Joel Filipe Rodrigues Sansana

SIMULATION, MONITORING AND DIAGNOSIS OF FAULTS AND EQUIPMENT DEGRADATION IN CHEMICAL PROCESSES

UNIVERSIDADE D COIMBRA

UNIVERSIDADE D
COIMBRA

Joel Filipe Rodrigues Sansana

SIMULATION, MONITORING AND DIAGNOSIS OF FAULTS AND EQUIPMENT DEGRADATION IN CHEMICAL PROCESSES

Dissertation in the scope of the Integrated Masters in Chemical Engineering, branch of Process, Environment and Energy, supervised by Professors Marco Reis and Lino Santos and submitted to the Department of Chemical Engineering, Faculty of Science and Technology, University of Coimbra

July 2018

Joel Filipe Rodrigues Sansana

Simulation, monitoring and diagnosis of faults and equipment degradation in chemical processes

Master's dissertation in the scientific area of Chemical Engineering, supervised by Professors Marco Reis and Lino Santos and submitted to the Department of Chemical Engineering, Faculty of Science and Technology, University of Coimbra

Supervisors:
Professor Marco Paulo Seabra dos Reis, PhD
and
Professor Lino de Oliveira Santos, PhD

Institution:
Departament of Chemical Engineering
Faculty of Sciences and Technology of the University of Coimbra

Universidade de Coimbra

Coimbra, 2018

À minha avó (in memorium),
Maria Rodrigues Margato

Acknowledgements/Agradecimentos

If I may, allow me the pleasure to thank for all the support, the help and all round contributions that many people gave me throughout not only the semester but also throughout the time I have been a proud student of the wonderful University of Coimbra. All these people enriched my experience here and my life, although tuition impoverished me. Without exception, I thank you all deeply. Nevertheless, as the great Fernando Pessoa used to say «*a minha pátria é a língua portuguesa*» and as a result all these people are fellow countrymen. That is why I make a point to address each and everyone in the beautiful Portuguese language.

Em primeiro lugar, quero agradecer aos meus pais, Cristina Guímaro e Gracílio Sansana. Obrigado pelo apoio que sempre me deram desde pequeno para poder ser alguém na vida. Obrigado por todos os conselhos, pelos valores, pelo financiamento e por me mostraram que com trabalho e integridade podemos levantar a cabeça em qualquer lado, não interessa de onde viemos. Obrigado por me terem posto na linha quando precisei. Obrigado por irem à Granja do Ulmeiro todas as sextas-feiras e domingos e obrigado por ficarem chateados quando não telefono. Muito obrigado!

À Ju, por me inspirar a trabalhar. Quando for grande quero estudar tanto como tu!

À Andreia, obrigado por estares sempre presente a meu lado, por me ajudares a manter a sanidade, por fazeres os melhores panados e panquecas e pela preciosa ajuda que me deste para fazer a análise de riscos para a saúde. Gosto de ti!

Aos camaradas das noites etílicas: Rita de la Vodka, Maria Leonor, Inocêncio, Maria Banaco, Mariana Alvim, afilhada Telma, Adriano Dias, Nélson Henrique e Eugeniu «Estrangeiro» Strelet. Obrigado (mulțumesc) por tornarem estes 5 anos interessantes e divertidos. A arte de fazer amigos é das mais difíceis, mas realmente também das mais proveitosas. Apesar de termos tido momentos dos quais não me recordo, a vós não vos esquecerei.

Agradeço a todos os professores do Departamento de Engenharia Química da FCTUC por me armarem com o conhecimento para redigir esta dissertação, especialmente ao Doutor Lino Santos por, apesar de longe, estar sempre disponível a ajudar e ao Doutor Marco Reis pelo acompanhamento e aconselhamento ao longo do semestre e por me ter mostrado o mundo da data science *que tanto gosto tenho tido em descobrir.*

Muito obrigado a todos!
P.S.: Se esqueci alguém peço desculpa e fico a dever um fino.

Resumo

O controlo estatístico de processos faz parte do dia-a-dia industrial desde os anos 30, mas a evolução tecnológica, que permite a medição e o registo de inúmeras variáveis, exige metodologias mais avançadas capazes de lidar com a atual escala de dados na gama dos *petabytes*. A Análise dos Componentes Principais (ACP) e o Controlo Estatístico Multivariado de Processo (CEMP) têm vindo a ser saudados como uma potencial resposta para monitorizar processos químicos e detetar eventos especiais. Além disso, existem reações e fenómenos complexos que ocorrem em processos químicos e que têm um efeito negativo, a longo prazo, na qualidade do produto, nas condições do equipamento ou no consumo de energia. No entanto, a monitorização de processos geralmente não procura essas dinâmicas, perdendo possíveis melhorias de processo. Nesta tese, os objetivos são detetar e diagnosticar falhas bruscas simuladas, usando métodos de CEMP, e detetar dinâmicas mais lentas que contribuem para a degradação de processos e equipamentos.

As metodologias usadas para atingir o primeiro objetivo são baseadas em ACP, cartas de controlo do CEMP e gráficos de contribuição. As duas primeiras são mais focadas na deteção, enquanto o segundo toma o lugar central no diagnóstico de falhas. No que diz respeito à degradação de processos e equipamentos, o ACP em conjunto com a Análise de Multirresolução (AMR) e as transformadas de onduleta tentam detetar as dinâmicas mais lentas ou de longo prazo que afetam os processos químicos, mas cujos efeitos são mascarados pelos controladores.

Para testar as metodologias, um sistema foi simulado para gerar os dados necessários, bem como ter acesso à verdade básica do processo que não está disponível na realidade. O sistema é um reator contínuo perfeitamente agitado (RCPA) não-isotérmico sob controlo de um ciclo de realimentação, tendo a flexibilidade necessária para simular várias condições defeituosas.

Em relação à detecção de falhas abruptas, as metodologias ACP-CEMP foram robustas e mostraram boas capacidades de deteção. O diagnóstico de falhas teve resultados mistos, com os gráficos de contribuição falhando em algumas classes de falhas.

Finalmente, o ACP-AMR teve resultados promissores, pois foi capaz de detetar as dinâmicas de degradação simuladas, abrindo caminho para o desenvolvimento de um modelo capaz de prever o estado de degradação de um processo com base em medições de variáveis.

Palavras-chave: Análise dos componentes principais, Controlo estatístico multivariado de processo, Análise multiresolução, Degradação de processo & equipamento, simulação.

Abstract

Statistical process control is part in the industrial landscape since the 1930's, but technological evolution which allows the measurement and recording of countless variables asks for more advanced methodologies able to cope with the current petabyte data scale. Principal Component Analysis (PCA) and Multivariate Statistical Process Control (MSPC) have been hailed as a potential answer to monitor chemical processes and detect special causes. Moreover, there are complex reactions and processes that occur in chemical processes and have a negative effect, in the long term, on product quality, equipment conditions or energy consumption. However, process monitoring usually does not look for these dynamics, leaving possible process improvements unattended. In this thesis, the goals are to detect and diagnose simulated abrupt faults using MSPC methods and to detect slower dynamics that contribute to process & equipment degradation.

The methodologies used to achieve the first goal are based on PCA, MSPC control charts and contribution plots. The first two are more focused on detection while the latter takes central stage in fault diagnosis. Regarding process & equipment degradation, PCA in conjunction with Multiresolution Analysis (MRA) and wavelet transforms attempt to detect the slower or long term dynamics that affect chemical processes but whose effects are masked by controllers.

To test the methodologies a system was simulated to generated the necessary data as well as to have access to the process's ground truth which is not available in reality. The system is a non-isothermal Continuous Stirred-Tank Reactor (CSTR) under feedback control with the needed flexibility to simulate an array of faulty conditions.

Concerning the detection of abrupt faults, the PCA-MSPC methodologies came through robust and showing good detection capabilities. Fault diagnosis had mixed results with the contribution plots failing in some failure classes.

Finally, PCA-MRA had promising results as it was able to detect the simulated process & degradation dynamics, paving the way to the development of a model that predicts the state of a process degradation based on variable measurements.

keywords: Principal component analysis (PCA), Multivariate statistical process control (MSPC), Multiresolution analysis (MRA), Process & equipment degradation, Simulation.

Contents

List of Tables

List of Figures

Abbreviations

ARL Average Run Length. 41, 65

ARMA Autoregressive Moving-Average. 20

ATS Average Time to Signal. 41, 65

CD Complete Decomposition. 8, 9

CSTR Continuous Stirred-Tank Reactor. ix, 17, 19, 21, 25

CUMSUM Cumulative Sum. 1

EWMA Exponentially Weighted Moving Average. 1

FAR False Alarm Rate. 8, 36, 41, 65

i.i.d. independent and identically distributed. 8

LCL Lower Control Limit. 7

MRA Multiresolution Analysis. ix, 3, 46, 51

MSPC Multivariate Statistical Process Control. ix, 1, 2, 3, 7, 11, 33, 36, 39, 40, 41, 43, 51

NOC Normal Operating Conditions. 5, 7, 11, 25, 33, 34, 36, 39

P Proportional. 23, 24

PC Principal Component. 6, 7, 8, 12, 34, 35, 36, 41, 42, 43, 44, 49, 50

PCA Principal Component Analysis. ix, 1, 3, 5, 6, 7, 11, 12, 13, 14, 33, 34, 36, 37, 39, 41, 44, 47, 49, 51

PI Proportional and Integral. 23, 24

PID Proportional, Integral and Derivative. 23

PLS Partial Least Squares. 1

PSE Process Systems Engineering. 2

SPE Squared Prediction Error. 1, 7

SPM Statistical Process Monitoring. 1

TDR True Detection Rate. 41, 65

UCL Upper Control Limit. 7, 8, 36, 41, 43, 51

var Variable. 33, 34, 35, 36, 37, 38, 40, 42, 49, 50

Symbols

A_b Cross-sectional area [m^2]. 18

A_t Heat transfer area [m^2]. 19

C_A Reactant A concentration [$\mathrm{kmol/m}^3$]. 17, 19, 20

C_A0 Feed concentration of reactant A [$\mathrm{kmol/m}^3$]. 17, 19, 20

C_P Mixture heat capacity at constant pressure [J/(kg K)]. 19, 20

C_Pj Thermal fluid heat capacity at constant pressure [J/(kg K)]. 19, 20

c_valve Discharge coefficient [$\mathrm{m}^{2.5}$/min]. 26

D Reactor diameter [m]. 19, 20

E_a Activation Energy divided by the ideal gas constant [K]. 18, 20

F Outflow rate [m^3/min]. 18, 26

F_0 Feed flow rate [m^3/min]. 17, 18, 19, 20, 22, 26

F_j Thermal fluid flow rate [m^3/min]. 17, 20, 22, 26

F_jv Thermal fluid flow rate after the jacket [m^3/min]. 17, 19, 20, 26

g Gravitational acceleration [m s^{-2}]. 26

h Liquid level [m]. 17, 18, 19, 20, 26

J Slipjump [-]. 28, 29

k_0 Pre-exponential factor [min^{-1}]. 18, 20

K_c Controller stationary constant [m^{-1}]/[$^\circ\mathrm{C}^{-1}$]. 23, 24

k_d Catalyst deactivation rate [min^{-1}]. 19, 31, 44

k_factor Kinetic constant [m^{-1}]. 18, 19

K_v Valve's stationary gain [-]. 21, 27, 29

m_d Deposit mass rate [m^2 °C $(\mathrm{J/min})^{-1}$ min^{-1}]. 30, 31, 44

$m\Delta H_\mathrm{R}$ Reaction enthalpy [J/kmol]. 19, 20

r Reaction rate [$\mathrm{mol/(m}^3$ s)]. 19

r_A Reaction rate of A [$\mathrm{mol/(m}^3$ s)]. 19

R_f Fouling Resistance [m^2 °C $(\mathrm{J/min})^{-1}$]. 30

S Stickband plus deadband [-]. 28, 29

t time [min]. 19, 21, 26, 30

T Mixture temperature [°C]. 17, 18, 19, 20, 26

T_0 Feed temperature [°C]. 17, 19, 20

T_j Thermal fluid temperature in the jacket [°C]. 17, 19, 20, 26, 37

$T_{\mathrm{j}0}$ Thermal fluid temperature before the valve [°C]. 17, 19, 20

$\boldsymbol{u}$ Loading variables vector. 17

U Global heat transfer coefficient [J/(min m^2 °C)]. 19, 20, 30

V Reactor volume [m^3]. 19

V_base Reactor's base volume [m^3]. 18, 20

V_j Jacket volume [m^3]. 19, 20

x Data value. 5

$\boldsymbol{x}$ State variables vector. 5, 17

$\boldsymbol{X}$ Data matrix. 5, 6, 7, 11, 14, 33, 34, 36, 41

$\overline{\boldsymbol{X}}$ Vector of average values of each variable. 5

$\boldsymbol{X}_C$ Centered data matrix. 5, 6

$\boldsymbol{Z}$ Auto-scaled matrix. 6

α_A Stoichiometric coefficient of A [-]. 19, 20

α_catalyst Catalyst activity factor [-]. 17, 19, 20, 31

ζ Valve's damping factor [-]. 21

λ Valve's position [-]. 21, 22, 28

ρ Mixture density [$\mathrm{kg/m}^3$]. 19, 20

ρ_j Thermal fluid density [$\mathrm{kg/m}^3$]. 19, 20

τ_I Integral time [min]. 24

τ_jacket Thermal fluid's residence time in the jacket [min]. 20, 26

τ_V Valve's time constant [min]. 21

Chapter 1

Introduction

«The world is full of obvious things which nobody by any chance ever observes.»
*(*Arthur Conan Doyle*)*

This chapter presents an holistic overview of the contents and challenges addressed in this thesis as well as its motivations and goals. It comprises three sections. The first one is focused on explaining the importance of Statistical Process Monitoring (SPM) in the most recent paradigm of chemical engineering, on exploring its methodologies and pointing out some fundamental work published over recent years and on which this thesis builds upon. Next, the thesis goals are outlined. Finally, on the third section, an overview of how this thesis is structured is provided.

1.1 Motivation

Ever since the 1930's, when Shewhart presented his groundbreaking works in control charts (Shewhart 1931), process engineers have monitored variables relevant to product quality with the goal of assessing process performance, and verify if it is in a state of statistical control. These efforts lead to a plethora of univariate control charts which are presented in control rooms and operators have to monitor them all. Besides the aforementioned Shewhart chart, Cumulative Sum (CUMSUM) and Exponentially Weighted Moving Average (EWMA) charts are also often used.

With the advances in sensors and instrumentation, collecting data is getting easier and cheaper, giving engineers more data to infer the state of the process. Therefore, industry and academia have been pressured to come up with better and more expedite ways for dealing with the massive amounts of data collected. Thus, latent variable methods entered the process control world, first to control continuous processes product quality and then batch processes (Kourti 2005).

MSPC techniques coupled with latent variable methodologies gained importance since the works of Kresta et al. (1991). The PCA-MSPC methodology was systematized in Kourti & MacGregor (1995) and MacGregor & Kourti (1995). Later, Wise & Gallagher (1996) summarized what become known as process chemometrics which makes use of PCA

(Bro & Smilde 2014), Partial Least Squares (PLS) (Wold et al. 2001) and multivariate control charts such as Hotelling's T^2 and Squared Prediction Error (SPE), also known as Q-statistic (Rato & Reis 2013*a*).

