveform detail (b)

Fig. 140 represents the original vibration signal spectrum measured from the motor in vertical direction at the end of the test Fig. 140 (a) and the spectrum of the combined signal after adding the simulated fault Fig. 140 (b). It is noticeable the increase in the amplitude of the 2X peak:
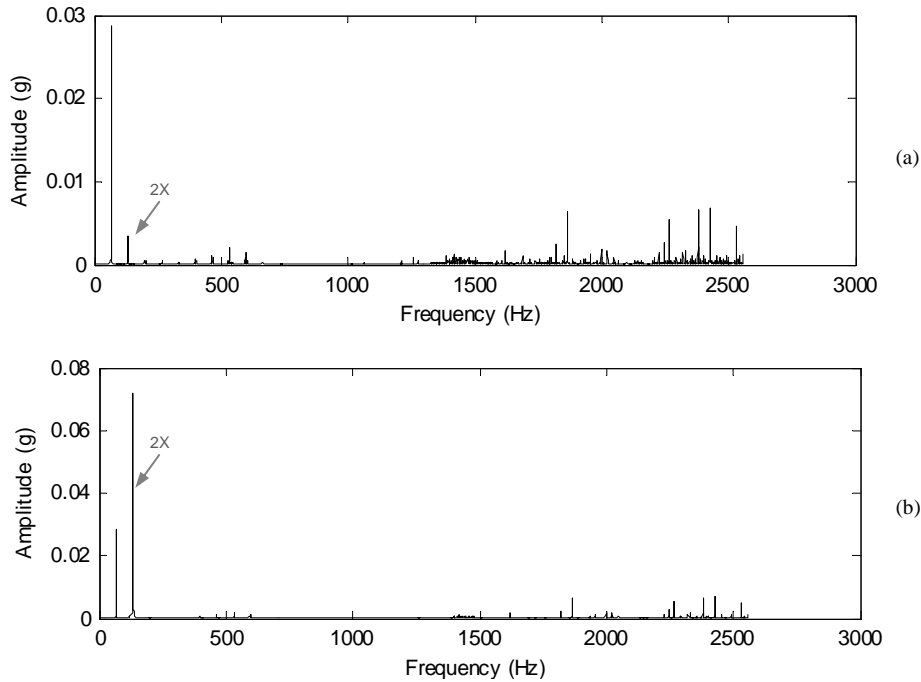
Fig. 140: Vibration frequency spectrum observed in the motor before (a) and after seeding the simulated fault (b)

Fig. 141 represents the original vibration signal spectrum measured from the compressor in vertical direction at the end of the test Fig. 141 (a) and the spectrum of the combined signal after adding the simulated fault Fig. 141 (b). It is noticeable the increment in the amplitude of the 2X peak:
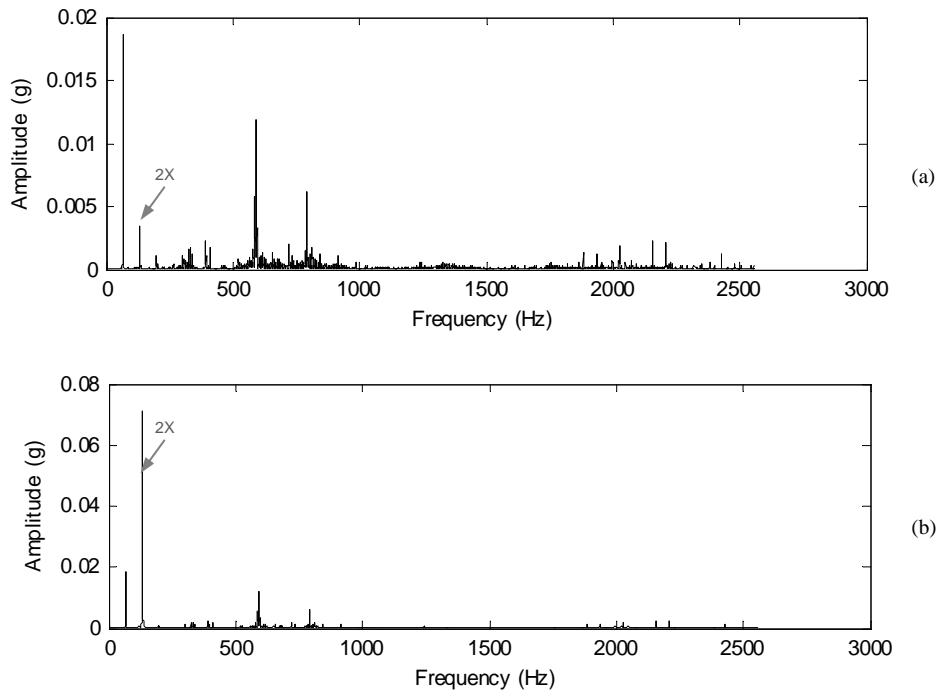
Fig. 141: Vibration frequency spectrum observed in the compressor before (a) and after seeding the simulated fault (b)

Fig. 142 represents the results obtained in terms of fault detection (a) and diagnosis (b) after the application of CVA to the data set. For this data set the first fault detection occurred at sample 1908 for both statistical indicators after three and five short false alarms respectively. This sample corresponds to a force amplitude of 50.48N. After the detection point the value of both indicators fluctuates above and below the UCL depending on the operational conditions. The cause is the dependency of the fault severity with the motor speed, as it can be seen in the shape of the simulated residual load in Fig. 139. Again there are oscillations in the indicators caused by the difference in phase between the simulated signal and the original component at 2X, but in this case the differences are reduced due to the low amplitude of the original 2X peak. The variables contributing more to the final value of the T^2 are the amplitude of the peak at 2X in the frequency spectrum of the vibration signal measured in the compressor (variable 19) and the motor (variable 22), the peak amplitude at 1X for both signals (variables 18 and 21), the RMS of the motor vibration signal (variable 23) and the outlet tank pressure (variable 14). In the case of the Q indicator the same variables show a high level of significance but with a

240

different rank. Fig. 143 shows the evolution of the amplitude of the 2X peak in the spectrum of the compressor vibration.
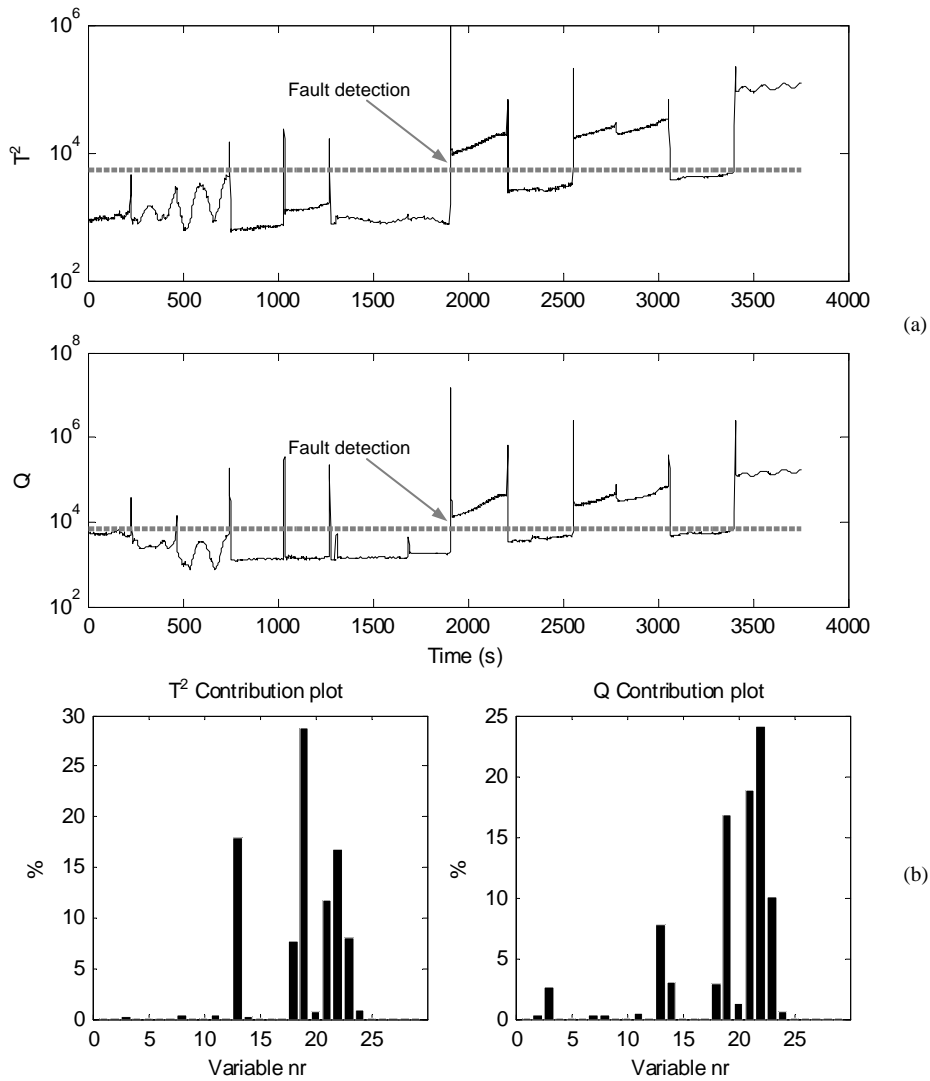
Fig. 142: Results for misalignment: T^2 and Q indicators (a) and contribution plots at sample 1908(b)
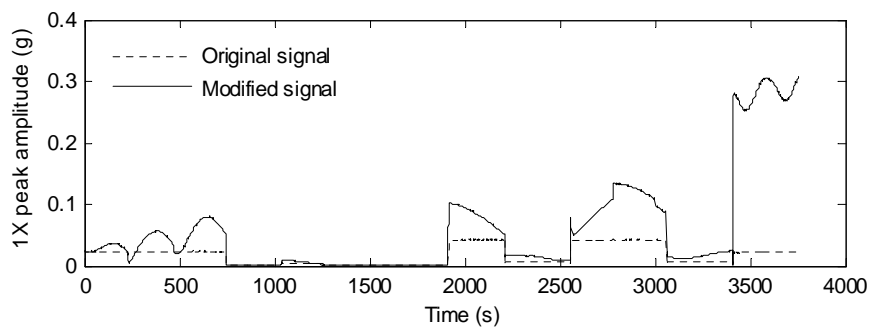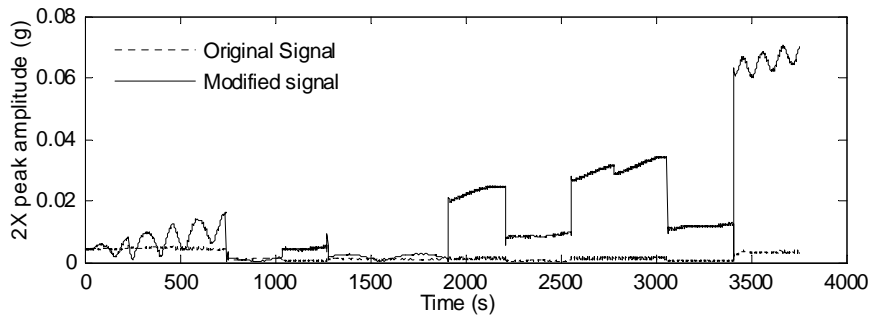