The advent of the digital era transformed science. It allowed calculations to be made much faster and with fewer costs (Aspuru-Guzik et al. 2018). Chemical engineering was no stranger to this revolution and expanded its reach through the advances made in Process Systems Engineering (PSE) and Molecular Modeling and Simulation (Ratcliff et al. 2017). PSE permitted the development of process simulators and granted engineers the possibility of simulating interactions with a process without disturbing reality. Engineers can now be more bold in testing new structures to improve process sustainability (Palmer & Debenedetti 2015).

Modern chemical industries are always on the look out for new ways to improve their energy efficiency, product quality and operation standards, focusing on reducing the variability of process variables around their primary target (Rato & Reis 2013*b*).

Chemical processes are influenced by longterm dynamics that are slowly adjusted by the controllers actions. These dynamics may not have an immediate effect on product quality, but in the long run they can severely affect process safety, maintenance and product quality. These aspects are often compartmentalized: process safety to Health and Safety Engineers, product quality to Quality Engineers and maintenance to Mechanical Engineers. However, a change is on the horizon. The amount of data collected on a chemical process may allow someone with the *savoir-faire* to turn data into knowledge and to interpret those results process-wise and create value in all of these fields. This leads to the belief that Chemical Engineers should study these areas looking for small improvements in the process which delay maintenance and keep process safety. A mere 1% improvement in energy efficiency or improved controller maintenance represents hundreds of millions of dollars in savings to the process industries (Desborough & Miller 2002).

1.2 Goals

The above mentioned motivations create conditions to evaluate data in a more intelligent way, taking wisdom from the process and paving the way for industry to achieve a state of smart manufacturing. That being stated and knowing the time limitations of a master's dissertation, it is proposed *to tackle the challenge of simultaneously monitoring and diagnosing abrupt process failures and to detect long term equipment degradation*. In order to achieve this goal, the following tasks are carried out:

- Building a MATLAB® simulator of normal process conditions.
- Simulating faults in the process, in sensors and in control valves.
- Simulating process and equipment degradation phenomena, such as catalyst activity loss and fouling.
- Monitor and diagnose faults through MSPC.
- Develop data analysis tools that are able to extract slower degradation dynamics changes in the process.

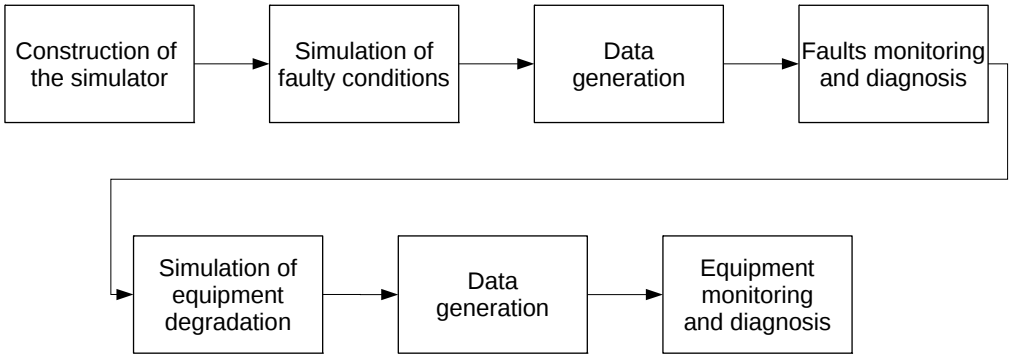


Figure 1.1: Work route followed in this thesis to achieve the proposed goals.

1.3 Thesis overview

The thesis is divided into 6 chapters. This organization reflects the progression of the work made during the semester, as well as the workflow depicted in Figure 1.1. At the beginning of each chapter, a small introduction will be presented to frame its topics.

Chapter 2 contains the description of the PCA-MSPC methodologies explored in this work to process monitoring and contribution plots used for fault detection and diagnosis.

Chapter 3 focus on strategies to monitor process & equipment degradation. These include PCA and wavelet-based MRA methodology.

In chapter 4 the simulator's making of, the methods used and their theoretical foundations are discussed. An explanation of each of the several faults added to the system is also given.

Chapter 5 presents the results of applying the methodologies presented in Chapters 2 and 3 to the data generated by the system simulated in Chapter 4.

Last but not least, in Chapter 6, some thesis' conclusions are outlined and possible future work is mentioned.

Chapter 2

Fault detection and diagnosis

«Success consists of going from failure to failure without loss of enthusiasm.»
*(*Winston Churchill*)*

When the number of variables analyzed becomes large, with tens or more variables, it is often the case that there are variables highly correlated with one another, causing problems in the inversion of the co-variance matrix during the Mahalanobis distance computation (Rato & Reis 2013*b*). To avoid this problem, a latent variable modeling framework such as PCA can be used to reduce the dimensionality of the variable space and stabilize the computation of the monitoring statistics (Jackson 1991).

2.1 Principal component analysis

Considering $\boldsymbol{X}_{n\times m}$, a dataset representative of Normal Operating Conditions (NOC) operation, as a single block of n observations of m variables, PCA determines the subspace of lower dimensionality that explains the maximum of variability of the original dataset (for the considered dimensionality) and makes a projection of the original variables into it. Firstly, it is advised to preprocess $\boldsymbol{X}$ through auto-scaling in order to remove the effect of different scales due to the existence of different units. Auto-scaling consists in computing first the average of each variable in $\boldsymbol{X}$,

$$\overline{\boldsymbol{X}}_j = \frac{1}{n} \cdot \sum_{i=1}^{n} x_{ij} \,. \tag{2.1}$$

Then the variables are centered to an average equal to zero,

$$\boldsymbol{X}_{Cij} = \boldsymbol{X}_{ij} - \overline{\boldsymbol{X}}_j \,. \tag{2.2}$$

Mean centering is a suitable method when variables have the same units of measurement. When that is not the case further calculations are needed. The standard deviation of the variables, defined as

$$s_X = \sqrt{\frac{\sum_{i=1}^{n}(\boldsymbol{x}_i - \overline{x})^2}{N-1}}\,. \tag{2.3}$$

is then used to compute an auto-scaled matrix $\boldsymbol{Z}$ as follows:

$$\boldsymbol{Z}_{ij} = \frac{\boldsymbol{X}_{Cij}}{s_{Xj}}\,, \tag{2.4}$$

Auto-scaled variables in this matrix have an average value equal to zero and a variance equal to one. Therefore, they have *a priori* the same weight in a PCA analysis. This preprocessing is a common solution to address the problem of handling different variables with different scales (Bro & Smilde 2014).

The co-variance matrix of $\boldsymbol{Z}$ which is equal to the correlations matrix of $\boldsymbol{X}$, is the quantity from which the eigenvalues and their respective eigenvectors are extracted, in order to obtain the fundamental PCA outcomes (Johnson & Wichern 2007). The eigenvalues are then ordered from the greatest to the smallest. Greater eigenvalues represent the most variability of the data, while small eigenvalues often represent noise. At the end, PCA decomposes $\boldsymbol{Z}$ as follows,

$$\boldsymbol{Z} = \boldsymbol{T}\boldsymbol{L}^T\,, \tag{2.5}$$

where the principal component scores, $\boldsymbol{T}_{n\times m}$, are defined as the observed values of the Principal Component (PC) for each of the n observation vectors and the loadings matrix, denoted as $\boldsymbol{L}_{m\times m}$, whose columns contain the weights of each variable in each PC (MacGregor & Kourti 1995).

To evaluate the number of PCs included in the PCA model, there are several methodologies that can be applied, such as the Kaiser Test, the Scree Test and fraction of variance explained. Any of these methodologies is based solely on the information contained in the eigenvalues (Jollife 2002).

The Kaiser Test considers that, in auto-scaled data, any PCs with an eigenvalue below one should be discarded. The Scree Test is a method based on the assumption that random noise levels off linearly as the numbers of components increase. Therefore, eigenvalues are plotted and the bottom ones that form a plateau are discarded. In some scenarios, interpretation of the data is needed to make the soundest choice. If the scree plot breaks to much lower eigenvalues, at *e.g.* 70%, and the level of variance achieved is sufficient to work with, the remaining PCs can be discarded (Bro & Smilde 2014).

Deciding how many PCs to use, allows the definition of the pseudo-rank, p, which is equal to number of chosen PCs. With this new information $\boldsymbol{Z}$ is defined by

$$\boldsymbol{Z} = \boldsymbol{T}_{n\times p}\ \boldsymbol{L}^T_{m\times p} + \boldsymbol{E}_{n\times m}\,, \tag{2.6}$$

where $\boldsymbol{E}_{n\times m}$ is the residual matrix that contains the accumulated contribution of the last PCs, given by

$$\boldsymbol{E}_{n\times m} = \boldsymbol{Z} - \hat{\boldsymbol{Z}} \tag{2.7}$$

where $\hat{\boldsymbol{Z}}$ is the reconstructed $\boldsymbol{Z}$ matrix from the scores and loadings with pseudo-rank p.

After choosing the PCs to build the PCA model, the loadings and the scores matrices should be analyzed. The loadings values indicate which variables contribute more to the variability along a given PC. Scores plotted against each other form clusters that may allow the detection of a special event with assignable cause. It also matters to look at the residues matrix to verify if none of the important features is missing in the model.

2.2 Multivariate control charts

PCA-MSPC is divided in two phases. In Phase I, the objective is to verify if the system is stable using a dataset representing NOC operation, also called training set. Phase II deals with proper detection in the future operation of the process. In this thesis, to achieve a more robust PCA-MSPC strategy and to avoid over-adjustment, Phase I is divided in two stages, a) and b). Therefore, the PCA-MSPC methods has the following sequence of phases/stages:

Phase I stage a)

- Training set (NOC) - PCA model.

Phase I stage b)

- Validation set (NOC) - MSPC statistics.

Phase II

- Test set - Fault detection.

In order to establish the control limits monitoring statistics can be computed using a validation set $\boldsymbol{X}_{n\times m}$, representing NOC operation but different from the dataset used in PCA.

The Hotelling's T^2 control chart is a multivariate generalization of the t-student statistic and it is related with the statistical distance between each observation and the variable's average vector. It is the squared Mahalanobis distance (Rato & Reis 2013*a*),

$$T^2_{\text{PCA}} = \sum_{i=1}^{p} \frac{\boldsymbol{t}_i^2}{\boldsymbol{\lambda}_i}, \tag{2.8}$$

where T^2_{PCA} is the Hotelling's statistic, $\boldsymbol{\lambda}$ is an eigenvalue and $\boldsymbol{t}$ is a scores vector. The Q-statistic, or SPE, looks at the eigenvalues that were not included the PCA model. Q chart's observations can be defined as a sum of squared errors of prediction (MacGregor & Kourti 1995) which gives,

$$Q = \sum_{i=1}^{m} \boldsymbol{e}_i^2, \tag{2.9}$$

$\boldsymbol{e}_i$ is a residuals vector, defined as

$$\boldsymbol{e}_i = \boldsymbol{x}_i - \hat{\boldsymbol{x}}_i \tag{2.10}$$

By definition both Lower Control Limits (LCLs) are zero. The Upper Control Limits (UCLs) can be computed with (Montgomery 2013),

$$\mathrm{UCL}_{T^2} = \frac{p(n-1)(n+1)}{n^2 - np} F_{\alpha,n,n-p}\,, \tag{2.11}$$

and

$$\mathrm{UCL}_Q = \theta_1 \left(\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)^{\frac{1}{h_0}}}{\theta_1^2} \right)^{\frac{1}{h_0}}, \tag{2.12}$$

where p is the number of PCs considered in the detection, n is the number of observations and $1-\alpha$ is the confidence interval. F is a continuous probability distribution, known as the Fisher-Snedecor distribution (Montgomery et al. 2012) and c is the common normal or Gaussian distribution, while h_0 is

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2}\,, \tag{2.13}$$

where θ_i is given by

$$\theta_i = \sum_{j=p+1}^{n} \lambda_j^i\,, i = 1, 2, 3\,. \tag{2.14}$$

These methods of computing UCLs, however, may not have a perfect liaison with the reality. For instance, in the case study investigated in this thesis the system has dynamics and not a static independent and identically distributed (i.i.d.) latent variable structure, as it is assumed in these formulations. On these conditions, the False Alarm Rate (FAR) from the control charts is not going to be around the value of the parameter α, meaning the UCLs are wrong.

Therefore, to achieve the desired FAR, the T^2 and Q UCLs were computed through the calculation of the $(1-\alpha)$th percentile of T^2 and Q observations from the validation dataset, thus attaining control charts with the targeted FAR.

2.3 Fault diagnosis through contribution plots

Upon detection of a special event, control charts do not provide any information about the cause of the event. A good way to investigate the cause of special events are contribution plots. A variety of methods are available for computing variables contributions, such as: Complete Decomposition (CD), partial decomposition and reconstruction-based contributions. In this thesis, the method CD is considered as it is the only one where all contributions are positive and the sum of these contributions is equal to the value of

the corresponding statistic. These properties make it easier to analyze and to interpret (Kerkhof et al. 2013).

The CD contribution for the mth variable of the T^2 statistic is given by,

$$\mathrm{CD}_m^{T^2} = \left[\xi_m^T \left(\boldsymbol{L}\, \boldsymbol{\Lambda}^{-1} \boldsymbol{L}^T \right)^{\frac{1}{2}} \boldsymbol{x} \right]^2 , \tag{2.15}$$

where ξ_m is the mth column of the $n \times n$ identity matrix, $\boldsymbol{L}$ is the loadings matrix, $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal holds the eigenvalues and $\boldsymbol{x}$ is a measurement vector. As for the CD contribution for the mth variable of the Q-statistic (Westerhuis et al. 2000), the definition gives

$$\mathrm{CD}_m^{Q} = \left[\xi_m^T \left(\boldsymbol{I}_{n\times n} - \boldsymbol{L}\, \boldsymbol{L}^T \right) \boldsymbol{x} \right]^2 , \tag{2.16}$$

where $\boldsymbol{I}$ is the $n \times n$ identity matrix.

When exploring contribution plots, it is important to be aware of the previous knowledge about the considered chemical process. Contributions will generally point out a variable or group of variables that contributed numerically to the out-of-control alarm. In a system with highly correlated variables, the smearing-out effect can take place, meaning that the effect of change in a variable may be spread over other variables (Kerkhof et al. 2013).

It can also happen that the fault is not a measured variable, making the analysis more envolved and complex. Moreover, complex phenomena such as fouling and catalyst activity loss manifest themselves on measured variables over long periods of time with slow dynamics that can be masked by a good controller. Therefore, these phenomena are not easy to identify strictly based on contribution plots analysis (Kourti 2005).

Chapter 3

Strategies for monitoring process & equipment degradation

«Even if you're not doing anything wrong, you are being watched and recorded.»
*(*Edward Snowden*)*

Chemical processes have dynamics spread through various time-scales. Knowing and understanding these dynamics can turn into a valuable asset for the whole manufacturing organization. While floor operators make decisions based on frequent high resolution observations, a process engineer looks at data summaries of hours or days to check if the process is stable and plans how to act on the process. On the other end, management concentrates its attention in quarterly or yearly averages to decide on the company's strategy (Reis 2005).

Figure 3.1 reflects what was discussed in the previous paragraph and focus the time-scale of the long term dynamics. Considering this time-scale, the methodology developed in this thesis aims at shining light on complex processes that happen throughout days and months and hold sway over decision making related to maintenance (*e.g.*, removing a foulant from a reactor wall), logistics (*e.g.*, buying fresh catalyst) and even at a certain degree strategy (*e.g.*, stopping the operation for general intervention).

3.1 PCA Analysis

Let us consider a matrix $\boldsymbol{X}_{n\times m}$, in NOC, holding a single block of n observations of m variables. Unlike in Chapter 2, $\boldsymbol{X}$ stores data from several days, at least. Auto-scaling and PCA is applied to the training set $\boldsymbol{X}$. With the PCA model and a validation set, the PCA-MSPC methodologies are carried out to generate the multivariate control charts, T^2 and Q-statistic.