Fig. 143: Evolution of 2X peak amplitude in compressor vibration spectrum

Bearing fault

The objective of this test is to simulate the vibrational response generated by an outer race bearing fault in the driven end of the motor. The fault was simulated as a series of impulses acting on the virtual system, following a similar approach as [152]. The obtained system response in terms of acceleration was added to the vibration measurement taken from the motor in the vertical direction and then processed for CVA analysis. The main bearing characteristics and the parameters used for the simulation are represented in Table 40 and Table 41:

Table 40: Main bearing characteristics

Manufacturer/model	SKF 6309-2Z
Inner diameter	45mm
Outer diameter	100mm
Pitch diameter	72.492mm
Ball diameter	17.462mm
Number of rolling elements	8
Contact angle	0°
Outer race defect frequency	3.036X
Cage rotating frequency	0.38X

Table 41: Parameters for simulation of bearing outer race defect

Parameter	Value
K	10^8 N/m
C	1000 Ns/m
m	105 Kg
Impulse frequency	3.036X
Impulse duration	0.000115s-0.0023s
Impulse amplitude	0-10N

Fig. 144 shows the force simulated during the test Fig. 144 (a) and a detail of the impulse shape at the beginning (b) and end of the test (c)
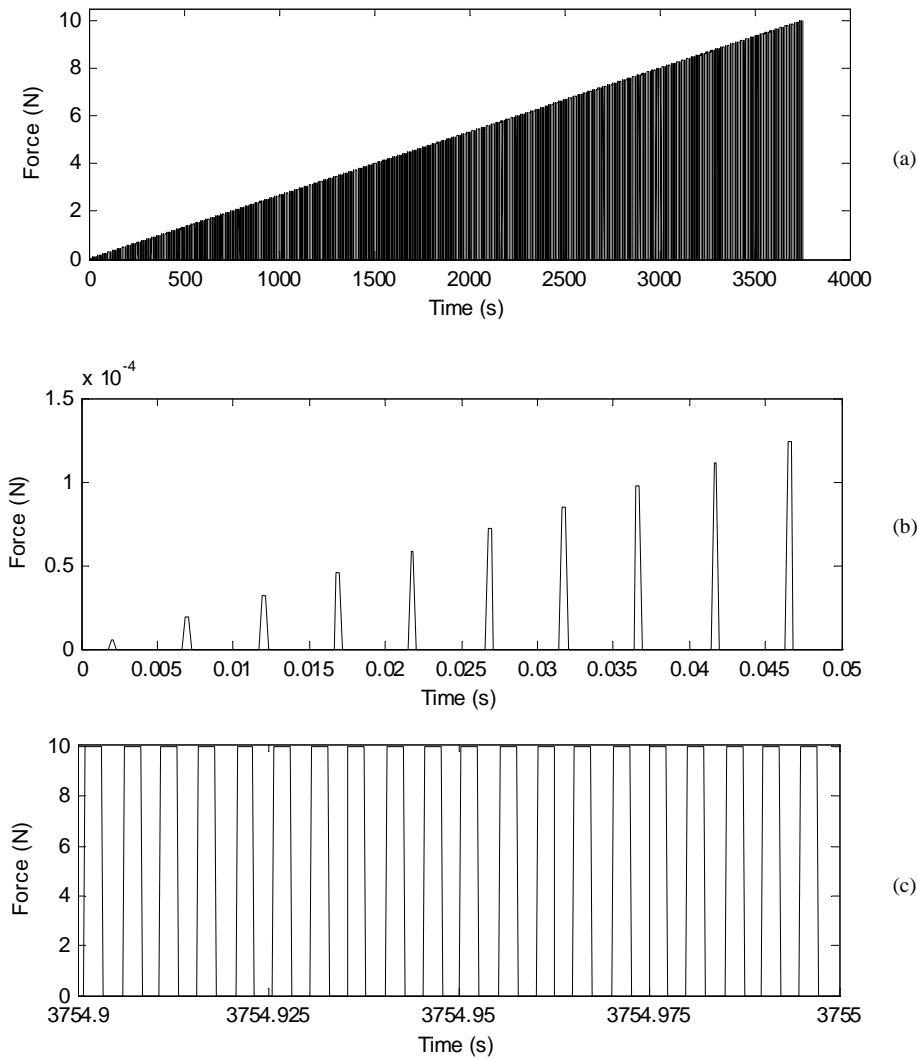
Fig. 144: Bearing fault force (a) and impulse detail at the beginning (b) and end (c) of the experiment

The impulse frequency was set to 3.036X in order to simulate a fault in the outer ring. The impact duration grew linearly during the test from 0.1157ms at the beginning and 2.3 ms at the end, which corresponds to a defect size between 0.5mm and 10mm if the rotor speed is 3000rpm. Fig. 145 represents the generated system response Fig. 145 (a) and a detail of the signal waveform Fig. 145 (b).
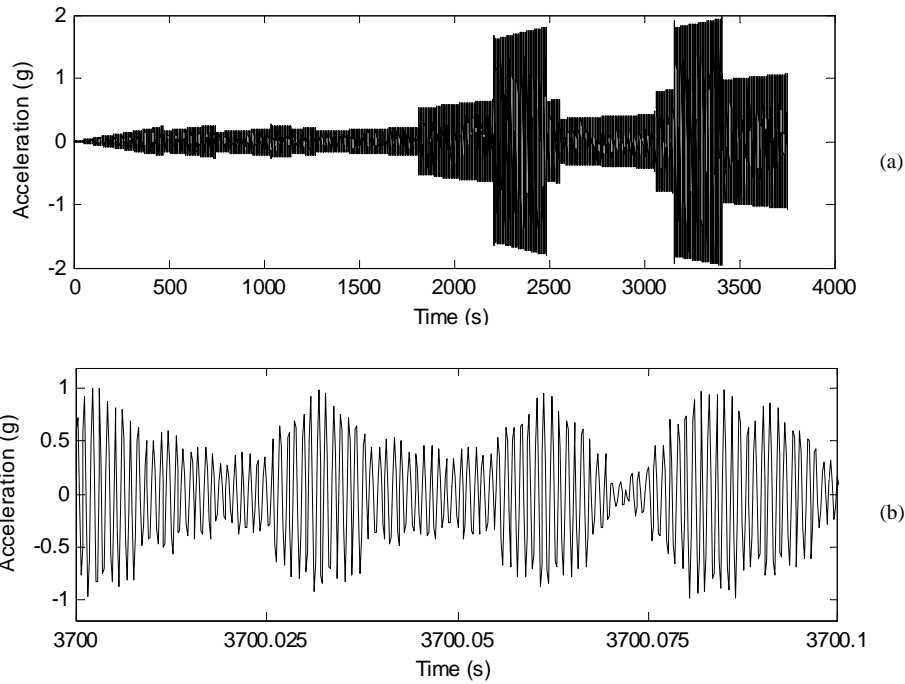
Fig. 145: System response to bearing fault (a) and waveform detail (b)

Fig. 146 represents the combined vibration signal observed in the motor vibration in vertical direction after adding the simulated fault:
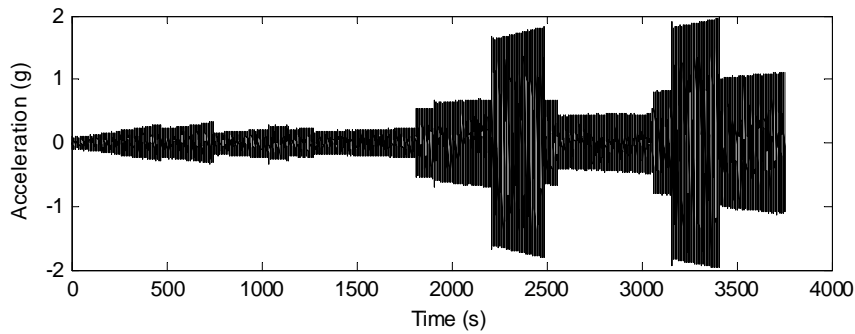
Fig. 146: System response to bearing fault

The effect of the seeded fault in the frequency spectrum of the vibration signal in the last stages of degradation (maximum unbalance) can be seen in Fig. 147.
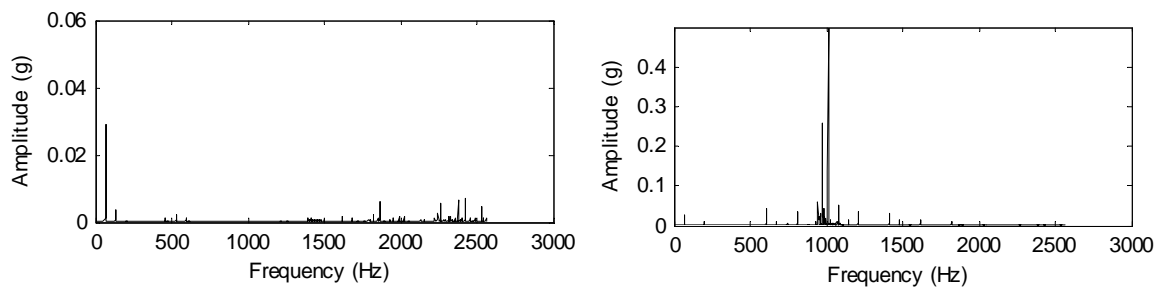
Fig. 147: Spectrum of the vibration signal before (left) and after (right) the introduction of bearing fault

The most significant changes observed after seeding the simulated fault are the activity increment around the virtual system natural frequency (975Hz) due to the resonant response to the impacts, and the appearance of a new peak at 3.036X. Fig. 148 represents the results obtained in terms of fault detection (a) and diagnosis (b) after the application of CVA to the data set.