The computed PCA model and control charts are applied to a given test set, containing data from weeks of operation. Plotting T^2 statistics against Q-statistics points out whether the charges are happening to variables included in the PCA model. Plotting scores versus time shows information about slower dynamics in a chemical process.

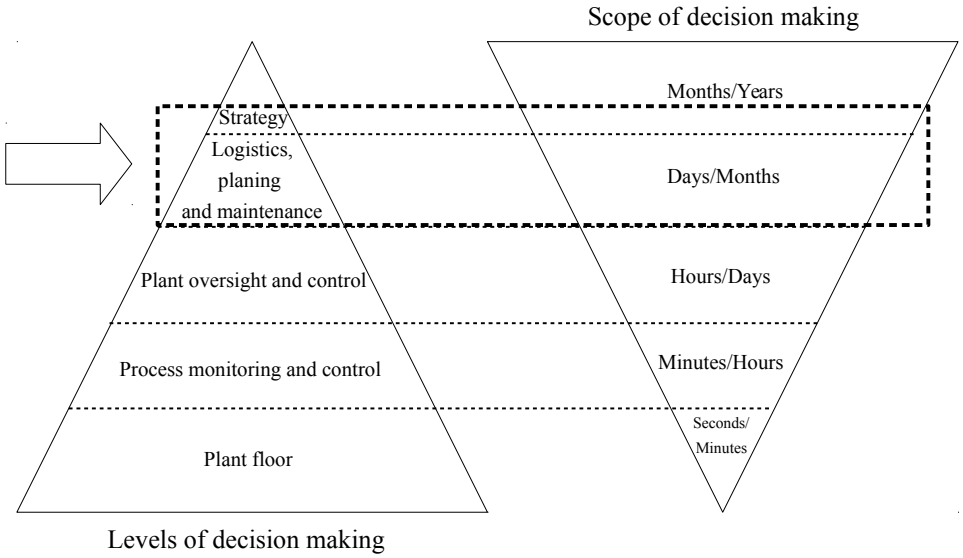


Figure 3.1: Levels of decision making in process plants and impact on the decision making process. The arrow and the dotted box highlights the focus of this section's methodology (adapted from Reis (2005)).

3.2 Window PCA

Having gathered data on several weeks of operation, it is possible to divide that data in smaller time frames. Considering $\boldsymbol{X}_{n\times m}$ as a single block of n observations of m variables, w sequential matrices, $\boldsymbol{X}_w$, are created from $\boldsymbol{X}$, each with $\frac{n}{w}$ observations of m variables.

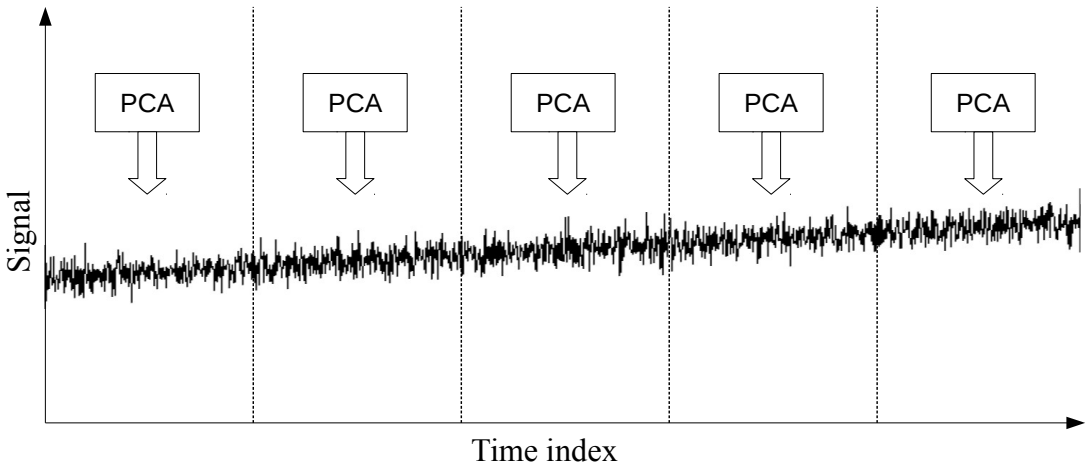


Figure 3.2: Schematic representation of window PCA.

Applying PCA to each data «window», as shown in Figure 3.2, gives w PCA models that can be compared. The different loadings matrices contain the weight of each variable along a given PC. The goal of window PCA is to access the change these weights undergo with the passing of time.

The loadings of each PCA model form a hypersurface, therefore one way to verify how the loadings change from window to window is to compute the angle between those hypersurfaces and the reference PCA model (Reis & Saraiva 2006).

The mininum angle between an arbitrary vector belonging to the PCA space of window w and the most nearly parallel to it in the PCA space of window $w+1$ is defined as θ. It is computed as follows,

$$\theta = \cos^{-1}\left(\lambda_1^{\frac{1}{2}}\right), \tag{3.1}$$

where λ_1 is the largest eigenvalue of

$$S = L_w^T \, L_{w+1} \, L_{w+1}^T \, L_w \,, \tag{3.2}$$

where L_w and L_{w+1} are the loadings matrix on windows w and $w+1$, respectively (Krzanowski 1979).

3.3 Multivariate & multiresolution analysis

Data collected from industrial processes are usually composed of complex patterns each representing different features such as sudden events, daily patterns or even monthly nuances. These patterns do not have the same time/frequency location and localization. However, they appear simultaneously in the signal and in data analysis procedure one has to make sure that the methodologies deal with all these multi-scale features without privileging certain features over others.

Therefore, in order to contend with these multi-scale features, a flexible method capable of coping with sudden high frequency events and long term low frequency dynamics is favorable. Such a flexible method is based on wavelet functions which leads to wavelet transforms (See Figure 3.3).

The wavelet transforms are an adequate tool to access the different frequencies that compose a signal exhibiting the following advantageous properties, summarized in Reis & Saraiva (2008):

1. Capacity to detect and describe localized features in the time/frequency plane;
2. Ability to extract deterministic features in a few wavelet coefficients;
3. Flexiblity to represent smooth functions as well as singularities;
4. Low computational complexity.

The projections (f_j) of a given signal in the spaces of approximation and detail of a given scale are defined as a linear combination of basis functions multiplied by the expansion coefficients, as follows,

$$f_j = \sum_k a_{k,j+N}\phi_{j+N,k} + \sum_{i=j+N}^{j+1}\sum_k d_{k,i}\psi_{i,k} \,, \tag{3.3}$$

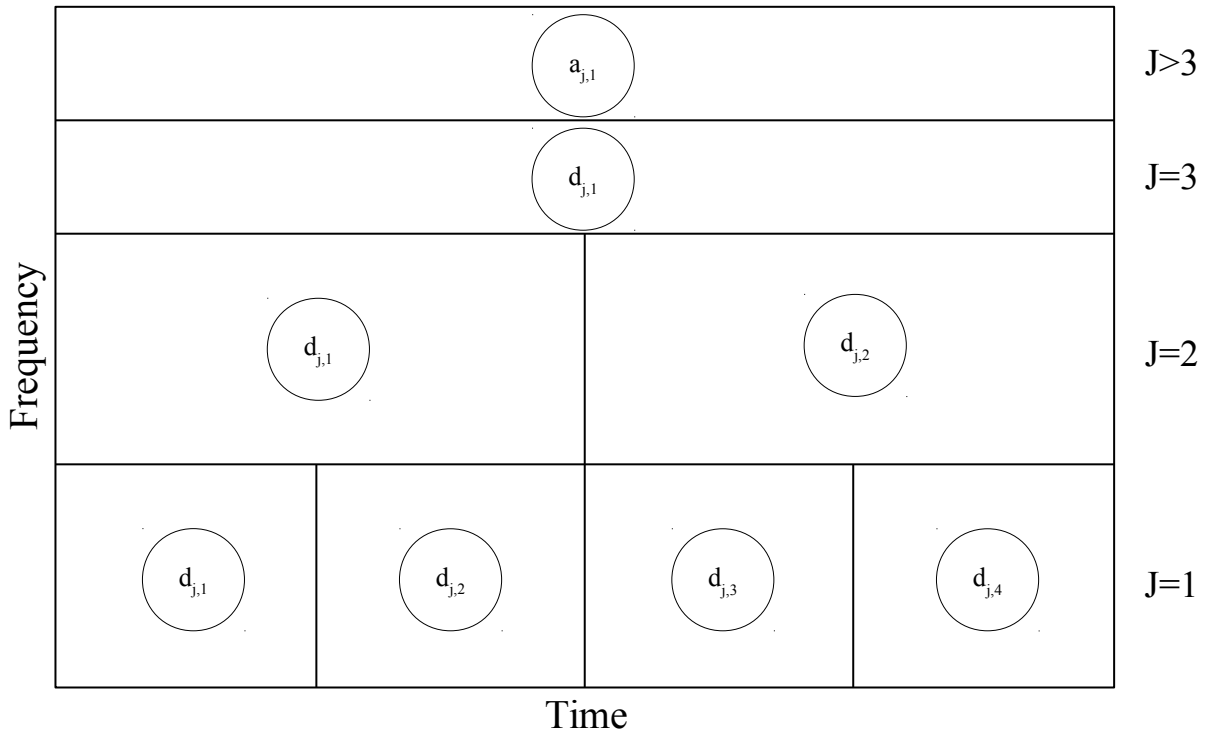


Figure 3.3: Simplified representation of the wavelet transform. J is the decomposition level, $d_{j,i}$ are the details coefficients and $a_{j,i}$ are the approximation coefficients.

where $\phi_{j,k}$ and $\psi_{i,k}$ are orthonormal basis functions, $a_{k,i}$ are the approximation coefficients and $d_{k,i}$ are the details coefficients (Reis & Saraiva 2000). To compute the wavelet transforms as well as the inverse wavelet transforms, the MATLAB® toolbox WaveLab 850 (Buckheit et al. 2005) is used.

In the methodology applied in the thesis, one takes the scores matrix described in Section 3.1 and apply the wavelet transforms to it. The scores act as a training set to choose key parameters of the wavelet transforms such as the filter, the maximum decomposition, J_{max}, and a suitable reconstruction in a way that it highlights the different dynamics, structuring the short and long range behavior of the system.

With the filter, J_{max} and reconstruction adjusted with the scores, wavelet transforms are applied to dataset $\boldsymbol{X}$ (See Figure 3.4). Then $\boldsymbol{X}$ is reconstructed in three matrices with the high, medium and low frequency signals, $\hat{\boldsymbol{X}}_{hf}$, $\hat{\boldsymbol{X}}_{mf}$ and $\hat{\boldsymbol{X}}_{lf}$, respectively. PCA is then applied again to all of the new matrices and the resulting loadings matrices are analyzed to investigate what variables are the most significant at each range of scales.

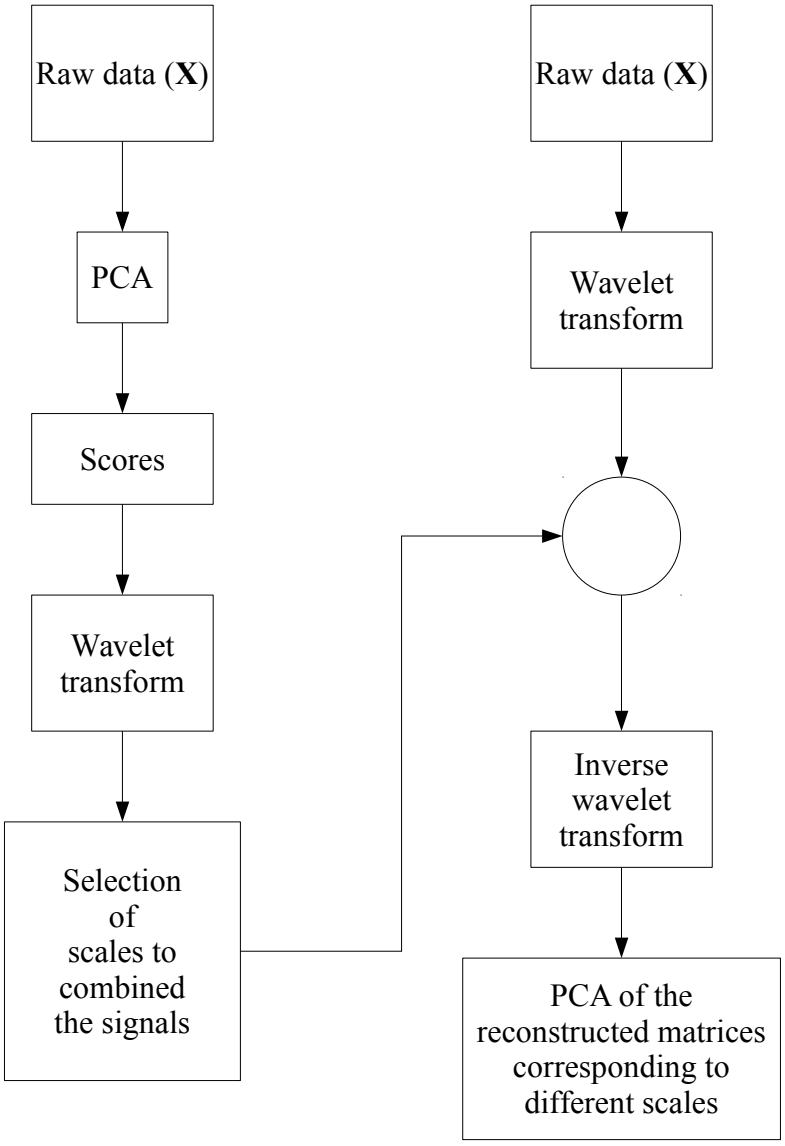


Figure 3.4: MRA methodology.

Chapter 4

Simulation of a CSTR

«All models are wrong;
some models are useful.»
*(*George E. P. Box*)*

In order to test monitoring methodologies, one needs to have absolute control over the process and access to the system's ground truth which is not available in industrial data. As proof of concept, there was the necessity of modeling a chemical system. The model needed to be simple enough to enable proper interpretation of the results and save some computational effort; yet it should contemplate the possibility to simulate a meaningful number of faults to analyze.

4.1 Reactor model

The system considered in this case study consists of a non-isothermal CSTR under feedback control based on the CSTR described by Luyben (1990) with some modifications. Although the system considered shows a quite complex dynamic under feedback control, it is easy to work with due to the reduced number of state variables (< 10). The system consists of a CSTR, where a simple first order exothermic reaction (4.1) is catalyzed by small spherical catalysts in suspension, like in a slurry reactor.

$$A \rightarrow B \tag{4.1}$$

Here it is assumed that the system should be as generic as possible, so that the results won't be defrated by the process. The introduction of the suspended catalyst, extending Luyben's (1990) system, is justified by the need of investigating the catalyst activity loss. It is assumed perfect mixing which means the concentration of the reactant A is the same in every point of the reactor's mixture. The reactor is fitted with a heating jacket. The outflow stream is in free discharge operating mode.

The main goal of the control system is to keep the liquid level at 45 cm and the temperature of the mixture at 45 °C with the help of two pneumatic control valves. The vector of state variables is given by $\boldsymbol{x} = [\ h\ C_\mathrm{A}\ T\ T_\mathrm{j}\ F_\mathrm{jv}\ \alpha_\mathrm{catalyst}\ ]^T$ and the vector of loading variables is $\boldsymbol{u}$

$= [\ F_0\ C_{A0}\ T_0\ T_{j0}\ F_j\ ]^T$, where F_0 and F_j are the manipulated variables, as it is illustrated in Figure 4.1.