For this data set the first fault detection happened at sample 164 for the T^2 statistical indicator and at sample 78 for the Q indicator. These samples correspond to an impact duration of 0.211ms and 0.1611ms, and an impact amplitude of 0.43N and 0.2N. The variable contributing most to the final value of both indicators is the RMS value of the motor vibration signal (variable 23). Fig. 149 shows the evolution of the RMS of the motor vibration, where it can be seen how the RMS value increases from the start of the test due to the introduced impulsive force, contributing to the early detection of the fault.

245
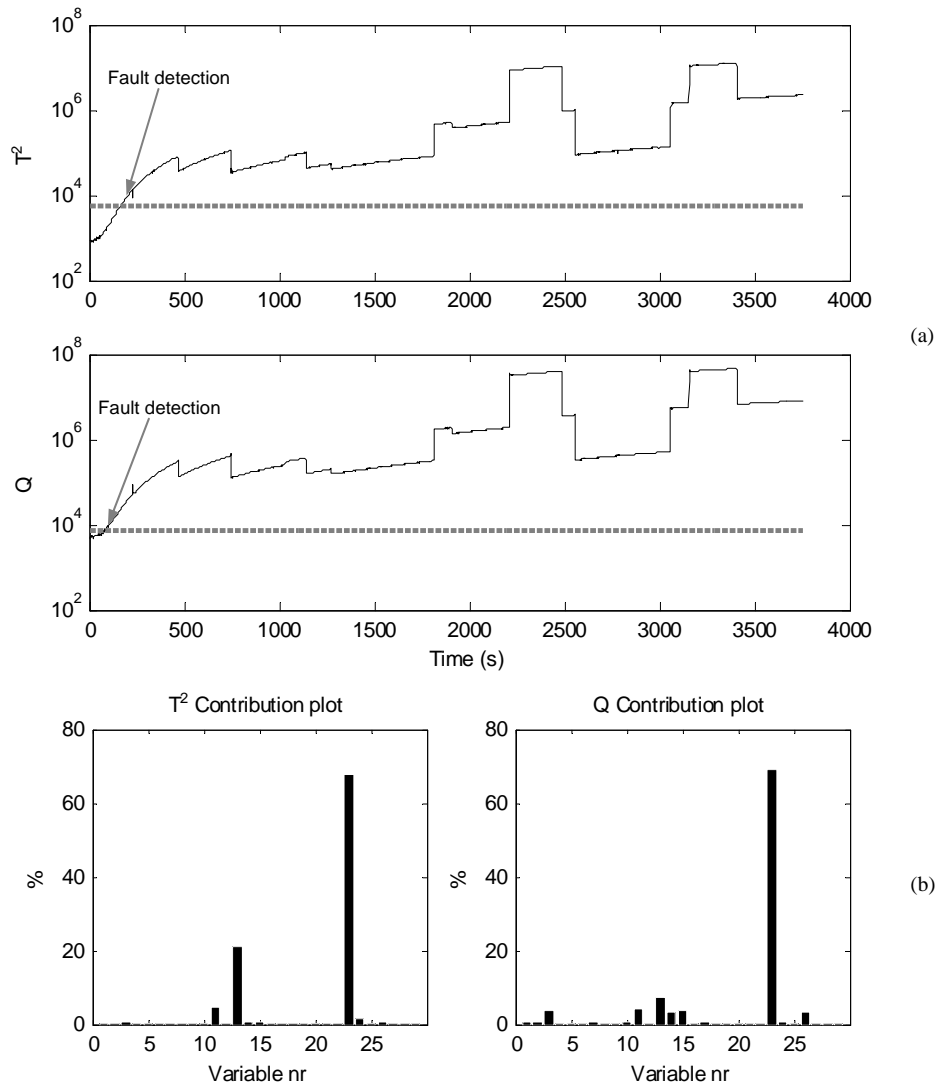
Fig. 148: Results for bearing fault: T^2 and Q indicators (a) and contribution plots at sample 78(b)
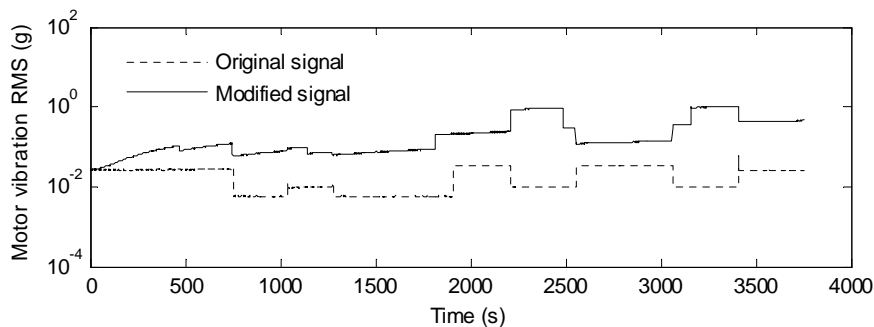
Fig. 149: Evolution of RMS in motor vibration spectrum

246

8.4 Conclusion

This study proposes a simple method to merge process data and vibration features to improve the detectability of mechanical faults in systems working under varying operational conditions. The proposed method was tested using experimental data acquired from a compressor test rig where data sets were acquired from the system working under different faulty conditions and varying operational points. Additionally mechanical faults were numerically simulated and added to vibration measurements.

The results show how CVA is able to detect changes in vibration features during the test, and detect deviations from normal operation even under operational conditions that were not considered in the training period. In addition contribution plots obtained at the time of fault detection provided valuable information about the variables most affected by the fault, which can help in root cause analysis. The analyses were carried out in a fast and robust manner, as the CVA is a well-known efficient and effective method and the proposed approach just requires the computation of short Fourier transforms and extraction of the desired features.

9 Conclusions

This chapter summarizes how the objectives stated in 1.2 have been fulfilled as presented in chapters 3 to 8, highlighting the most significant contributions.

The literature review identified the lack of a common experimental case study that can be used for the comparative assessment of different process monitoring techniques. For this reason the development of such benchmark case study was the first target of the investigation. As a result, a case study composed of experimental data acquired from a large scale test rig with different process faults were introduced was presented in chapter 3. The characteristics of this case study included some of the main challenges to be faced in the field of process monitoring, such as multivariate data, process nonlinearities and varying operational conditions. The data obtained for the case study as well as the guidelines presented in chapter 3 will be made publicly available, allowing its use by other researches for the assessment of new approaches for process monitoring with real data.

This benchmark case study was used in chapter 4 to compare the capabilities of three currently used algorithms for process monitoring (PCA, PLS and CVA) in terms of fault detection and diagnosis using real data. Another objective of this analysis was the assessment of the benefits of using dynamic approaches (DPCA and DPLS) and KDE for the calculation of health indicator thresholds. The main conclusion of the data analysis was that the combination of CVA and KDE produced better results than any other tested method, which corroborates the results provided by other researchers using computer simulated data. Additionally this study proved the applicability and robustness of the different methods tested when applied to real data acquired from a complex system.

Another research opportunity identified in the literature review was the need of developing algorithms that can evaluate performance degradation caused by faults, so that this information can be taken into account when planning optimal production and maintenance schedules. One example of this degradation can be the effect of pipe blockages or accumulation of dirt in compressor blades,

which increases the energy consumed by pumps and compressors and can affect the product quality if the flow rate requirements are not met. The application of CVA for identification of a real complex system in order to provide information about its performance degradation was studied in chapter 5. The results showed how the algorithm can be trained using data collected from the system under normal operating conditions to build a model that estimates accurately process measurements for different operating conditions. The differences between the estimations assuming normal operation and actual process measurements can be used to obtain an estimation of the degradation in the system performance. In addition, the algorithm was trained using data collected during the early stages of degradation after the introduction of faults in order to build a model that can predict the behaviour of the faulty system for different operational conditions. These approaches allow the system operators to forecast the losses caused by performance degradation and the capabilities of the faulty system to fulfil safety and quality requirements, and use this information to plan the appropriate production and maintenance schedules.

The potential benefits of merging different types of data to obtain more robust and reliable condition monitoring systems have been reported by several authors. Due to the popularity of vibration-based condition monitoring methods, the opportunity of merging features extracted from vibration signals and process data typically used in process monitoring was studied in this investigation. In a first instance, the capabilities of three signal processing techniques (LP, SANC and SK) to enhance fault bearing features in the vibration signal were investigated in chapter 6. The analysis showed that when high noise level is present in the signal it can be difficult to identify bearing fault features in the frequency spectrum in the early stages of degradation due to their relative low intensity. However, using algorithms for noise reduction and signal enhancement it is possible to unmask the desired components and detect the faults earlier. The objective of the analysis was to compare the performance of the algorithms tested using real data from a bearing test rig were natural degradation occurred. The result was that using Spectral Kurtosis it was

possible to detect and identify the fault much earlier than with any other technique.

Due to the benefits of Spectral Kurtosis and the popularity of bearing diagnostic techniques based on Acoustic Emission (AE), this technique was applied on acoustic emission data acquired from a bearing test rig were faults of different size were seeded. This analysis was presented in chapter 7 with the objective of demonstrating the sensitivity of the method to faults of different size. The results obtained showed that it is possible to improve the signal to noise ratio in AE signals with Spectral Kurtosis, especially when the defect is small and difficult to observe in the raw signal.

Once the capabilities of vibration-based condition monitoring for rotating machinery were reviewed and tested, a procedure for the combination of process and vibration data for CVA application was presented in chapter 8. Experimental data acquired from a compressor test rig which included process, vibration, and electrical measurements were used for the investigation. The compressor was operated at different speeds and pressure ratios during the tests to ensure the detection capabilities of the method in systems working under varying operational conditions. Various process and mechanical faults were simulated to provide a realistic test condition for the proposed algorithm. The results obtained showed that it is possible to improve the detectability of certain types of faults using CVA when different types of data are combined.

The contribution to knowledge of the research work developed for this thesis can be summarized in three key points:

- The capabilities of currently used multivariate algorithms for detection and diagnosis of process faults using real process data were tested and compared. The data used was acquired from an experimental test rig, generating a case study which included varying operational conditions and system nonlinearities. CVA resulted to be the most effective algorithm tested.

- System identification though CVA was applied to experimental data acquired from a large experimental facility to produce a model that can estimate performance degradation and predict the behaviour of a faulty system working under varying operational conditions.

- The fault detection capabilities of CVA were improved by merging process and vibration data extracting key features from the vibration signal spectrum. This approach makes CVA more sensitive to mechanical faults and any other type of fault that manifests as a fast oscillation in any of the variables measured.

These achievements are completely in line with the objectives of the Energy-Smartops project, which aims to find solutions to improve the efficiency of large complex processes through improved maintenance and operation, taking into consideration the connections between the different system interfaces.

As a result of the innovative work carried out in this investigation the following papers have been submitted for publication in specialized journals:

- "Use of Spectral Kurtosis for improving signal to noise ratio of acoustic emission signal from defective bearings" (Journal of Failure Analysis and Prevention, published June 2014)

- "Application of Linear Prediction, Self-Adaptive Noise Cancellation and Spectral Kurtosis in Identifying Natural Damage of a Rolling Element Bearing in a Gearbox" (International Journal of Acoustics and Vibration, accepted on May 2014)

- "Statistical process monitoring of a multiphase flow facility, part I: The benchmark case" (Control Engineering Practice, submitted on March 2014)

- "Statistical process monitoring of a multiphase flow facility, part II: Application of canonical variate analysis for fault detection and diagnosis" (Control Engineering Practice, submitted on March 2014)

- "Combination of process and vibration data for improved condition monitoring of industrial systems working under variable operating conditions" (Mechanical systems and signal processing, submitted on May 2014)

- "Estimation of process performance degradation under faulty conditions using canonical variate analysis" (Signal Processing, submitted on July 2014)

Additionally this work has been presented in the following international conferences:

- "Application of linear prediction, self-adaptive noise cancellation and spectral kurtosis in identifying natural damage of rolling element bearing in a gearbox" (World Congress on Engineering Asset Management, Daejeon (Korea) September 2012)

- "A benchmark application of canonical variate analysis for fault detection and diagnosis" (International Conference on Control, Loughborough (UK) July 2014)

- "Application of canonical variate analysis for fault detection and performance control" (World Congress on Intelligent Control and Automation, Shenyang (China) July 2014)

- "Application of Spectral Kurtosis on acoustic emission signals for the detection of faults in rolling element bearings" (Congress on Condition Monitoring and Diagnostic Engineering Management, Brisbane (Australia) September 2014)

- "Improved condition monitoring using fast-oscillating measurements" (International Conference on Automation and Computing, Cranfield (UK) September 2014)

The recommendations for the continuation of this research in future work are:

- To explore the applicability in real processes of nonlinear extensions of multivariate algorithms for condition monitoring, particularly those based on nonlinear transformations of the input data using kernel methods and assess their performance.

- To extend the use of CVA for prediction of system behaviour under faulty conditions to cases where the fault evolves over time, making use of time-varying models for system identification and assess their capabilities using real process data.

- To collect experimental data including process and vibration data from a rig where mechanical faults can be physically seeded in the system to assess the benefits of combining process and vibration data in multivariate algorithms in a real (not simulated) case study.

REFERENCES

[1]   Randall, R. B. (2011), *Vibration-based condition monitoring,* 1st ed, John Wiley & Sons, Singapore.

[2]   L.H. Chiang and E.L. Rusell and R.D. Braatz (2000), *Fault detection and diagnosis in industrial systems,* 1st ed, Springer, London UK.

[3]   Chen, R. Q. (2013), "Advances in data-driven monitoring methods for complex process", *3rd International Conference on Applied Mechanics, Materials and Manufacturing, ICAMMM 2013;* Vol. 423-426, 24 August 2013, Dalian; China, pp. 2448.

[4]   Ge, Z., Song, Z. and Gao, F. (2013), "Review of recent research on data-based process monitoring", *Industrial and Engineering Chemistry Research,* vol. 52, no. 10, pp. 3543-3562.

[5]   Commission of the European Communities (2006)*, A European Strategy for Sustainable, Competitive and Secure Energy*, available at: http://ec.europa.eu/energy/green-paper-energy/doc/2006_03_08_gp_document_en.pdf (accessed May 2012).

[6]   Zhang, Q. (2013), "Case study of cost benefits of condition based maintenance used in medical devices", *Proceedings - Annual Reliability and Maintainability Symposium,* January 2013, Orlando, FL; United States.

[7]   Rajan, B. S. and Roylance, B. J. (2000), "Condition-based maintenance: A systematic method for counting the cost and assessing the benefits", *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical,* vol. 214, no. 2, pp. 97-108.

[8]   van der Weide, J. A. M., Pandey, M. D. and van Noortwijk, J. M. (2010), "Discounted cost model for condition-based maintenance optimization", *Reliability Engineering and System Safety,* vol. 95, no. 3, pp. 236-246.

[9]   Golmakani, H. R. (2011), "Cost-effective condition-based inspection scheme for condition-based maintenance", *Proceedings of the 2011 IEEE International Conference on Information Reuse and Integration, IRI 2011,* pp. 327.

[10]   Wang, P. and Wang, Y. (2013), "Cost benefit analysis of condition monitoring systems for optimal maintenance decision making", *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, .*

[11]   Amari, S. V., McLaughlin, L. and Pham, H. (2006), "Cost-effective condition-based maintenance using Markov decision processes", *Proceedings - Annual Reliability and Maintainability Symposium,* pp. 464.

[12]   Maillart, L. M. and Pollock, S. M. (2002), "Cost-optimal condition-monitoring for predictive maintenance of 2-phase systems", *IEEE Transactions on Reliability,* vol. 51, no. 3, pp. 322-330.

[13]   Verma, N. K. and Subramanian, T. S. S. (2012), "Cost benefit analysis of intelligent condition based maintenance of rotating machinery", *Proceedings of the 2012 7th IEEE Conference on Industrial Electronics and Applications, ICIEA 2012,* pp. 1390.

[14]   Al-Najjar, B. (2012), "On establishing cost-effective condition-based maintenance: Exemplified for vibration-based maintenance in case companies", *Journal of Quality in Maintenance Engineering,* vol. 18, no. 4, pp. 401-416.

[15]   Ku, W., Storer, R. H. and Georgakis, C. (1995), "Disturbance detection and isolation by dynamic principal component analysis", *Chemometrics and Intelligent Laboratory Systems,* vol. 30, no. 1, pp. 179-196.

[16]   Komulainen, T., Sourander, M. and Jämsä-Jounela, S. -. (2004), "An online application of dynamic PLS to a dearomatization process", *Computers and Chemical Engineering,* vol. 28, no. 12, pp. 2611-2619.

[17]   Odiowei, P. P. and Cao, Y. (2009), "Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations", *Computer Aided Chemical Engineering,* vol. 27, no. C, pp. 1557-1562.

[18]   Juricek, B. C., Seborg, D. E. and Larimore, W. E. (2004), "Fault Detection Using Canonical Variate Analysis", *Industrial and Engineering Chemistry Research,* vol. 43, no. 2, pp. 458-474.

[19]   Russell, E. L., Chiang, L. H. and Braatz, R. D. (2000), "Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis", *Chemometrics and Intelligent Laboratory Systems,* vol. 51, no. 1, pp. 81-93.

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

[21]   Schaper, C. D., Larimore, W. E., Seborg, D. E. and Mellichamp, D. A. (1994), "Identification of chemical processes using canonical variate analysis", *Computers and Chemical Engineering,* vol. 18, no. 1, pp. 55-69.

[22]   Yang, Y., Chen, Y., Chen, X. and Liu, X. (2012), "Multivariate industrial process monitoring based on the integration method of canonical variate analysis and independent component analysis", *Chemometrics and Intelligent Laboratory Systems,* vol. 116, pp. 94-101.

[23]   Choi, S. W., Park, J. H. and Lee, I. (2004), "Process monitoring using a Gaussian mixture model via principal component analysis and discriminant analysis", *Computers and Chemical Engineering,* vol. 28, no. 8, pp. 1377-1387.

[24]   Juricek, B. C., Seborg, D. E. and Larimore, W. E. (2004), "Fault Detection Using Canonical Variate Analysis", *Industrial and Engineering Chemistry Research,* vol. 43, no. 2, pp. 458-474.