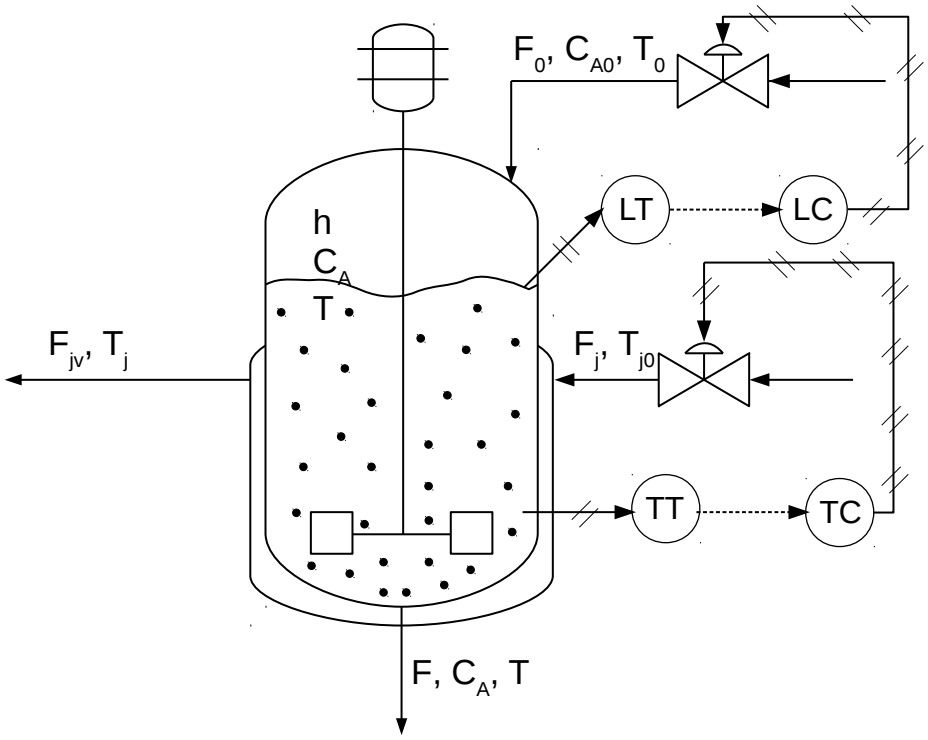


Figure 4.1: Simplified Process and Instrumentation Diagram of the simulated system.

Mass and energy balances are formulated assuming constant physical properties of compounds A and B and of the mixture (Seborg et al. 2004).

The simplified global mass balance to the reactor liquid mixture gives

$$\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{F_0 - F}{A_\mathrm{b}}\,, \tag{4.2}$$

where h is the liquid level, A_b is the cross sectional surface area of the tank, and F_0 and F are the inlet feed and outlet volumetric flow rates, respectively. The outlet stream leaves the tank by free discharge and its flow rate is calculated from

$$F = c_\mathrm{valve}\sqrt{h}\,. \tag{4.3}$$

The volume of the liquid mixture is given by

$$V = V_\mathrm{base} + A_\mathrm{b} \cdot h\,, \tag{4.4}$$

where V_base is a residual liquid volume at the bottom of the tank.

The temperature dependence of the reaction (4.1) is a first order Arrhenius' Law

$$k_\mathrm{factor} = k_0 \cdot exp\left(\frac{-E_\mathrm{a}}{T + 273.15}\right)\,, \tag{4.5}$$

where T is the mixture temperature in °C, k_0 is the pre-exponential factor and E_a is the activation energy of the reaction in investigation divided by the ideal gas constant. The catalyst activity loss balance gives

$$\frac{d\alpha_{\text{catalyst}}}{dt} = -k_d\,, \tag{4.6}$$

where α_{catalyst} is the catalyst activity factor and k_d is the catalyst deactivation rate which will be further discussed in section 4.6. The reaction rate is described as

$$r = k_{\text{factor}} \cdot C_A \cdot \alpha_{\text{catalyst}}\,. \tag{4.7}$$

where C_A is the reactant A concentration. As A is the sole reactant, its reaction rate is modeled as

$$r_A = \alpha_A \cdot r\,. \tag{4.8}$$

The partial mass balance gives

$$\frac{dC_A}{dt} = F_0 \cdot \frac{C_{A0} - C_A}{V} + r_A \tag{4.9}$$

To get the heat transfer area (A_t),

$$A_t = \pi \cdot \frac{D}{2} \cdot \left(\frac{D}{2} + 2 \cdot h\right) \cdot 1.25\,, \tag{4.10}$$

one must add the CSTR's bottom area and its lateral area. A geometry factor of 1.25 is included due to the base's round shape.

A thermodynamic system's total energy is given by the sum of kinetic energy, potential energy and internal energy. As the internal energy's contribution dwarfs both kinetic and potential energy, these two parcels can be left out. In liquids, at constant pressure, internal energy can be approximated with enthalpy. An energy balance is made which gives

$$\frac{dT}{dt} = F_0 \cdot \frac{T_0 - T}{V} + m\Delta H_R \cdot \frac{r}{\rho \cdot C_P} + U \cdot A_t \cdot \frac{T_j - T}{\rho \cdot C_P \cdot V}\,, \tag{4.11}$$

T and T_0 are the mixture's and feed flow's temperature, respectively. $m\Delta H_R$ is the reaction's enthalpy, ρ is the liquid's density and C_P is the mixture's heat capacity at constant pressure. In the last parcel of the equation, the heat transfer, U is the global heat transfer coefficient and T_j is the thermal fluid temperature in the jacket.

The thermal fluid's energy balance is given by

$$\frac{dT_j}{dt} = F_{jv} \cdot \frac{T_{j0} - T_j}{V_j} + U \cdot A_t \cdot \frac{T - T_j}{\rho_j \cdot C_{Pj} \cdot V_j}\,. \tag{4.12}$$

This balance is similar to the one in equation (4.11) except the reaction's parcel. The physical properties are from the thermal fluid. Finally, a mass balance to the thermal fluid in the jacket gives

$$F_{\mathrm{jv}} = F_{\mathrm{j}}\,, \tag{4.13}$$

where F_{jv} is the stream that leaves the jacket. As liquids are incompressible, the system is in steady-state.

The equations parameters, as well as initial values were taken from an example case study in Luyben (1990) and are presented in Table 4.1.

Table 4.1: Model parameters and initial values

Variable	Value	Variable	Value
F_0	8.0×10^{-3} m^3/min	V_{base}	5.0×10^{-3} m^3
C_{A0}	6.0 kmol/m^3	k_0	2.0×10^{10} min^{-1}
T_0	26.0 °C	E_{a}	8677.0 K
T_{j0}	75.0 °C	α_{A}	−1.0
F_{j}	8.0×10^{-3} m^3/min	ρ	1000.0 kg/m^3
h	0.45 m	C_{P}	4184.0 J/(kg K)
C_{A}	4.5 kmol/m^3	$m\Delta H_{\mathrm{R}}$	30.0 J/kmol
T	45.3 °C	U	48000.0 J/(min m^2 K)
T_{j}	57.3 °C	ρ_{j}	950.0 kg/m^3
F_{jv}	F_{j}	C_{Pj}	4800.0 J/(kg K)
$\alpha_{\mathrm{catalyst}}$	1.0	V_{j}	2.5×10^{-3} m^3
D	0.5 m	τ_{jacket}	$\frac{1}{12}$ min

In real life systems, variables do not behave deterministically. Several physical characteristics such as pipe roughness, impurities, instrumentation noise or temperature variations constitute a source of disturbances that deviate a variable from the expected value. To incorporate these disturbances in the model, a linear stochastic model is added to three load variables (C_{A0}, T_0, T_{j0}) as a linear aggregation of random shocks. One way to simulate these shocks is through times series such as a first order Autoregressive Moving-Average (ARMA) (Box et al. 1994),

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}\,, \tag{4.14}$$

where ϵ is the stochastic term, $\phi_1 \tilde{z}_{t-1}$ is the autoregressive term and $\theta_1 \epsilon_{t-1}$ is the moving average term. ϵ was generated multiplying a MATLAB® function, that returns pseudorandom values drawn from the standard normal distribution, by an adjusted constant equal to the desired standard deviation of the normal distribution. These values are presented in Table 4.2.

Manual tuning was then carried out to adjust ϕ and θ (See Table 4.3). These terms vary between 0 and 1 and determine how much the last iteration is valued and how much the last iteration's stochastic term is valued for ϕ and θ, respectively.

The values on Table 4.3 are assumed by the author as a model parameter to simulate realistic stochastic behavior.

Table 4.2: Standard deviation of the noise introduced in the ARMA model.

Variable	Value
C_{A0}	0.1 kmol m^{-3}
T_0	0.5 °C
T_{j0}	0.5 °C

Table 4.3: ARMA parameters.

Parameter	Value
$\phi_{C_{A0}}$	0.90
ϕ_{T_0}	0.90
$\phi_{T_{j0}}$	0.90
$\theta_{C_{A0}}$	0.75
θ_{T_0}	0.75
$\theta_{T_{j0}}$	0.75

4.2 Valve model

In industrial processes, the importance of control hardware cannot be overlooked. For example, it is the final control element that executes the controller's command. Therefore reliable instrumentation hardware is essential to ensure the system's stability and normal operation.

In this CSTR system, there are two final control elements, which are control valves. The dynamic behavior of a pneumatic valve is best described with an underdamped second order system (Smith & Corripio 1985).

Stephanopoulos (1984) models an underdamped second order system as

$$\lambda(t) = K_v \cdot \left[1 - \frac{\cos(\omega t) - \sin(\omega t) \cdot \zeta}{\sqrt{1-\zeta^2}} \cdot e^{\frac{-\zeta t}{\tau_V}}\right] \cdot \Delta u\,, \tag{4.15}$$

$$\omega = \frac{\sqrt{1-\zeta^2}}{\tau_V}\,, \tag{4.16}$$

$$\phi = \tan^{-1}\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right). \tag{4.17}$$

The equation (4.15) describes the response of valve's position to an input signal from a controller, where λ is the valve's position, K_v is the valve's stationary gain, τ_V is the valve's time constant and ζ is the damping factor. Δu is input signal.

For both valves, parameter values are presented in Table 4.4. These parameters are considered as model inputs to simulate an acceptable behavior of a control valve. Step tests were conducted to verify the valve model behavior, as shown in Figure 4.2.

Table 4.4: Valve model parameters for both valves.

Parameter	Value
ζ	0.5
K_{v}	1.0
τ_{V}	1.0

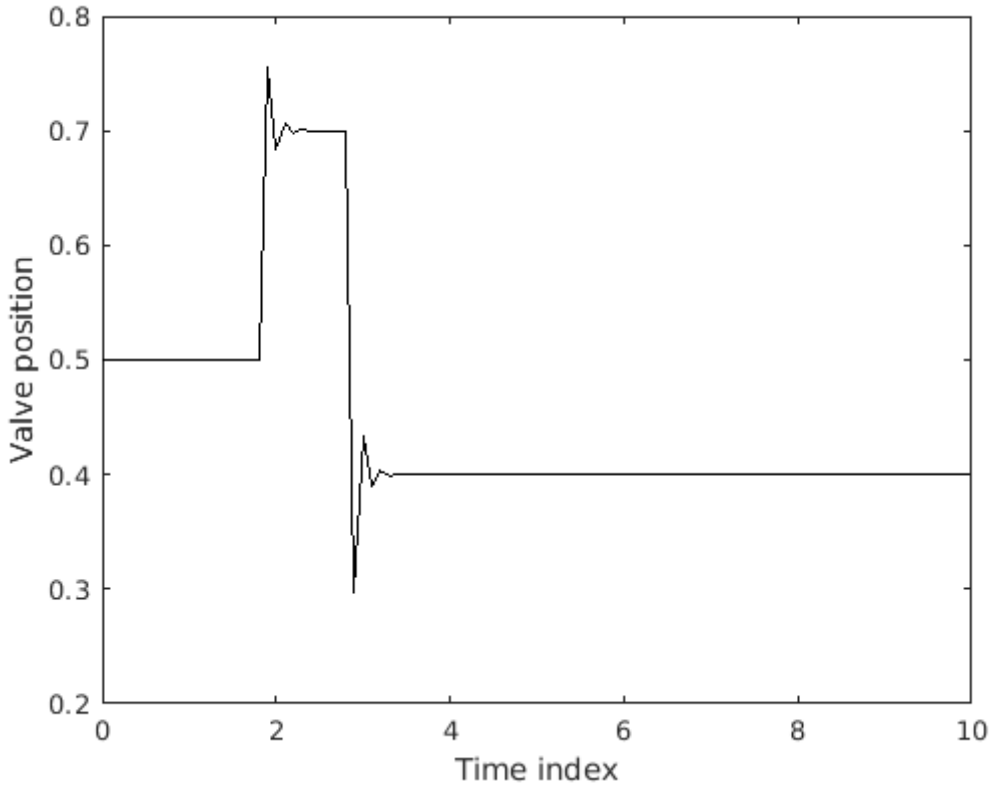


Figure 4.2: Valve's response to a sequence of two step tests in the command position signal.

Finally, the valves positions enter in equations

$$F_0 = \lambda_{\mathrm{F0}} \cdot F_{0\mathrm{max}}\,, \tag{4.18}$$

where $F_{0\mathrm{max}}$ is the feed flow's maximum output, and

$$F_{\mathrm{j}} = \lambda_{\mathrm{Fj}} \cdot F_{\mathrm{jmax}} \tag{4.19}$$

where F_{jmax} is the thermal fluid's maximum output. These equations (4.18) (4.19) describe the feed and the thermal fluid flow rates. The starting point of both valves is 0.50.

4.3 Controller design

A chemical process has to fulfill the requirements imposed by designers and process engineers to achieve the product quality needed, in spite of the ever changing economic and

technical conditions and external influences. A control system plays an essential role in the operation of a chemical process at the specified operating conditions, by compensating the effect of external disturbances, ensuring its stability, seeking its optimal performance (Stephanopoulos 1984).

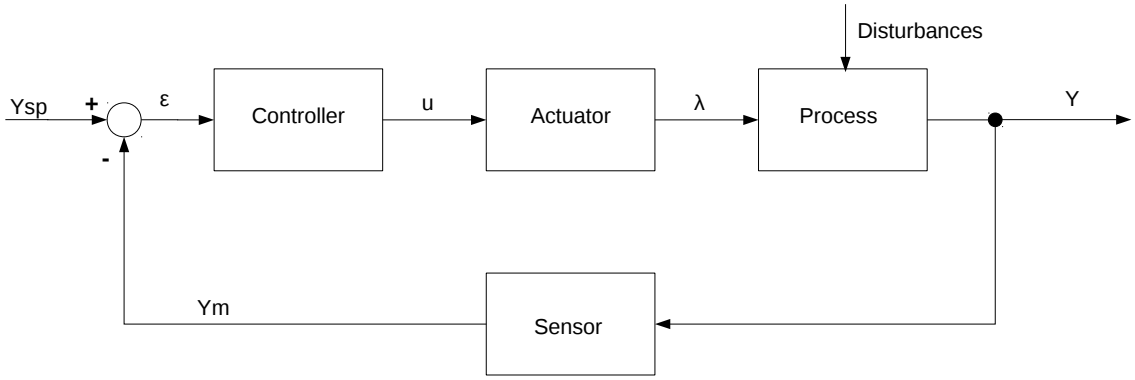


Figure 4.3: General structure of feedback control configurations.