[25]   Ma, L. , Wang, J. and Song, Y. (2004), "Process monitoring method based on multi-PCA models", *Beijing Ligong Daxue Xuebao/Transaction of Beijing Institute of Technology,* vol. 24, no. 1, pp. 64-68.

[26]   Qi, Y. , Wang, P., Fan, S. , Gao, X. and Jiang, J. (2009), "Enhanced batch process monitoring using kalman filter and multiway kernel principal component analysis", *2009 Chinese Control and Decision Conference, CCDC 2009,* pp. 5289.

[27]   Simoglou, A., Martin, E. B. and Morris, A. J. (2002), "Statistical performance monitoring of dynamic multivariate processes using state space modelling", *Computers and Chemical Engineering,* vol. 26, no. 6, pp. 909-920.

[28]   Yunus, M. Y. M. and Zhang, J. (2010), "Multivariate process monitoring using classical multidimensional scaling and procrustes analysis", *IFAC Proceedings Volumes (IFAC-PapersOnline),* Vol. 9, pp. 165.

[29]   Zhang, J., Martin, E. B. and Morris, A. J. (1997), "Process monitoring using non-linear statistical techniques", *Chemical Engineering Journal,* vol. 67, no. 3, pp. 181-189.

[30]   Zhou, K., Li, Q. and Guo, R. (2012), "Improving monitoring accuracy of process based on SPC method", *Proceedings of the 2012 24th Chinese Control and Decision Conference, CCDC 2012,* pp. 1488.

[31]   Ricker, N. L. *, Tennessee Eastman Challenge Archive*, available at: http://depts.washington.edu/control/LARRY/TE/download.html (accessed 21/03/2014).

[32]   Fan, J., Qin, S. J. and Wang, Y. (2014), "Online monitoring of nonlinear multivariate industrial processes using filtering KICA-PCA", *Control Engineering Practice,* vol. 22, no. 1, pp. 205-216.

[33]   Gao, Q. and Jia, Z. (2010), "An application study of the improved PCA method in the TE process fault detection", *Proceedings - 2010 IEEE International Conference on Intelligent Computing and Intelligent Systems, ICIS 2010,* Vol. 2, pp. 706.

[34]   Jiang, Q. and Yan, X. (2012), "Multivariate statistical process monitoring using modified factor analysis and its application", *Journal of Chemical Engineering of Japan,* vol. 45, no. 10, pp. 829-839.

[35]   Lau, C. K., Ghosh, K., Hussain, M. A. and Che Hassan, C. R. (2013), "Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS", *Chemometrics and Intelligent Laboratory Systems,* vol. 120, no. 0, pp. 1-14.

[36]   Ma, H., Hu, Y., Yan, X. and Shi, H. (2012), "An improved PLS (IPLS) method utilizing local standardization strategy for multimode process monitoring", *Journal of Donghua University (English Edition),* vol. 29, no. 4, pp. 288-294.

[37]   Mori, J. and Yu, J. (2014), "A quality relevant non-Gaussian latent subspace projection method for chemical process monitoring and fault detection", *AIChE Journal,* vol. 60, no. 2, pp. 485-499.

[38]   Ning, L., Zhang, Y. and Wen, D. (2010), "Industrial process fault detection based on wavelet denoising PCA", *4th International Symposium on Computational Intelligence and Industrial Applications, ISCIIA 2010,* pp. 134.

[39]   Odiowei, P. P. and Cao, Y. (2010), "Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations", *IEEE Transactions on Industrial Informatics,* vol. 6, no. 1, pp. 36-45.

[40]   Qing, Y., Feng, T. and Dazhi, W. (2010), "Nonlinear dynamic process monitoring based on lifting wavelets and dynamic kernel PCA", *Proceedings of the World Congress on Intelligent Control and Automation (WCICA),* pp. 5712.

[41]   Rato, T. J. and Reis, M. S. (2013), "Fault detection in the Tennessee Eastman benchmark process using dynamic principal components analysis based on decorrelated residuals (DPCA-DR)", *Chemometrics and Intelligent Laboratory Systems,* vol. 125, pp. 101-108.

[42]   Song, K., Wang, H. -. and Li, P. (2005), "PLS quality monitoring and its application for Tennessee Eastman process", *Zhejiang Daxue Xuebao (Gongxue Ban)/Journal of Zhejiang University (Engineering Science),* vol. 39, no. 5, pp. 657-662.

[43]   Stubbs, S., Zhang, J. and Morris, J. (2012), "Fault detection in dynamic processes using a simplified monitoring-specific CVA state space modelling approach", *Computers and Chemical Engineering,* vol. 41, pp. 77-87.

[44]   Wang, Q. and Li, H. (2012), "A process singular value recognition based recursive PCA approach to mode-transition process monitoring", *Huagong Xuebao/CIESC Journal,* vol. 63, no. 9, pp. 2948-2952.

[45]   Xiangrong Shi, Jun Liang, Lubin Ye and Bin Hu (2010), "A method of fault diagnosis based on PCA and Bayes classification", *Intelligent Control and Automation (WCICA), 2010 8th World Congress on,* pp. 5628.

[46]   Xu, X., Xie, L. and Wang, S. (2011), "Multi-mode process monitoring method based on PCA mixture model", *Huagong Xuebao/CIESC Journal,* vol. 62, no. 3, pp. 743-752.

[47]   Yang, Q., Tian, F. and Wang, D. (2010), "Online approach of fault diagnosis based on lifting wavelets and moving window PCA", *Proceedings of the World Congress on Intelligent Control and Automation (WCICA),* pp. 2909.

[48]   Yi, H., Hehe, M. and Hongbo, S. (2013), "Fault detection for chemical process based on robust PLS", *2013 25th Chinese Control and Decision Conference, CCDC 2013,* pp. 4947.

[49]   Yin, S., Ding, S. X., Zhang, P., Hagahni, A. and Naik, A. (2011), "Study on modifications of PLS approach for process monitoring", *IFAC Proceedings Volumes (IFAC-PapersOnline),* Vol. 18, pp. 12389.

[50]   Yu, J. (2010), "Hidden Markov models combining local and global information for nonlinear and multimodal process monitoring", *Journal of Process Control,* vol. 20, no. 3, pp. 344-359.

[51]   Liu, D., Jiang, D., Chen, X., Luo, A. and Xu, G. (2012), "Research on fault identification for complex system based on generalized linear canonical correlation analysis", *2012 IEEE International Conference on Automation Science and Engineering: Green Automation Toward a Sustainable Society, CASE 2012,* 20 August 2012 through 24 August 2012, Seoul, pp. 474.

[52]   Borsje, H. J. (1999), "Fault detection in boilers using canonical variate analysis", *Proceedings of the 1999 American Control Conference (99ACC),* Vol. 2, 2 June 1999 through 4 June 1999, San Diego, CA, USA, IEEE, Piscataway, NJ, United States, pp. 1167.

[53]   Stander, C. J., Heyns, P. S. and Schoombie, W. (2002), "Using vibration monitoring for local fault detection on gears operating under fluctuating load conditions", *Mechanical Systems and Signal Processing,* vol. 16, no. 6, pp. 1005-1024.

[54]   Zimroz, R., Bartelmus, W., Barszcz, T. and Urbanek, J. (2014), "Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings", *Mechanical Systems and Signal Processing,* vol. 46, no. 1, pp. 16-27.

[55]   Bartelmus, W. and Zimroz, R. (2009), "A new feature for monitoring the condition of gearboxes in non-stationary operating conditions", *Mechanical Systems and Signal Processing,* vol. 23, no. 5, pp. 1528-1534.

[56]   Braun, S. (2011), "The synchronous (time domain) average revisited", *Mechanical Systems and Signal Processing,* vol. 25, no. 4, pp. 1087-1102.

[57]   Zhan, Y., Makis, V. and Jardine, A. K. S. (2006), "Adaptive state detection of gearboxes under varying load conditions based on parametric modelling", *Mechanical Systems and Signal Processing,* vol. 20, no. 1, pp. 188-221.

[58]   Parker Jr., B. E., Ware, H. A., Wipf, D. P., Tompkins, W. R., Clark, B. R., Larson, E. C. and Poor, H. V. (2000), "Fault diagnostics using statistical change detection in the bispectral domain", *Mechanical Systems and Signal Processing,* vol. 14, no. 4, pp. 561-570.

[59]   McFadden, P. D. (1986), "Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration", *Journal of vibration, acoustics, stress, and reliability in design,* vol. 108, no. 2, pp. 165-170.

[60]   McFadden, P. D. (1988), "Determining the location of a fatigue crack in a gear from the phase of the change in the meshing vibration", *Mechanical Systems and Signal Processing,* vol. 2, no. 4, pp. 403-409.

[61]   Heng, A., Zhang, S., Tan, A. C. C. and Mathew, J. (2009), "Rotating machinery prognostics: State of the art, challenges and opportunities", *Mechanical Systems and Signal Processing,* vol. 23, no. 3, pp. 724-739.

[62]   Jardine, A. K. S., Lin, D. and Banjevic, D. (2006), "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mechanical Systems and Signal Processing,* vol. 20, no. 7, pp. 1483-1510.

[63]   Ji, Z., Zhang, X. and Wang, C. (2010), "Monitoring of continuous steel casting process based on independent component analysis", *2010 Chinese Control and Decision Conference, CCDC 2010,* pp. 3920.

[64]   Rotem, Y., Wachs, A. and Lewin, D. R. (2000), "Ethylene compressor monitoring using model-based PCA", *AIChE Journal,* vol. 46, no. 9, pp. 1825-1836.

[65]   Vanhatalo, E. (2010), "Multivariate process monitoring of an experimental blast furnace", *Quality and Reliability Engineering International,* vol. 26, no. 5, pp. 495-508.

[66]   Yoo, C. K., Lee, D. S. and Vanrolleghem, P. A. (2004), "Application of multiway ICA for on-line process monitoring of a sequencing batch reactor", *Water research,* vol. 38, no. 7, pp. 1715-1732.

[67]   Alkaya, A. and Eker, İ (2011), "Variance sensitive adaptive threshold-based PCA method for fault detection with experimental application", *ISA transactions,* vol. 50, no. 2, pp. 287-302.

[68]   Borsje, H. J. (1999), "Fault detection in boilers using canonical variate analysis", *Proceedings of the American Control Conference, 1999,* Vol. 2, pp. 1167.

[69]   Chen, G., Liang, J. and Qian, J. (2005), "Support vector classifier based PCA with application to process", *Yi Qi Yi Biao Xue Bao/Chinese Journal of Scientific Instrument,* vol. 26, no. 1, pp. 54-58.

[70]   Eserin, P. (1999), "Application of canonical variate analysis to the dynamical modeling and control of drum level in an industrial boiler", *Proceedings of the American Control Conference, 1999,* Vol. 2, pp. 1163.

[71]   Gunther, J. C., Conner, J. S. and Seborg, D. E. (2009), "Process monitoring and quality variable prediction utilizing PLS in industrial fed-batch cell culture", *Journal of Process Control,* vol. 19, no. 5, pp. 914-921.

[72]   Lee, H. W., Lee, M. W. and Park, J. M. (2009), "Multi-scale extension of PLS algorithm for advanced on-line process monitoring", *Chemometrics and Intelligent Laboratory Systems,* vol. 98, no. 2, pp. 201-212.

[73]   Liren Yan (2006), "A PCA-Based PCM Data Analyzing Method for Diagnosing Process Failures", *IEEE Transactions on Semiconductor Manufacturing,* vol. 19, no. 4, pp. 404-410.

[74]   Peng, K., Li, G. and Zhang, K. (2012), "Strip thickness monitoring in hot strip mill processes based on dynamic total projection to latent structures (T-PLS) algorithm", *Kongzhi Lilun Yu Yingyong/Control Theory and Applications,* vol. 29, no. 11, pp. 1446-1451.

[75]   Emerson Electric Co. *, The DeltaV Digital Automation System* , available at: http://www2.emersonprocess.com/en-us/brands/deltav/pages/index.aspx (accessed 13 February 2014).

[76]   Jansen, F. E., Shoham, O. and Taitel, Y. (1996), "The elimination of severe slugging - Experiments and modeling", *International Journal of Multiphase Flow,* vol. 22, no. 6, pp. 1055-1072.

[77]   Malekzadeh, R., Henkes, R. A. W. M. and Mudde, R. F. (2012), "Severe slugging in a long pipeline–riser system: Experiments and predictions", *International Journal of Multiphase Flow,* vol. 46, no. 0, pp. 9-21.

[78]   MacGregor, J. F. and Kourti, T. (1995), "Statistical process control of multivariate processes", *Control Engineering Practice,* vol. 3, no. 3, pp. 403-414.

[79]   Jackson, J. E. and Mudholkar, G. S. (1979), "Control procedures for residuals associated with principal component analysis", *Technometrics,* vol. 21, no. 3, pp. 341-349.

[80]   Deng, X. and Tian, X. (2011), "A new fault isolation method based on unified contribution plots", *Proceedings of the 30th Chinese Control Conference, CCC 2011,* pp. 4280.

[81]   Liu, J. and Chen, D. -. (2012), "Multiple Sensor Fault Isolation Using Contribution Plots without Smearing Effect to Non-Faulty Variables", *Computer Aided Chemical Engineering,* vol. 31, pp. 1517-1521.

[82]   Mnassri, B., El Adel, E. M., Ananou, B. and Ouladsine, M. (2009), "Fault detection and diagnosis based on PCA and a new contribution plots", *IFAC Proceedings Volumes (IFAC-PapersOnline),* pp. 834.

[83]   Ramírez, A. W. and Llinàs, J. C. (2011), "Fault diagnosis of batch processes release using PCA contribution plots as fault signatures", *ICEIS 2011 - Proceedings of the 13th International Conference on Enterprise Information Systems,* Vol. 1 DISI, pp. 223.

[84]   Negiz, A. and Çinar, A. (1998), "Monitoring of multivariable dynamic processes and sensor auditing", *Journal of Process Control,* vol. 8, no. 5-6, pp. 375-380.

[85]   Juricek, B. C., Seborg, D. E. and Larimore, W. E. (2002), "Identification of multivariable, linear, dynamic models: Comparing regression and subspace techniques", *Industrial and Engineering Chemistry Research,* vol. 41, no. 9, pp. 2185-2203.

[86]   Juricek, B. C., Seborg, D. E. and Larimore, W. E. (2001), "Identification of the Tennessee Eastman challenge process with subspace methods", *Control Engineering Practice,* vol. 9, no. 12, pp. 1337-1351.

[87]   Larimore, W. E. (1990), "Canonical variate analysis in identification, filtering, and adaptive control", *Proceedings of the IEEE Conference on Decision and Control,* Vol. 2, pp. 596.

[88]   Ho, D. and Randall, R. B. (2000), "Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals", *Mechanical Systems and Signal Processing,* vol. 14, no. 5, pp. 763-788.

[89]   Tan, C. C. (1987), "An Adaptive Noise Cancellation Approach for Condition Monitoring of Gear Box Bearings", *International Tribology Conference (1987 : Melbourne, Vic.),* Vol. 87/18 1987, 2-4 December 1987, Melbourne, pp. 360-365.

[90]   Randall, R. B. and Antoni, J. (2011), "Rolling element bearing diagnostics—A tutorial", *Mechanical Systems and Signal Processing,* vol. 25, no. 2, pp. 485-520.

[91]   Behzad, M., Bastami, A. R. and Mba, D. (2012), "Rolling bearing fault detection by short-time statistical features", *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering,* vol. 226, no. 3, pp. 229-237.

[92]   Li, H., Zhang, Y. and Zheng, H. (2009), "Hilbert-Huang transform and marginal spectrum for detection and diagnosis of localized defects in roller bearings", *Journal of Mechanical Science and Technology,* vol. 23, no. 2, pp. 291-301.

[93]   Li, F., Meng, G., Ye, L. and Chen, P. (2008), "Wavelet transform-based higher-order statistics for fault diagnosis in rolling element bearings", *JVC/Journal of Vibration and Control,* vol. 14, no. 11, pp. 1691-1709.

[94]   Makhoul, J. (1975), "Linear prediction: a tutorial review.", *Proceedings of the IEEE,* vol. 63, no. 4, pp. 561-580.

[95]   Widrow, B., Glover Jr., J. R. and McCool, J. M. (1975), "Adaptive noise cancelling: principles and applications", *Proceedings of the IEEE,* vol. 63, no. 12, pp. 1692-1716.

[96]   Da Silva, S. and Dias Junior, M. (2007), "Statistical damage detection in a stationary rotor systems through time series analysis", *Latin American Applied Research,* vol. 37, no. 4, pp. 243-246.

[97]   Dron, J. -., Rasolofondraibe, L., Chiementin, X. and Bolaers, F. (2010), "A comparative experimental study on the use of three denoising methods for bearing defect detection", *Meccanica,* vol. 45, no. 2, pp. 265-277.

[98]   Lu, B., Nowak, M., Grubic, S. and Habetler, T. G. (2009), "An adaptive noise-cancellation method for detecting generalized roughness bearing faults under dynamic load conditions", *2009 IEEE Energy Conversion Congress and Exposition, ECCE 2009,* pp. 1091.

[99]   Sawalhi, N. and Randall, R. B. (2008), "Helicopter gearbox bearing blind fault identification using a range of analysis techniques", *Australian Journal of Mechanical Engineering,* vol. 5, no. 2, pp. 157-168.

[100]   Sui, W. and Zhang, D. (2010), "DWT-based adaptive filter and its application on canceling noise in mechanical signals", *Key Engineering Materials,* vol. 439-440, pp. 6-11.

[101]   Wang, W. (2008), "Autoregressive model-based diagnostics for gears and bearings", *Insight: Non-Destructive Testing and Condition Monitoring,* vol. 50, no. 8, pp. 414-418.

[102]   Patel, V. N., Tandon, N. and Pandey, R. K. (2012), "Improving defect detection of rolling element bearings in the presence of external vibrations using adaptive noise cancellation and multiscale morphology", *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology,* vol. 226, no. 2, pp. 150-162.

[103]   Antoni, J. (2006), "The spectral kurtosis: A useful tool for characterising non-stationary signals", *Mechanical Systems and Signal Processing,* vol. 20, no. 2, pp. 282-307.

[104]   Antoni, J. and Randall, R. B. (2006), "The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines", *Mechanical Systems and Signal Processing,* vol. 20, no. 2, pp. 308-331.

[105]   Antoni, J. (2007), "Fast computation of the kurtogram for the detection of transient faults", *Mechanical Systems and Signal Processing,* vol. 21, no. 1, pp. 108-124.

[106]   Chen, J., Zi, Y., He, Z. and Yuan, J. (2012), "Improved spectral kurtosis with adaptive redundant multiwavelet packet and its applications for rotating machinery fault detection", *Measurement Science and Technology,* vol. 23, no. 4.

[107]   Li, H., Zheng, H. and Tang, L. (2012), "Bearing fault diagnosis based on kurtogram of dual-tree complex wavelet packet transform", *Zhendong yu Chongji/Journal of Vibration and Shock,* vol. 31, no. 10, pp. 13-18.

[108]   Bechhoefer, E., Kingsley, M. and Menon, P. (2011), "Bearing envelope analysis window selection Using spectral kurtosis techniques", *2011 IEEE International Conference on Prognostics and Health Management, PHM 2011 - Conference Proceedings,* .

[109]   Lei, Y., Lin, J., He, Z. and Zi, Y. (2011), "Application of an improved kurtogram method for fault diagnosis of rolling element bearings", *Mechanical Systems and Signal Processing,* vol. 25, no. 5, pp. 1738-1749.