Load variables introduce variability and disturbances in the process. Therefore to keep the controlled variables within specifications a controller must be designed and installed. Being the most used controller technology in industry, with a percentage of about 90% (Åström & Hägglund 2001), for its feasibility and easy implementation, a feedback (Figure 4.3) Proportional, Integral and Derivative (PID) controller is considered for this case study.

To design controllers, first one must find the transfer function in open cycle of each variable one wishes to control. Step tests of known magnitude are conducted.

The response of the tank level, in Figure 4.4a, exhibits a first order behavior, whereas the dynamic behavior of the mixture temperature, in Figure 4.4b, can be approximated by a first order plus delay behavior. Characterizing the step tests responses and using the Sundaresan & Krishnaswamy method (Seborg et al. 2004), the following transfers functions are obtained to describe the dynamics of the level and temperature, respectively:

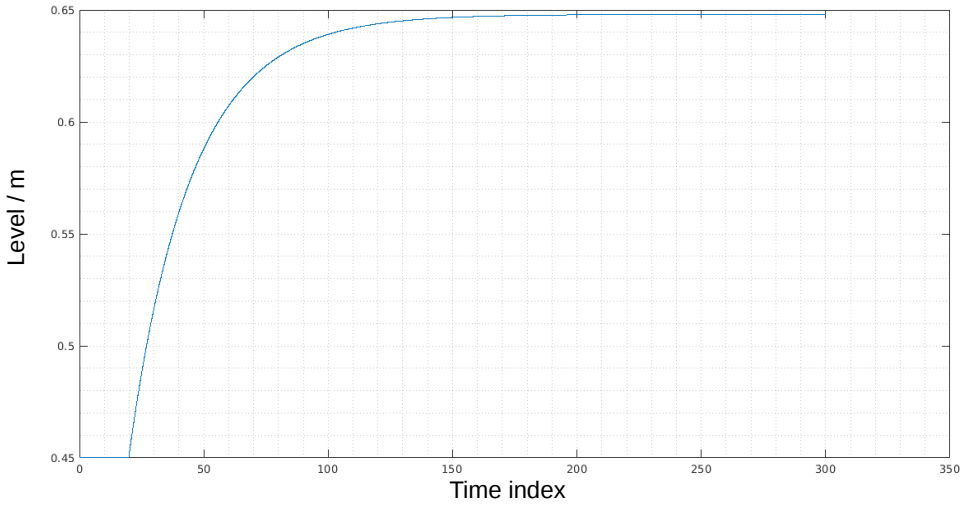
$$G_a(s) = \frac{2}{25.4s + 1} \tag{4.20}$$

$$G_b(s) = \frac{12 \cdot e^{-1.45s}}{5.63s + 1} \tag{4.21}$$

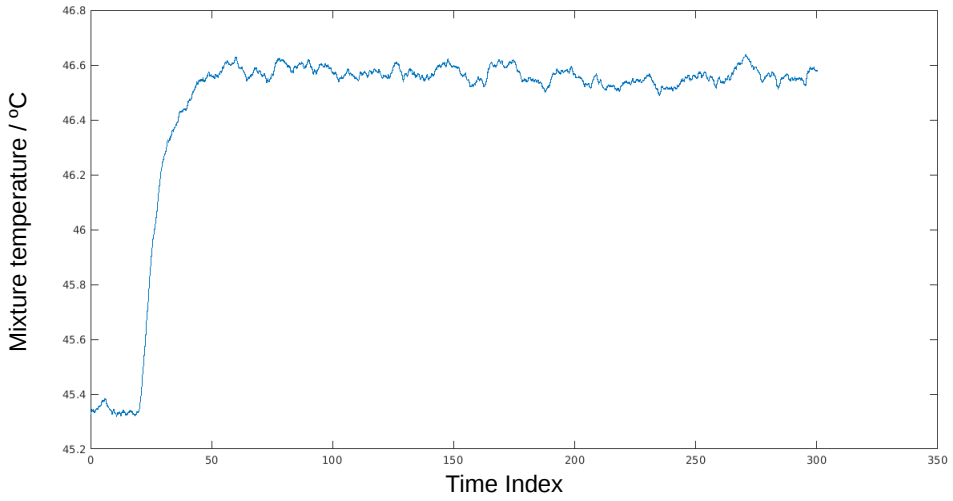
As it was stated, the level system has first order behavior. In these conditions the offset is acceptable, hence a controller with just the Proportional (P) mode can be used with a high stationary constant. Manual tunning led to $K_c = 5\ \mathrm{m}^{-1}$. Thus, the level controller transfer function is simply:

$$G_{Ca}(s) = K_c = 5 \tag{4.22}$$

The mixture temperature system, having a first order with delay dynamic, the offset caused by a P controller is not acceptable, therefore a Proportional and Integral (PI)



(a) Level's response.



(b) Temperature's response.

Figure 4.4: Level and temperature's response to a 10% increase in the aperture of the feed flow valve and the thermal fluid's flow valve, respectively.

controller was designed. To design this controller, the IMC method (Seborg et al. 2004) is used. This leads to the following temperature controller transfer function:

$$G_{Cb}(s) = K_c \left(1 + \frac{1}{\tau_I s}\right) = 0.1618 \times \left(1 + \frac{1}{5.63s}\right) \tag{4.23}$$

Table 4.5: Controller parameters.

Controlled variable	Controller	K_c	τ_I
Tank level	P	5.0 m^{-1}	-
Mixture temperature	PI	0.1618 $°C^{-1}$	5.63 min

4.4 Simulation strategy

In general, real systems evolve and suffer disturbances between control actions. Therefore it is important to establish a suitable strategy to realistically simulate control actions and dynamics of loads. The control cycle should take place at short time intervals enough to keep the controlled variables in their target values, but long enough to avoid unnecessary valve stress. The control step must also be smaller than the residence time of liquid in the tank, which is equal to 11.7 min in this simulation, to keep the system stable. After testing by simulation a range of possibilities, the value of 0.2 minutes was the selected interval for the control step.

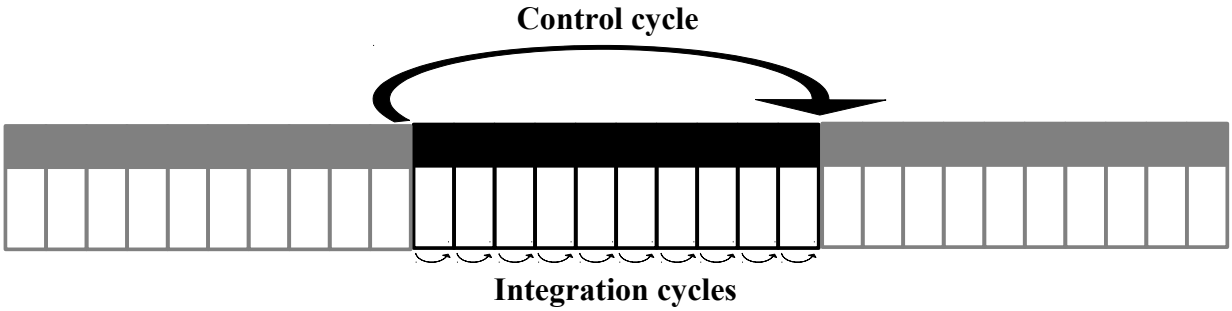


Figure 4.5: Simulation cycles architecture.

Furthermore, to simulate the process evolution within each control cycle, an inner integration cycle for simulating the CSTR with a smaller sampling time period was considered as presented in Figure 4.5. This inner loop has a step ten times lower than the control step. The goal of this simulation scheme is to provide a more realistic scenario in the simulation of the CSTR.

4.5 Abrupt faults

Conceptually, chemical processes are designed to run smoothly 24 hours a day for years. In reality, during the process operation problems keep arising and red blinking lights are ubiquitous on control panels. Since part of this project's goal is to detect and diagnose abrupt faults, these must be included in the model formulation in order to generate the data needed to conduct the analysis. This work is focused on three main failure sources: sensors, process and valves.

4.5.1 Sensor failure

Sensors are engineers' eyes and ears in a chemical process. Through them a system can be monitored and their outputs fed to digital controllers. This means that if a controlled variable sensor is not well calibrated, the controller will issue a wrong action. To study this effect, a bias was simulated on two critical sensors which are the level and mixture temperature, and on one non-critical sensor, the jacket temperature sensor. The first two sensors were dubbed critical because they measure controlled variables and their performance directly impacts the controller's action. The latter sensor was dubbed non-critical because it is an output variable without a direct relation to the controllers.

Each sensor failure was programmed to begin at the 150^{th} iteration, this is 30 minutes of operation. Three intensity levels where defined for each bias based on the standard deviation of values measured in the training set, NOC conditions. (See Table 4.6).

Table 4.6: Biases intensity defined for each sensor fault.

	Sensor h / m	**Sensor** T / °C	**Sensor** T_j / °C
σ	0.0025	0.2572	0.3136
.....	.....	.....	.....
Intensity 1	$-1 \cdot \sigma$	$-1 \cdot \sigma$	$-1 \cdot \sigma$
Intensity 2	$-2 \cdot \sigma$	$-2 \cdot \sigma$	$-2 \cdot \sigma$
Intensity 3	$-3 \cdot \sigma$	$-3 \cdot \sigma$	$-3 \cdot \sigma$

4.5.2 Process failure

Process safety must be a priority on a chemical site. Unfortunately, incidents occur and the control room must be aware of them as soon as possible. To account for these type of incidents, small misadventures were simulated at the 150^{th} iteration or 30 minutes of operation. The incident scenarios considered in this study are pipe and tank ruptures. Three different scenarios were devised.

In the first scenario, a worker driving a forklift unwittingly causes a small rupture in the feed flow pipe causing the maximum inflow, before the valve, to be reduced by 5% for intensity 1, 15% for intensity 2 and 25% for intensity 3. This means that in equation (4.18), F_{0max} decreases.

In the second scenario, another worker driving a forklift gets distracted and crashes into the tank reactor causing a small orifice with sharp borders in the heating jacket, at 20 cm from the top of the jacket. In the simulation framework, equation (4.13) is changed to:

$$\frac{dF_{\text{jv}}}{dt} = \frac{-F_{\text{jv}} - \sqrt{2g \times 0.2} \cdot A_{jorifice} \cdot C_0 + F_{\text{j}}}{\tau_{\text{jacket}}} \tag{4.24}$$

where $A_{jorifice}$ is the area of the orifice and C_0 is the discharge coefficient. $A_{jorifice}$ is equal to 3×10^{-5} m^2 for intensity 1, 3×10^{-4} m^2 for intensity 2 and 3×10^{-3} m^2 for intensity 3. C_0 is equal to 0.61 because the orifice has sharp borders.

In the third scenario, the same worker crashes again into the tank reactor with a forklift. This time, besides rupturing the heating jacket, he also ruptures the tank reactor, 10 cm above the bottom of the tank, causing it to leak. Equation (4.24) maintains the same form and equation (4.3) is changed to:

$$F = c_{\text{valve}} \cdot \sqrt{h} + \sqrt{2g \cdot (h - 0.1)} \cdot A_{orifice} \cdot C_0 \tag{4.25}$$

$A_{orifice}$ is equal to 3×10^{-5} m^2 for intensity 1, 3×10^{-4} m^2 for intensity 2 and 3×10^{-3} m^2 for intensity 3. C_0 maintains the same value as before.

The process faults scenarios are summarized in Table 4.7.

Table 4.7: Summary of process faults.

	Max. feed flow / (m^3/min)	**Jacket orifice area** / m^2	**Tank orifice area** / m^2
Intensity 1	0.95×0.016	3×10^{-5}	3×10^{-5}
Intensity 2	0.85×0.016	3×10^{-4}	3×10^{-4}
Intensity 3	0.75×0.016	3×10^{-3}	3×10^{-3}

4.5.3 Valve failure

Good process control, as it has been stated, needs final control elements operating well. However, this is not always the case. It is important for operators to detect easily and rapidly faulty valves. Common problems in control valves are oversizing, undersizing, corrosion, defective diaphragm, hysteresis, deadband, valve lock, valve offset, stiction (Choudhury et al. 2008). The last three problems are addressed and simulated in this study. In the simulation framework they are triggered at the 150^{th} iteration, this is, at 30 minutes of operation.

Valve lock happens when corrosion locks the valve in place, preventing it to follow the controller orders. Other times, a valve simply locks because there is a problem in the transmission line and the order to move does not reach the valve. In any case, it is important to know when this happens so the problem can be solved. To simulate this scenario, one simply overwrites the controller orders so that the valve stays in place. For each intensity level the following events are considered. For intensity 1, the level control valve is locked. For intensity 2 the temperature control valve is locked. Finally, for intensity 3 both valves are locked.

Under the effect of corrosion and/or material depositions, valve actuators start to present offset (Smith & Corripio 1985). These phenomena prevent the valve to reach the desired aperture, making the control system less responsive and efficient.

In mathematical terms, the valve offset can be simulated by a decrease in its stationary gain. So to simulate this failure on both valves, K_v is reduced from 1 to 0.75 for intensity 1, to 0.50 for intensity 2 and to 0.25 for intensity 3.

Static friction or stiction is a phenomenon that occurs in control valves that prevents them from moving, accumulating strength and suddenly releasing it. Many definitions have been proposed over the years. The definition proposed by Ruel (2000) is considered to be one of the closest to the stiction measured online:

> «a combination of the words stick and friction, created to emphasize the difference between static and dynamic friction. Stiction exists when the static (starting) friction exceeds the dynamic (moving) friction inside the valve. Stiction describes the valve's stem (or shaft) sticking when small changes are attempted. Friction of a moving object is less than when it is stationary. Stiction can keep the stem from moving for small control input changes, and then the stem moves when there is enough force to free it. The result of stiction is that the force required to get the stem to move is more than is required to go to the desired stem position. In presence of stiction, the movement is jumpy.»

The stiction phenomenon is a ubiquitous problem in control valves which is hard to detect and difficult to model. That is why there has been great interest in being able to detect stiction with data-driven methodologies. Figure 4.6 shows a simplified version of the

algorithm proposed by Choudhury et al. (2005) to simulate stiction.