[110]   Wang, Y. and Liang, M. (2011), "An adaptive SK technique and its application for fault detection of rolling element bearings", *Mechanical Systems and Signal Processing,* vol. 25, no. 5, pp. 1750-1764.

[111]   Su, W., Wang, F., Zhang, Z., Guo, Z. and Li, H. -. (2010), "Application of EMD denoising and spectral kurtosis in early fault diagnosis of rolling element bearings", *Zhendong yu Chongji/Journal of Vibration and Shock,* vol. 29, no. 3, pp. 18-21.

[112]   Yule, G. U. (1927), "On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers", *Philosophical Transactions of the Royal Society of london,* vol. 226-A, pp. 267-298.

[113]   Randall, R. B. (2011), "3.6.3 Linear Prediction", in John Wiley and Sons (ed.) *Vibration-based condition monitoring: Industrial, aerospace and automotive applications,* 2011th ed, Wiley, Singapore, pp. 122-125.

[114]   Ljung, L. (1999), "10.1 Linear Regressions and Least Squares", in Pentrice-Hall (ed.) *System identification: Theory for the user,* 2nd ed, Pentrice-Hall, New Jersey, pp. 321-324.

[115]   Chaturvedi, G. K. and Thomas, D. W. (1982), "Bearing fault detection using adaptive noise cancelling.", *TRANS.ASME J.MECH.DES.,* vol. 104, no. 2 , Apr. 1982, pp. 280-289.

[116]   Antoni, J. and Randall, R. B. (2004), "Unsupervised noise cancellation for vibration signals: Part I - Evaluation of adaptive algorithms", *Mechanical Systems and Signal Processing,* vol. 18, no. 1, pp. 89-101.

[117]   Ho, D. and Randall, R.B. (1997), "Effects of time delay, order of FIR filter and convergence factor on self adaptive noise cancellation", *Fifth International Congress on Sound and Vibration,* December 15-18 1997, Adelaide, pp. 945-952.

[118]   Randall, R. B. (2011), "3.1. Probability Distribution and Density", in John Wiley and Sons (ed.) *Vibration-based condition monitoring: Industrial, aerospace and automotive applications,* 2011th ed, Wiley, Singapore, pp. 63-66.

[119]   Dwyer, R. F. (1983), "Detection of non-gaussian signals by frequency domain kurtosis estimation.", *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings,* Vol. 2, pp. 607.

[120]   Randall, R. B. (2011), "5.3 Spectral Kurtosis and the Kurtogram", in John Wiley and Sons (ed.) *Vibration-based condition monitoring: Industrial, aerospace and automotive applications,* 2011th ed, Wiley, Singapore, pp. 122-125.

[121]   Antoni, J. *, Jérôme Antoni personal page*, available at: http://www.utc.fr/~antoni/ (accessed 24/07/2012).

[122]   Smith, S. W. (1997), *The scientist and engineer's guide to digital signal processing,* 1st ed, California Technical Publishing, California.

[123]   Al-Dossary, S., Hamzah, R. I. R. and Mba, D. (2009), "Observations of changes in acoustic emission waveform for varying seeded defect sizes in a rolling element bearing", *Applied Acoustics,* vol. 70, no. 1, pp. 58-81.

[124]   Al-Ghamd, A. M. and Mba, D. (2006), "A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size", *Mechanical Systems and Signal Processing,* vol. 20, no. 7, pp. 1537-1571.

[125]   Behzad, M., AlandiHallaj, A., Bastami, A. R., Eftekharnejad, B., Charnley, B. and Mba, D. (2009), "Defect size estimation in rolling element bearings using vibration time waveform", *Insight: Non-Destructive Testing and Condition Monitoring,* vol. 51, no. 8, pp. 426-430.

[126]   Elforjani, M. and Mba, D. (2010), "Accelerated natural fault diagnosis in slow speed bearings with Acoustic Emission", *Engineering Fracture Mechanics,* vol. 77, no. 1, pp. 112-127.

[127]   Elforjani, M. and Mba, D. (2008), "Detecting the onset, propagation and location of non-artificial defects in a slow rotating thrust bearing with acoustic emission", *Insight: Non-Destructive Testing and Condition Monitoring,* vol. 50, no. 5, pp. 264-268.

[128]   Hawman, M. W. and Galinaitis, W. S. (1988), "Acoustic emission monitoring of rolling element bearings", *Ultrasonics Symposium Proceedings,* Vol. 2, pp. 885.

[129]   Kilundu, B., Chiementin, X., Duez, J. and Mba, D. (2011), "Cyclostationarity of Acoustic Emissions (AE) for monitoring bearing defects", *Mechanical Systems and Signal Processing,* vol. 25, no. 6, pp. 2061-2072.

[130]   Mba, D. and Rao, R. B. K. N. (2006), "Development of acoustic emission technology for condition monitoring and diagnosis of rotating machines: Bearings, pumps, gearboxes, engines, and rotating structures", *Shock and Vibration Digest,* vol. 38, no. 2, pp. 3-16.

[131]   Shiroishi, J., Li, Y., Liang, S., Kurfess, T. and Danyluk, S. (1997), "Bearing condition diagnostics via vibration and acoustic emission measurements", *Mechanical Systems and Signal Processing,* vol. 11, no. 5, pp. 693-705.

[132]   Yoshioka, T., Korenaga, A., Mano, H. and Yamamoto, T. (1999), "Diagnosis of rolling bearing by measuring time interval of AE generation", *Journal of Tribology,* vol. 121, no. 3, pp. 468-472.

[133]   Al-Balushi, K. R., Addali, A., Charnley, B. and Mba, D. (2010), "Energy index technique for detection of acoustic emissions associated with incipient bearing failures", *Applied Acoustics,* vol. 71, no. 9, pp. 812-821.

[134]   Chiementin, X., Mba, D., Charnley, B., Lignon, S. and Dron, J. P. (2010), "Effect of the denoising on acoustic emission signals", *Journal of Vibration and Acoustics, Transactions of the ASME,* vol. 132, no. 3, pp. 0310091-0310099.

[135]   Couturier, J. and Mba, D. (2008), "Operational bearing parameters and acoustic emission generation", *Journal of Vibration and Acoustics, Transactions of the ASME,* vol. 130, no. 2.

[136]   Elforjani, M. and Mba, D. (2011), "Condition monitoring of slow-speed shafts and bearings with Acoustic emission", *Strain,* vol. 47, no. SUPPL. 2, pp. 350-363.

[137]   Elforjani, M. and Mba, D. (2010), "Acoustic emissions observed from a naturally degrading slow speed bearing and shaft", *COMADEM 2010 - Advances in Maintenance and Condition Diagnosis Technologies Towards Sustainable Society, Proc. 23rd Int. Congr. Condition Monitoring and Diagnostic Engineering Management,* pp. 355.

[138]   Elforjani, M. and Mba, D. (2009), "Assessment of natural crack initiation and its propagation in slow speed bearings", *Nondestructive Testing and Evaluation,* vol. 24, no. 3, pp. 261-275.

[139]   Elforjani, M. and Mba, D. (2008), "Monitoring the onset and propagation of natural degradation process in a slow speed rolling element bearing with acoustic emission", *Journal of Vibration and Acoustics, Transactions of the ASME,* vol. 130, no. 4.

[140]   Elforjani, M. and Mba, D. (2008), "Observations and location of acoustic emissions for a naturally degrading rolling element thrust bearing", *Journal of Failure Analysis and Prevention,* vol. 8, no. 4, pp. 370-385.

[141]   Mba, D. (2008), "The use of acoustic emission for estimation of bearing defect size", *Journal of Failure Analysis and Prevention,* vol. 8, no. 2, pp. 188-192.

[142]   Mirhadizadeh, S. A., Moncholi, E. P. and Mba, D. (2010), "Influence of operational variables in a hydrodynamic bearing on the generation of acoustic emission", *Tribology International,* vol. 43, no. 9, pp. 1760-1767.

[143]   Tandon, N. and Nakra, B. C. (1992), "Comparison of vibration and acoustic measurement techniques for the condition monitoring of rolling element bearings", *Tribology International,* vol. 25, no. 3, pp. 205-212.

[144]   Eftekharnejad, B., Carrasco, M. R., Charnley, B. and Mba, D. (2011), "The application of spectral kurtosis on Acoustic Emission and vibrations from a defective bearing", *Mechanical Systems and Signal Processing,* vol. 25, no. 1, pp. 266-284.

[145]   Matthews, J. R. (1983), *Acoustic emission,* Gordon and Breach Science Publishers, New York.

[146]   Tan, C. K. and Mba, D. (2005), "Limitation of acoustic emission for identifying seeded defects in gearboxes", *Journal of Nondestructive Evaluation,* vol. 24, no. 1, pp. 11-28.

[147]   Williams, T., Ribadeneira, X., Billington, S. and Kurfess, T. (2001), "Rolling element bearing diagnostics in run-to-failure lifetime testing", *Mechanical Systems and Signal Processing,* vol. 15, no. 5, pp. 979-993.

[148]   Dalpiaz, G., Rivola, A. and Rubini, R. (2000), "Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears", *Mechanical Systems and Signal Processing,* vol. 14, no. 3, pp. 387-412.

[149]   Baydar, N. and Ball, A. (2001), "A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distribution", *Mechanical Systems and Signal Processing,* vol. 15, no. 6, pp. 1091-1107.

[150]   Pennacchi, P. and Vania, A. (2008), "Diagnostics of a crack in a load coupling of a gas turbine using the machine model and the analysis of the shaft vibrations", *Mechanical Systems and Signal Processing,* vol. 22, no. 5, pp. 1157-1178.

[151]   Ahmed, M., Baqqar, M., Gu, F. and Ball, A. D. (2012), "Fault detection and diagnosis using Principal Component Analysis of vibration data from a reciprocating compressor", *Proceedings of the 2012 UKACC International Conference on Control, CONTROL 2012,* pp. 461.