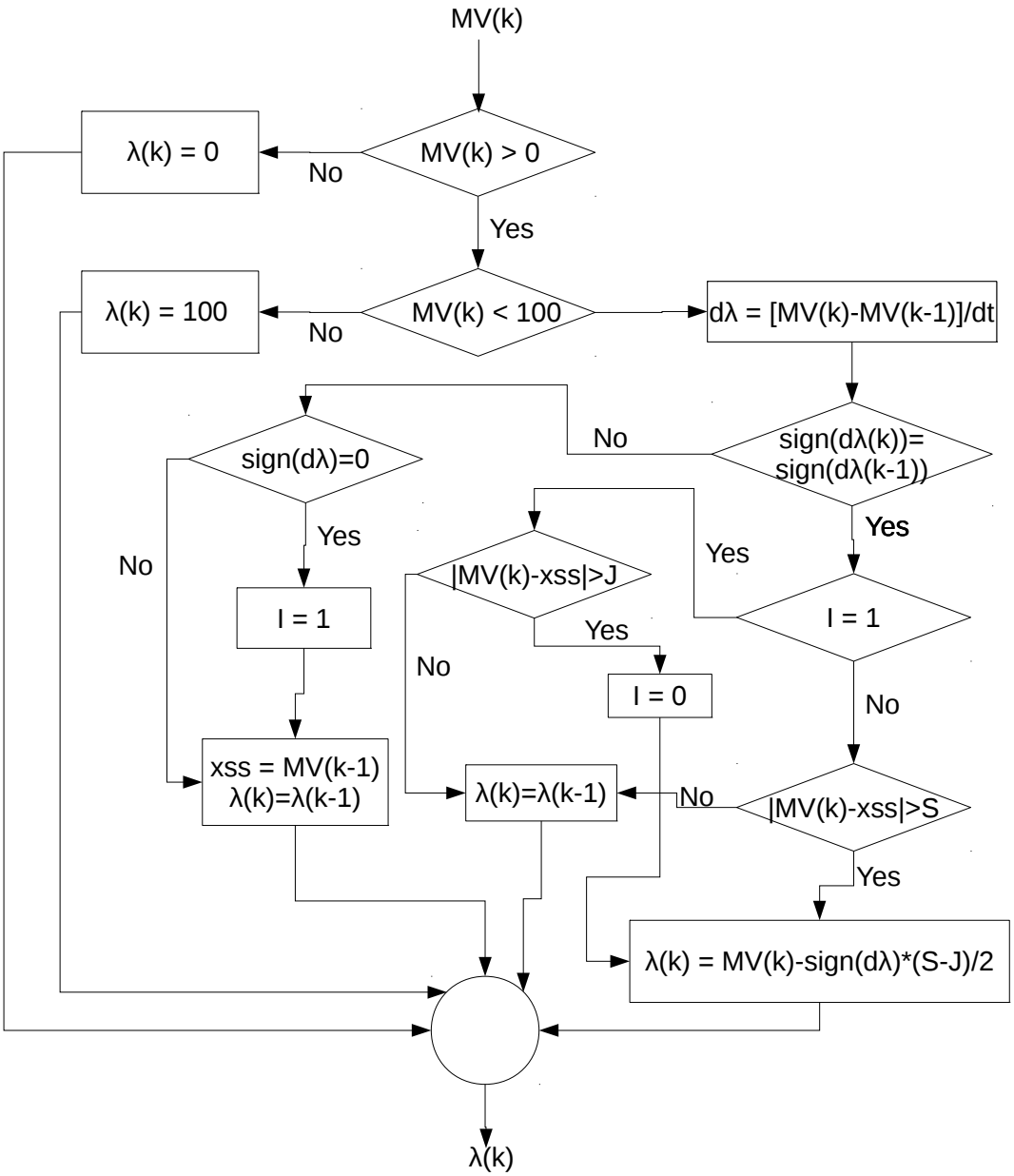


Figure 4.6: Logic flow chart for the data-driven stiction model. I is the locking indicator, if it is 1 the valve sticks, otherwise it moves. xss stores the input signal for when the valve gets stuck (adapted from Choudhury et al. (2005)).

Incorporating this algorithm in the model numerical framework was a challenge and involved some fine tunning. After calculating the valve's stem position with equation (4.15), the resulting value is used in equation

$$\lambda_{stiction} = \lambda \pm \frac{S - J}{2}, \tag{4.26}$$

where S and J are model parameters that must be tuned to the desired level and type of

stiction. S is the stickband plus deadband which characterizes the valve behavior while not moving, though the controller orders otherwise. J is the slipjump which represents the sudden release of energy stored in the valve aperture mechanism. These concepts are illustrated in Figure 4.7.

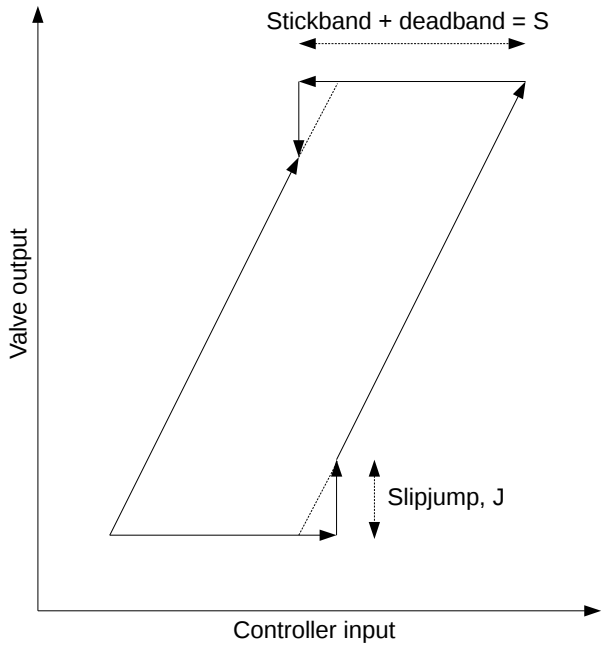


Figure 4.7: Typical input–output behavior of a sticky valve (adapted from Choudhury et al. (2005)).

If J is equal to S or they are both zero, as it happens for the scenario with intensity 1, the algorithm produces a pure stick-slip behavior with no offset. If J is greater than S, overshoot occurs. If J is smaller than S, undershoot occurs which means the valve output may never reach the valve input (Choudhury et al. 2005). For the scenario with intensity 2, stiction was simulated without slipjump, thus $J = 0$ and $S = 0.02$ in both valves. For the scenario with intensity 3, stiction with undershoot was achieved setting $S = 0.02$ and $J = 0.01$.

Table 4.8: Summary of valve faults.

	Valve locked	K_v	**Stiction type**	S	J
Intensity 1	Feed flow	0.75	Pure stick-slip	0.00	0.00
Intensity 2	Thermal fluid flow	0.50	+ Deadband	0.02	0.00
Intensity 3	Both	0.25	+ Undershoot	0.02	0.01

4.6 Simulating process & equipment degradation

In this section, the inclusion of slower process dynamics in the model simulation is described. Although many different slow dynamic phenomena can occur in a chemical process unit, only three are considered in this thesis: external temperature variations,

fouling formation and catalyst activity loss. These are three slow dynamic phenomena that are widely reported in the literature on these subjects.

4.6.1 External temperature influence

The external temperature variation is modeled using the temperature profile in Coimbra from 12pm of the 26^{th} of April 2018 to 12pm of the 27^{th} of April 2018 (OpenWeatherMap 2018). This data was adjusted to a sum of sines, the wave amplitude was tweaked to accommodate fluid's thermal inertia and a stochastic term which gives

$$T_{external} = 0.05\epsilon + 0.75\sin(0.003536t - 3.266) + 0.75\sin(0.0001t - 3)\,, \tag{4.27}$$

where ϵ is the stochastic term generated by a MATLAB® function that returns pseudo-random values drawn from the standard normal distribution, t is time in minutes and $T_{external}$ is the temperature in °C.

4.6.2 Fouling

Heat transfer is an ordinary phenomenon in chemical industry and when its efficiency reduces thousands of dollars are lost in fuel. The accumulation of deposits on the walls of heat exchanging surfaces leads to extra resistance to heat transfer, *ergo* reducing heat transfer efficiency, and it is called fouling (Bott 1995).

Fouling resistance, R_f, can be defined by the difference between the global heat transfer coefficient at a given time and that coefficient at $t = 0$ (Yeap et al. 2004, Schreier & Fryer 1995) which gives

$$R_\mathrm{f} = \frac{1}{U(t)} - \frac{1}{U_0}\,. \tag{4.28}$$

To know the value of R_f a mass value is performed,

$$\frac{dR_\mathrm{f}}{dt} = m_\mathrm{d} - m_\mathrm{r}\,, \tag{4.29}$$

where m_d is the deposit rate and m_r is deposit dragged away.

Considering the m_d as constant and the m_r as negligible (Melo et al. 1988), fouling resistance varies linearly with time, which gives

$$R_\mathrm{f} = m_\mathrm{d} \cdot t\,, \tag{4.30}$$

where m_d is an adjustable parameter. Rearranging equation (4.28) and substituting equation (4.30) results in

$$U(t) = \left(m_\mathrm{d} \cdot t + \frac{1}{U_0}\right)^{-1}\,. \tag{4.31}$$

Being calcium sulfate a common foulant in most industrial water systems (Muller-Steinhagen 2000), its deposit rate, m_d, is used to fit equation (4.31). The experimental value is

$$m_d = 3.33 \times 10^{-11}\ \mathrm{m^2\ ^\circ C\ (J/min)^{-1} min^{-1}} \tag{4.32}$$

and was obtained by Peyghambarzadeh et al. (2012).

4.6.3 Catalyst activity loss

Catalysts are an important degree of freedom when operating or designing a chemical process. They allow reactions to occur at milder conditions, reducing operating costs as long as they remain active. As time goes by catalysts lose activity resulting in slower reaction rates and/or lower selectivity. Many factors can contribute to catalyst activity loss, being poisoning, sintering and coking the main ones (Rase 1990). In this work, catalyst activity loss will be treated in an independent way as is the system's reaction.

The catalyst activity factor ($\alpha_{\mathrm{catalyst}}$), present in equations 4.6 and 4.7, is described by

$$\alpha_{\mathrm{catalyst}} = \frac{\text{reactant A conversion rate}}{\text{reactant A conversion rate in a fresh catalyst}} \tag{4.33}$$

as defined by Levenspiel (1999). $\alpha_{\mathrm{catalyst}}$ varies between 1, for a fresh catalyst, and 0, for complete loss of activity, at a velocity linearly dependent of its deactivation rate, k_d.

The literature is very rich on data about experimental deactivation rates on petrochemical processes, from asphaltene coking kinetics (Rahmani et al. 2003) to coking in Fluid Catalytic Crackers (Xu et al. 2004). The problem with these harsh petrochemical conditions is that catalyst activity loss is extremely rapid and this study aims to detect softer catalyst activity loss dynamics. A milder deactivation rate, describing catalyst poisoning by a feed impurity

$$k_d = 5.77 \times 10^{-7}\ \mathrm{min^{-1}} \tag{4.34}$$

was found in Levenspiel (1999) and it was chosen to the model framework.

Chapter 5

Results and discussion

«It is the mark of a truly intelligent person to be moved by statistics.»
*(*George Bernard Shaw*)*

In this chapter, results are presented and discussed. As it has been the rule throughout the thesis, results regarding abrupt faults are presented first, followed by slower dynamics. The data generated in NOC will be followed by the PCA model with some of the considerations outlined. Then computation of multivariate control charts parameters are displayed. Finally, fault diagnosis will be addressed with contribution plots of all faults. When its comes to slower degradation dynamics, the presentation of results will follow chapter's 3 outline.

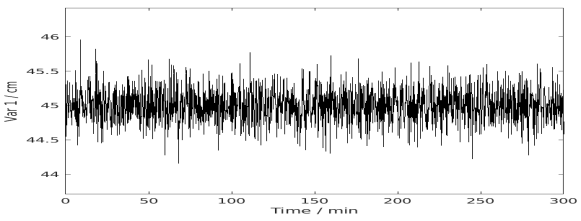
5.1 Abrupt faults monitoring and diagnosis

Following the MSPC-PCA methodology, described in section 2.2, data representing NOC operations was generated for Phase I stage a) from the system simulated in chapter 4. Data represents 300 minutes of operation in which the tank level, the concentration of A in the mixture, the mixture temperature and the jacket temperature are the measured state variables. The feed's flow rate, concentration of A and temperature as well as the thermal fluid's flow rate and temperature are the measured load variables (See Figure 5.1). The process & equipment degradation, such as the influence of external temperature variations, catalyst activity loss and fouling, was not simulated for this dataset, given the short horizon of the analysis.

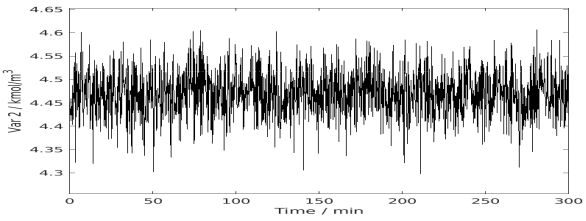
The data matrix $\boldsymbol{X}_{n\times m}$ has $n = 1501$ observations and $m = 9$. In the plots presented in this chapter variables are designated by indexes as follows:

- Variable (var) 1 represents the tank level;
- var 2 represents the concentration of A in the mixture;
- var 3 represents the mixture temperature;
- var 4 represents the jacket temperature;
- var 5 represents the feed flow rate;

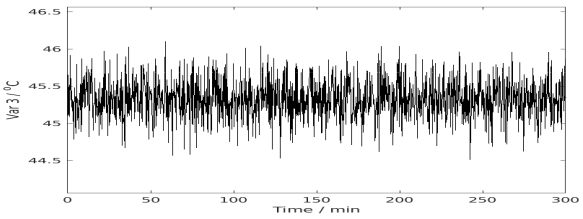
- var 6 represents the feed concentration of A;
- var 7 represents the feed temperature;
- var 8 represents the thermal fluid's temperature before the valve;
- var 9 represents the thermal fluid's flow rate.



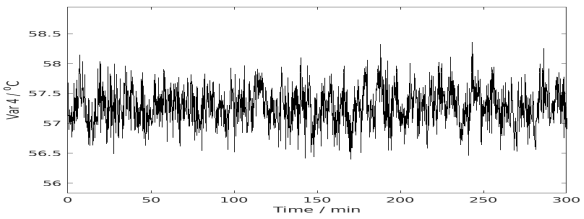
(a) Level measurements.



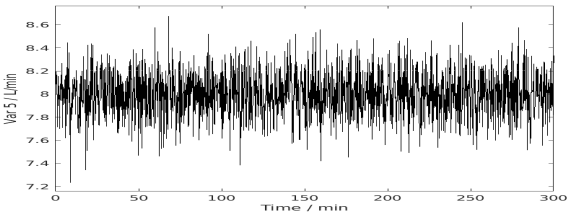
(b) Concentration of A measurements.



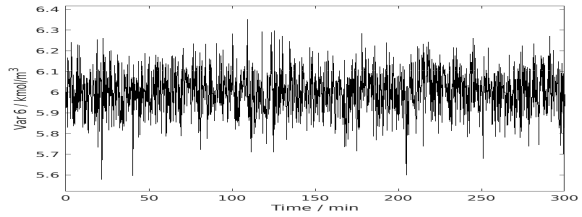
(c) Mixture temperature measurements.



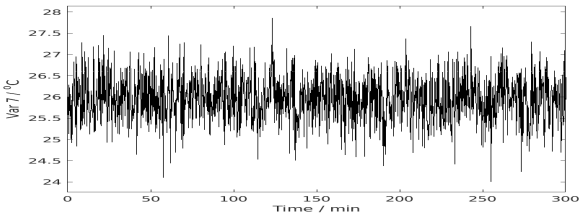
(d) Jacket temperature measurements.



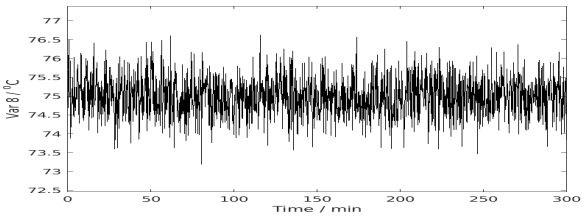
(e) Feed flow rate measurements.



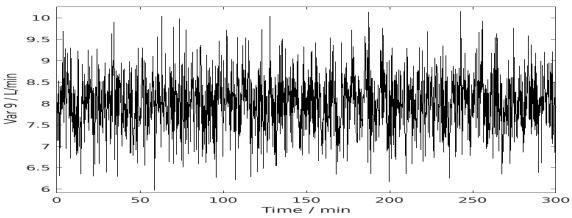
(f) Feed concentration of A measurements.



(g) Feed temperature measurements.



(h) Feed jacket temperature measurements.



(i) Thermal fluid flow rate measurements.

Figure 5.1: Measurement signals of variables in NOC.