[152]   Cui, L., Wang, J. and Lee, S. (2014), "Matching pursuit of an adaptive impulse dictionary for bearing fault diagnosis", *Journal of Sound and Vibration,* vol. 333, no. 10, pp. 2840-2862.

[153]   Sekhar, A. S. and Prabhu, B. S. (1995), "Effects of coupling misalignment on vibrations of rotating machinery", *Journal of Sound and Vibration,* vol. 185, no. 4, pp. 655-671.

[154]   Patel, T. H. and Darpe, A. K. (2009), "Vibration response of misaligned rotors", *Journal of Sound and Vibration,* vol. 325, no. 3, pp. 609-628.

[155]   Avendano, R. D. and Childs, D. W. (2013), "One explanation for two-times running speed response due to misalignment in rotors connected by flexible couplings", *Journal of Engineering for Gas Turbines and Power,* vol. 135, no. 6.

[156]   Saavedra, P. N. and Ramírez, D. E. (2004), "Vibration analysis of rotors for the identification of shaft misalignment part 1: Theoretical analysis", *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science,* vol. 218, no. 9, pp. 971-985.

[157]   Saavedra, P. N. and Ramírez, D. E. (2004), "Vibration analysis of rotors for the identification of shaft misalignment part 2: Experimental validation", *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science,* vol. 218, no. 9, pp. 987-999.

[158]   Gibbons, C. B. (1976), "Coupling missalignment forces", *Proceedings of the 5th Turbomach Symposium,* October 1976, Texas, USA, pp. 111.

[159]   Peng, Z. K., Chu, F. L. and Tse, P. W. (2007), "Singularity analysis of the vibration signals by means of wavelet modulus maximal method", *Mechanical Systems and Signal Processing,* vol. 21, no. 2, pp. 780-794.

[160]   Sekhar, A. S. (2005), "Identification of unbalance and crack acting simultaneously in a rotor system: Modal expansion versus reduced basis dynamic expansion", *JVC/Journal of Vibration and Control,* vol. 11, no. 9, pp. 1125-1145.

Appendix A Summary of estimation results for T2
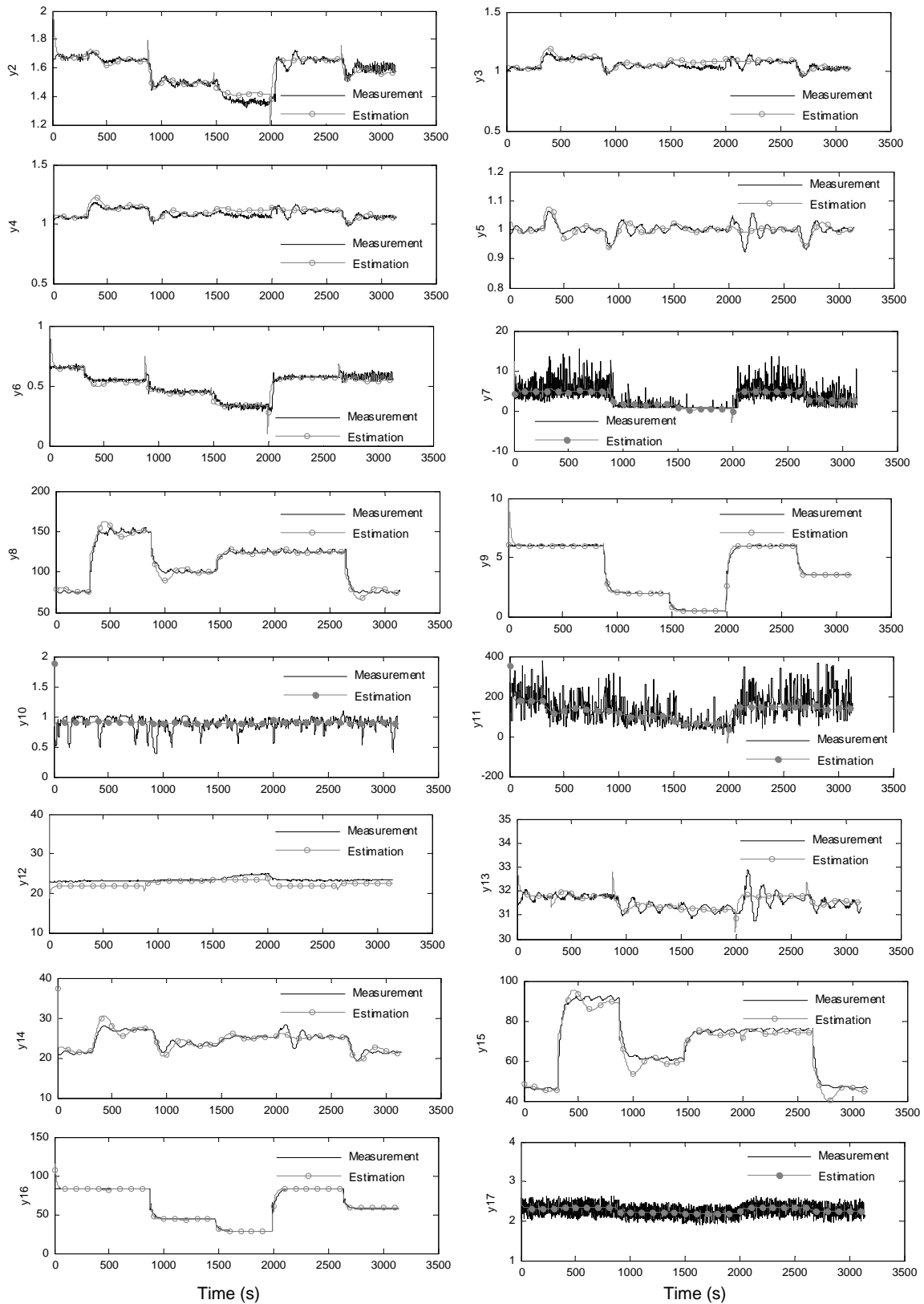
Appendix B Summary of estimation results for T3
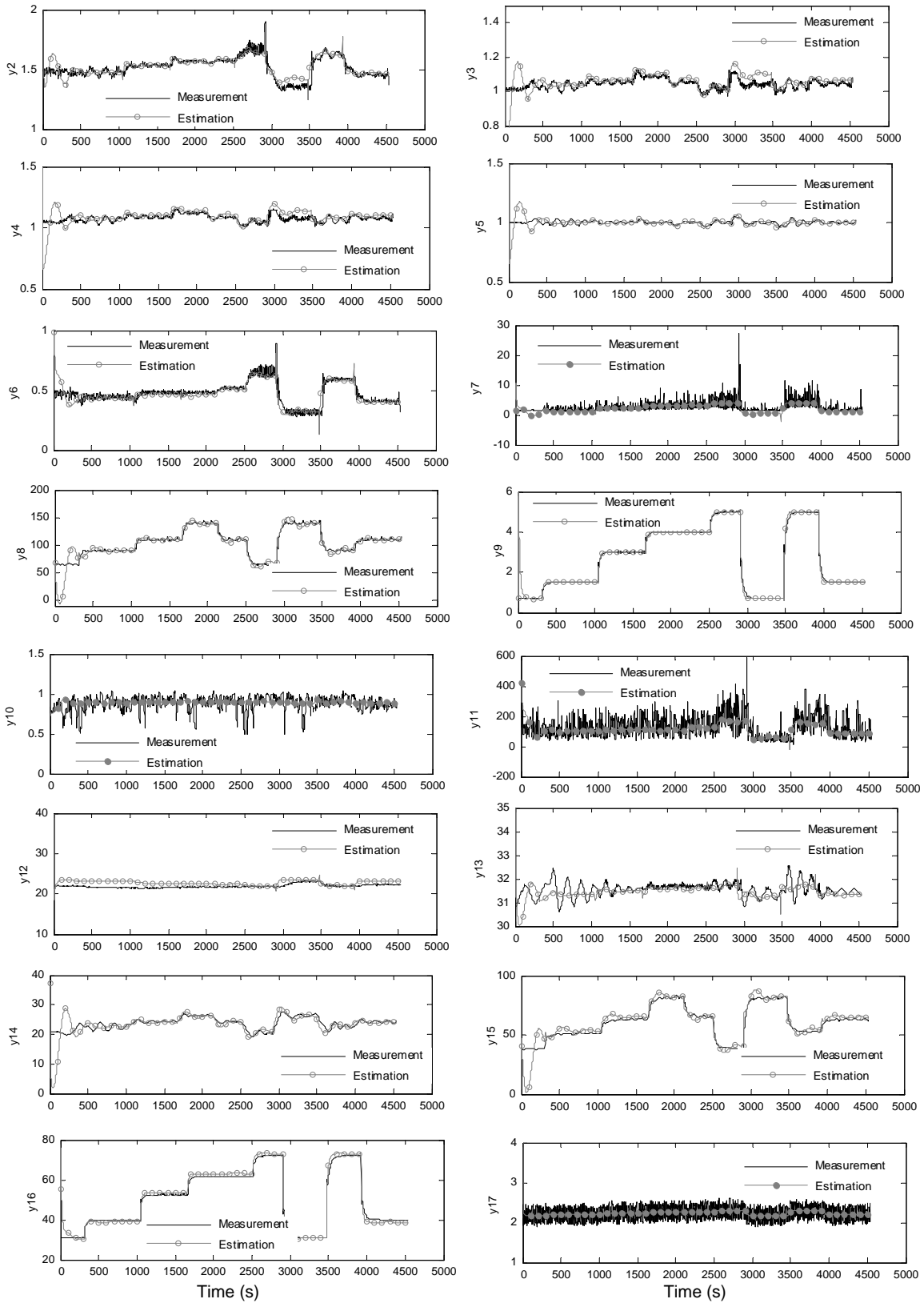
Appendix C Kurtograms from chapter 7