With this NOC dataset, Phase I stage a). When applying PCA to $\boldsymbol{X}$, one must choose how many PCs should be included in the model. The strategy used was the scree test.

Figure 5.2 shows a Scree Plot where the eigenvalues drawn from the co-variance matrix of autoscalled $\boldsymbol{X}$ are sorted by magnitude and data variability explained. It is observable

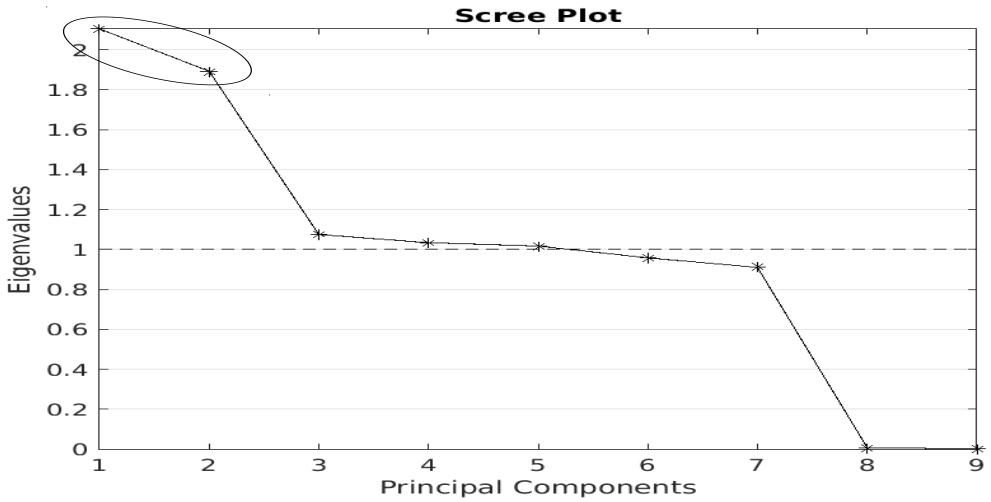


Figure 5.2: Scree Plot with the eigenvalues from data matrix $\boldsymbol{X}$ in NOC.

that the first two PCs clearly explain more data variability than the subsequent five PCs and that the last two PCs have almost no weight in data variability. Therefore, the PCA model is built using only the first two PCs which together explain 44.41% of data variability.

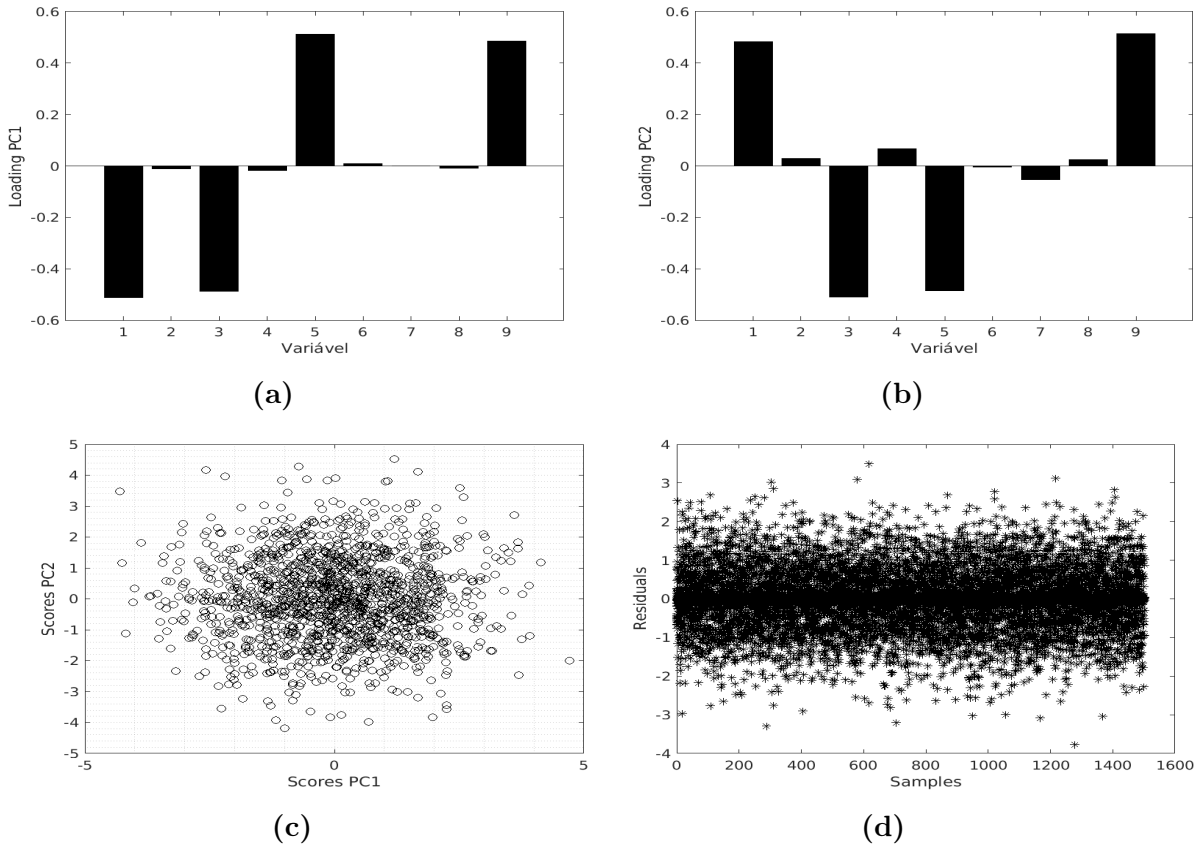
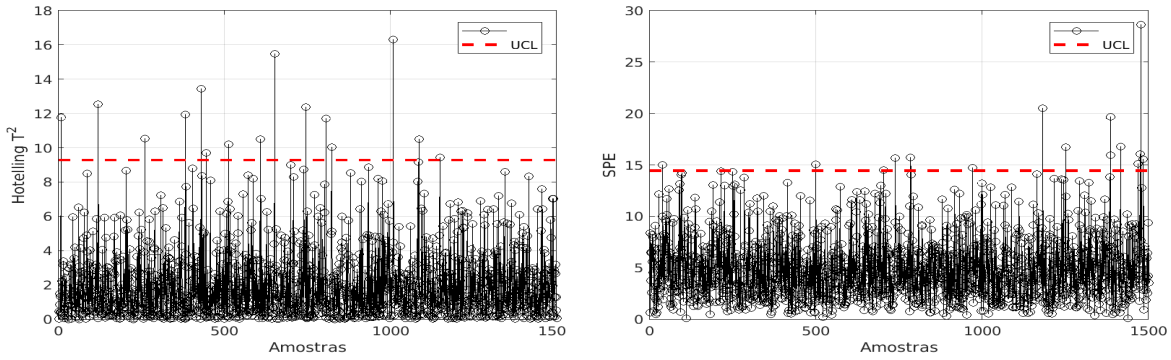


Figure 5.3: a) PC1 loadings, b) PC2 loadings, c) scores of PC1 against scores of PC2 and d) residuals from the PCA model

In Figures 5.3a and 5.3b, it is verifiable that vars 1, 3, 5, and 9 are responsible by the variability of the data along PC1 and PC2. This is expected since vars 1 and 3 - tank

level and mixture temperature - are the controlled variables and vars 4 and 9 - feed flow rate and thermal fluid flow rate - are the manipulated variables.

Figure 5.3c shows the scores of PC1 against PC2 concentrated around the origin, evenly dispersed along both axis and without any special cluster. This means process variables are variating around an average value and the system is apparently stable. Figure 5.3d shows the residues with no significant outliers, indicating the system is stable even when looking outside the PCA model. Here ends Phase I stage a).



(a) **(b)**

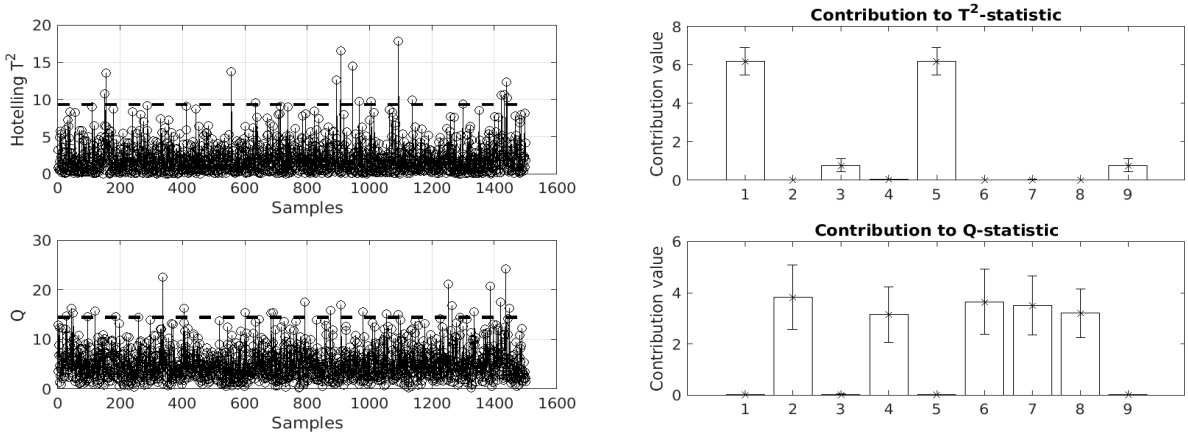
Figure 5.4: PCA-MSPC control charts, a) Hotelling's T^2 and b) Q-statistic.

For Phase I stage b), a validation set $\boldsymbol{X}_{1501\times 9}$ representing NOC operations is used to compute the PCA-MSPC control charts with $\alpha = 0.01$, aiming at a FAR of 0.01 as well (See Figure 5.4). The UCLs are 9.2941 for T^2 and 14.4375 for Q. This ends Phase I stage b).

For Phase II, fault detection, a test set is needed. Therefore, 100 runs representing 300 minutes of operation were simulated for each fault and each fault intensity, totalizing 2700 data matrices with 1501 observations of 9 variables. MSPC charts of each fault in its maximum intensity will be presented alongside contribution plots for fault diagnosis. It was decided to present only the maximum intensity faults because these examples are easier to draw conclusions from. The contributions are the average of the 100 runs with the respective three times the standard deviation of the average computed as $\pm\frac{3\hat{\sigma}}{\sqrt{n}}$. The MSPC charts displayed are just one of the 100 runs, but the overall performance of these charts is evaluated in appendix B and an example of charts of intensities 1 and 2 are presented in appendix A which also contain the contribution plots of intensities 1 and 2.

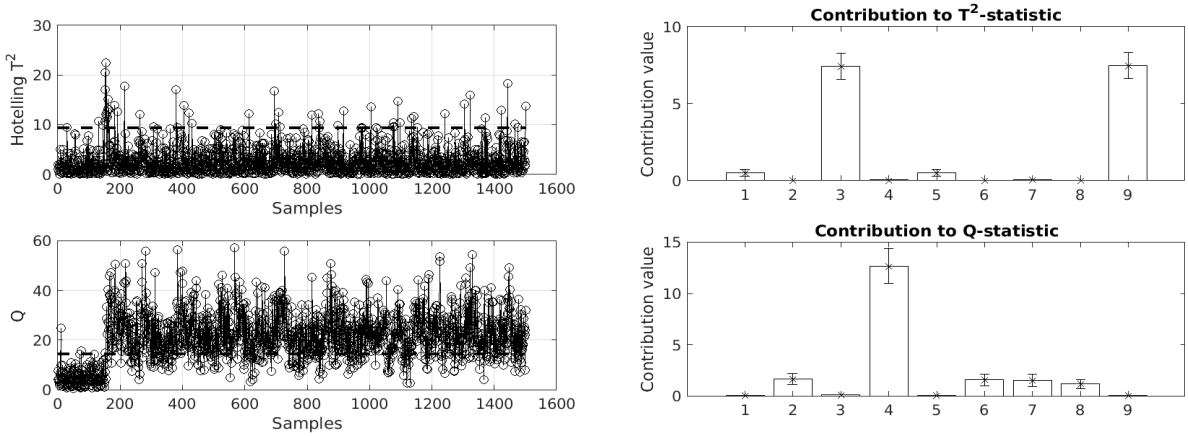
Level sensor fault is somewhat visible with the Q-statistics (See Figure 5.5a) given that the variables in the model are either not affected, vars 3 and 9, or the fault is masked by the controller's action, vars 1 and 5. The different magnitudes did not cause a big impact on the charts which indicates a certain robustness of the system level-wise (See Figure A.1).

Figure 5.5b shows indication that vars 1 and 5 are deviating from their NOC operation in T^2 contribution. Q contributions feel some of effect of the level sensor fault, although all variables are evenly affected. The contribution plot analysis for the level sensor fault points to a problem in the level and feed flow rate vars that has an influence on other measured variables. With some process knowledge one can infer that it should be the level, an important controlled variable.



(a) **(b)**

Figure 5.5: Level sensor fault - intensity 3 - a) MSPC control charts and b) contribution plots.



(a) **(b)**

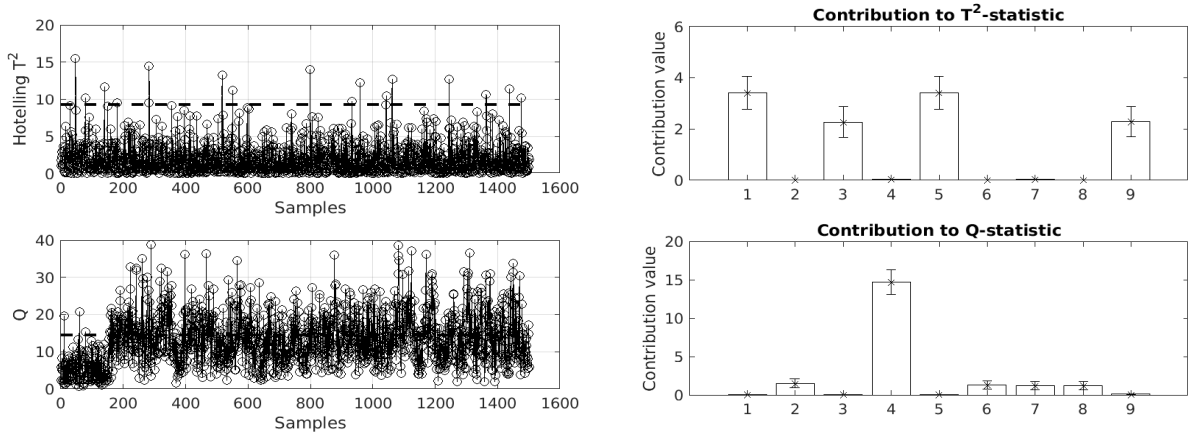
Figure 5.6: Mixture temperature sensor fault - intensity 3 - a) MSPC control charts and b) contribution plots.

With the mixture temperature sensor fault, the controller's action also masks the effect on the T^2 control chart. The fault, however, is visible on Figure 5.6a in the Q chart and it increases steadily with the fault intensity (See Figure A.2).

The contribution plots, in Figure 5.6b, show deviations in vars 3, 4 and 9, all temperature related variables.

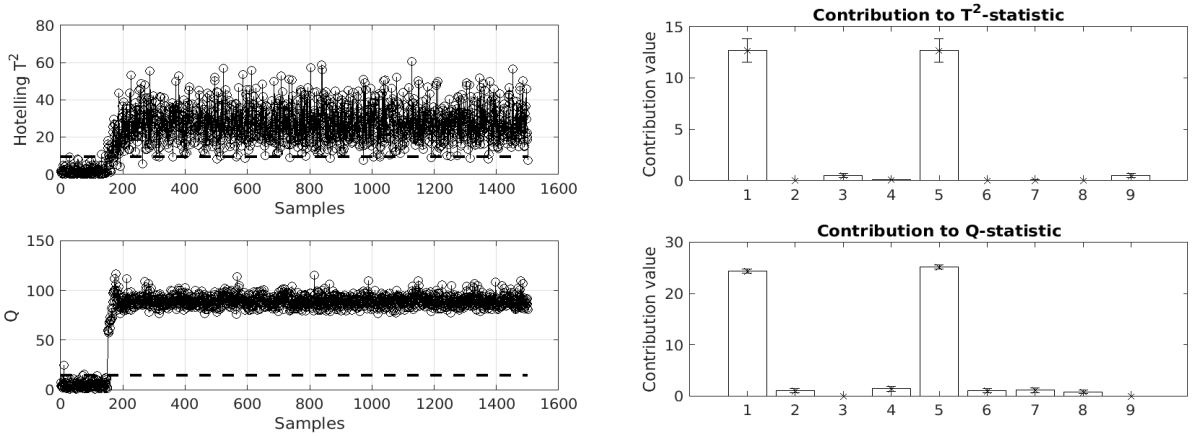
Figure 5.7a shows good fault detection in the Q control chart that increases with fault intensity (See Figure A.3). The T^2 chart does not detect the fault as var 4, T_j, is not influential in the PCA model.

Contribution plot of the Q statistic shows that there was a deviation in var 4, while the T^2 contributions are inconclusive. With this information, one can only deduce that a change in T_j has occured, but not much else (See 5.7b).



(a) (b)

Figure 5.7: Jacket temperature sensor fault - intensity 3 - a) MSPC control charts and b) contribution plots.



(a) (b)

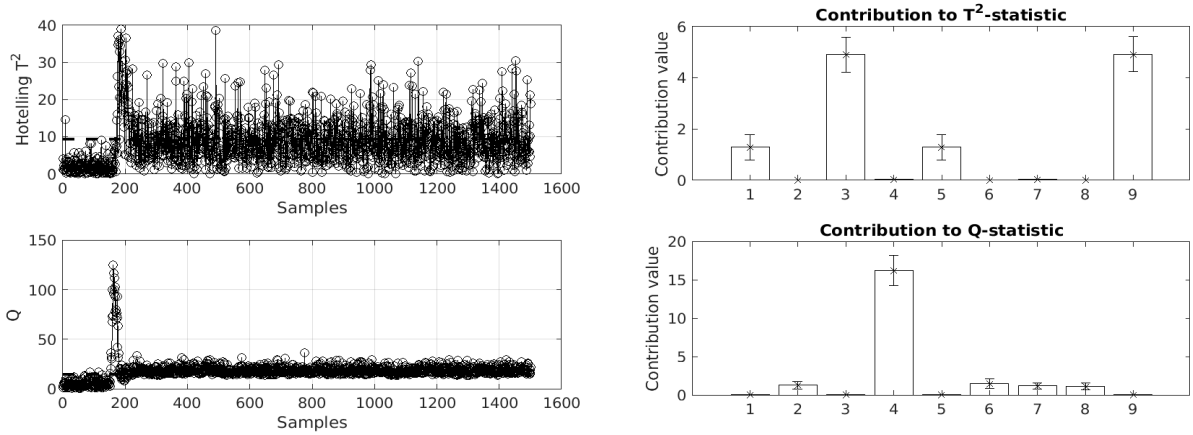
Figure 5.8: Pipe rupture fault - intensity 3 - a) MSPC control charts and b) contribution plots.

The pipe rupture fault is well detected in both control charts, as shown in Figure 5.8a. The effect of fault intensity is present in the detection charts (See Figure A.4).

On both contributions plots of Figure 5.8b the dominance from the contributions of vars 1 and 5 is evident. This indicates a problem with the level and/or the feed flow, but contrary to the level sensor fault that indication is present on both statistics.

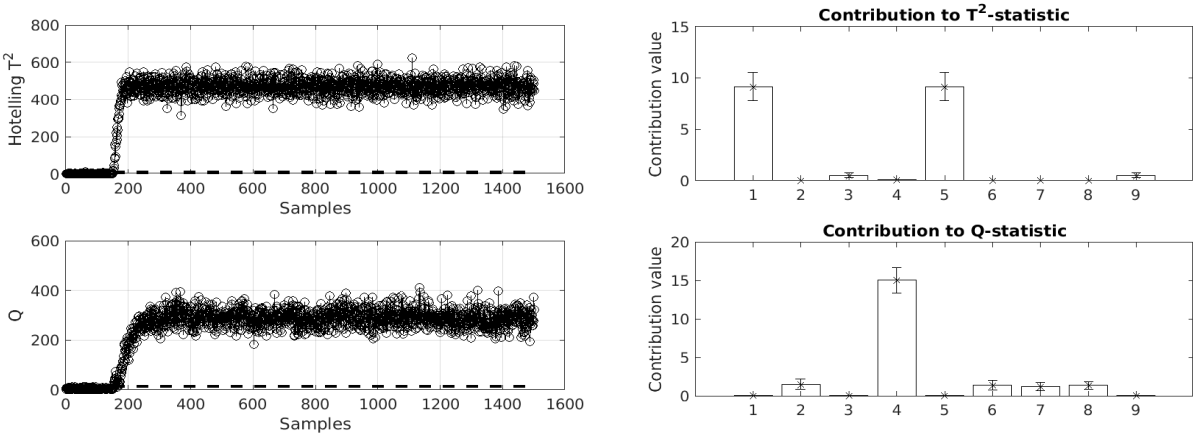
In Figure 5.9a, there is a spike in both control charts. This spike matches the period of adjustment of the system to the new conditions. The detection of this fault is competent. The first two intensities are not detected at all (See Figure A.5).

The contribution plots in Figure 5.9b are similar to the ones for the mixture temperature sensor fault. They inform that something is happening heat transfer-wise, but further conclusions need a deeper analysis.



(a) **(b)**

Figure 5.9: Jacket hole fault - intensity 3 - a) MSPC control charts and b) contribution plots.



(a) **(b)**

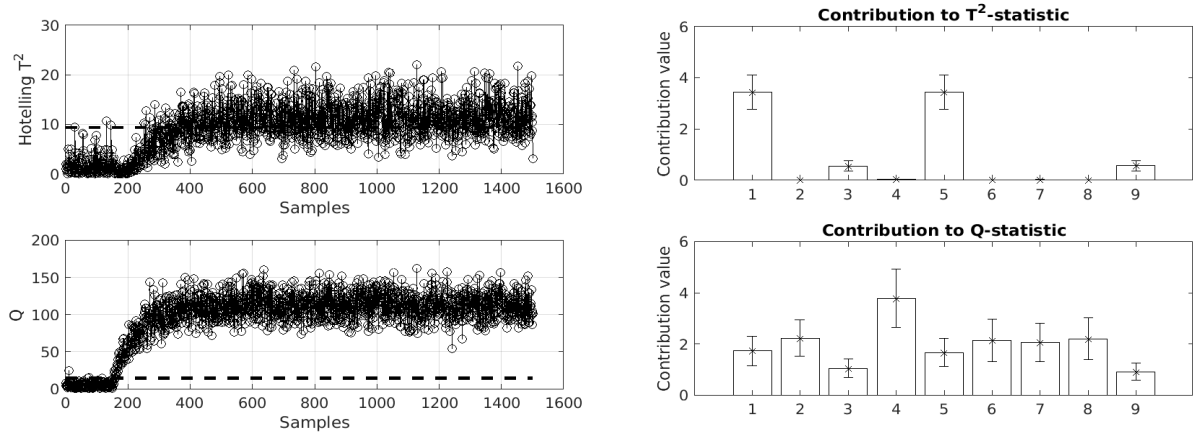
Figure 5.10: Tank and jacket holes fault - intensity 3 - a) MSPC control charts and b) contribution plots.

To detect the holes in the tank and the jacket, Figure 5.10a shows that the MSPC charts give a loud alarm. There is a consistency with different fault intensities (See Figure A.6). The statistics value soar at stay high throughout the operation time.

However, the contributions plots are inconclusive (See Figure 5.10b). The T^2 contribution plots points to changes in the level and the feed flow rate and the Q contribution plots highlight the jacket temperature. One gets the idea that something is up with the tank level and heat transfer, but no further illations can be drawn.

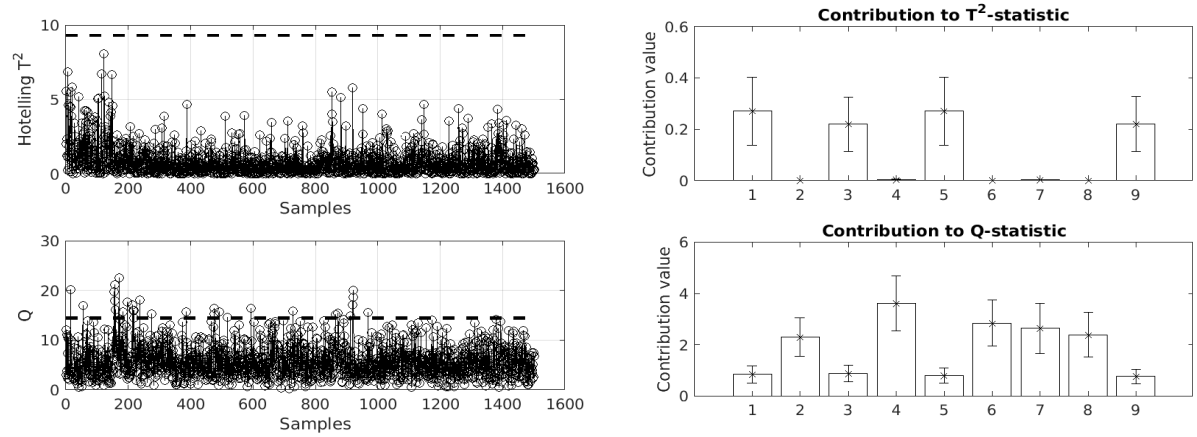
In valve lock fault, intensity 3, both control valves lock. The PCA-MSPC methodologies work as it is shown in Figures 5.11a and A.7. The detection is better in the Q control chart, but is also clear in the T^2 chart.

Figure 5.11b shows a misleading T^2 contribution plot as it only points to changes in the level and feed flow rate when in reality both control valves have stopped working. In the Q contribution plot all variables have almost equal modest contributions which means this plot is inconclusive.



(a) **(b)**

Figure 5.11: Valve lock fault - intensity 3 - a) MSPC control charts and b) contribution plots.



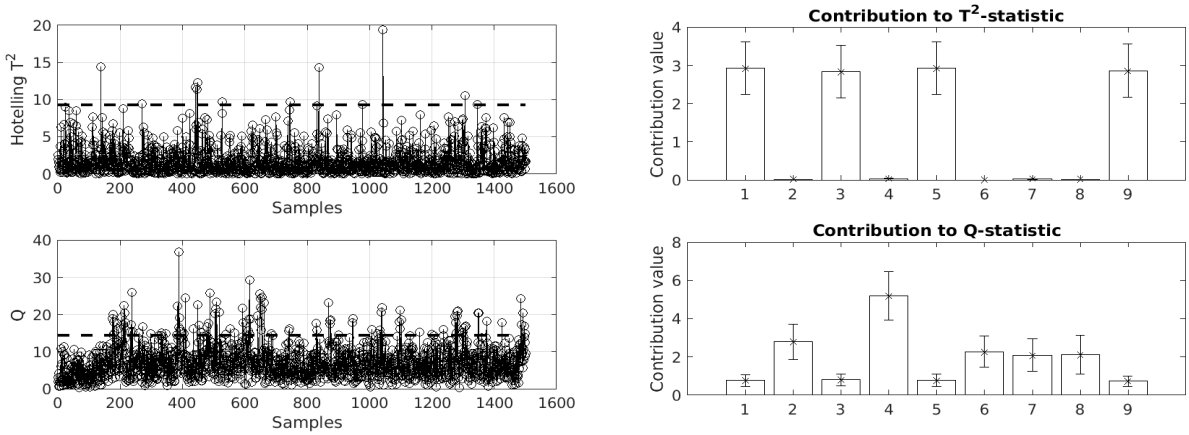
(a) **(b)**

Figure 5.12: Valve offset fault - intensity 3 - a) MSPC control charts and b) contribution plots.

Valve offset has a different behavior from the previous discussed faults in the T^2 charts. This is due to the lack of other special events in this simulations the valves try to comply with the controllers orders much slower in a otherwise NOC process, the controlled variables tend to stabilize around their average. In the Q chart the some change is sightly noted. Contribution plots are totally inconclusive in this fault's analysis (See Figure 5.12).

MSPC control charts are able to detect some change due to the stiction phenomenon, especially in the Q chart (See Figure 5.13a). The intensity increase is translated in better detection by the control charts (See Figure A.9).

Contribution plots in Figure 5.13b show small equal contributions for the T^2 statistic as well as for the Q statistic where var 4 barely sticks out. It is possible to stat that no conclusion that points to stiction can be extracted from these contribution plots.



(a) (b)

Figure 5.13: Valve stiction fault - intensity 3 - a) MSPC control charts and b) contribution plots.

The MSPC-PCA methodology has proven to be a reliable way of fault detection with poor results for valve lock and sensor level, but overall good performance. Although, T^2 and Q control charts are complementary and must be used together, it is often the case that the Q control chart is more sensitive to changes in the process and therefore is to be monitored carefully and calibrated with care, so it does not incur in too many false alarms.

Contribution plots have mixed results. They perform well for sensors, as stated by (Kerkhof et al. 2013), and give fairly good indicators for process faults. However, contribution plots perform poorly for instrumentation.

Detection evaluation was computed to have a global picture of the performance of the MSPC control charts. All of the computed FARs are around 0.01, as defined in the UCLs computation. Furthermore, True Detection Rates (TDRs), Average Run Lengths (ARLs) and Average Time to Signals (ATSs) are computed, but their detailed analysis would not bring new information to this thesis. Nevertheless, the results can be consulted and attested in appendix B.

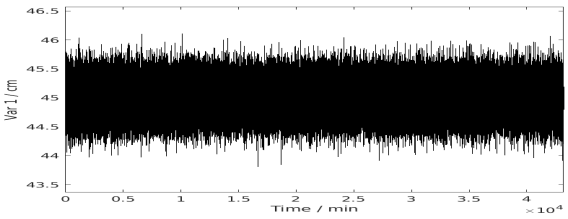
5.2 Monitoring of process & equipment degradation

To monitor process and equipment degradation, thirty days of operation of the system are simulated. This time, catalyst activity loss, external temperature variations and fouling are added to the simulation routine. A data matrix $\boldsymbol{X}_{n\times m}$ was generated containing 216001 observations and 9 measured variables. The state and load measured variables are the same as in section 5.1 and the simulation results can be seen in Figure 5.14.

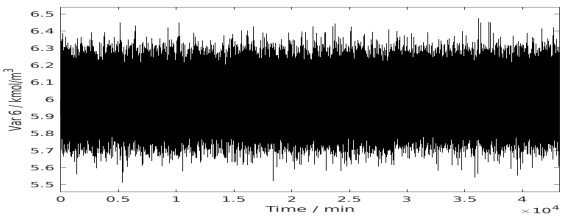
The original 30-day operation dataset is divided in three different sets: the training set with the first 5 days of operation, the validation set with the next 2 days of operation and the test set with the remaining 23 days of operation.

Taking the training and applying PCA, one is presented with the following scree plot:

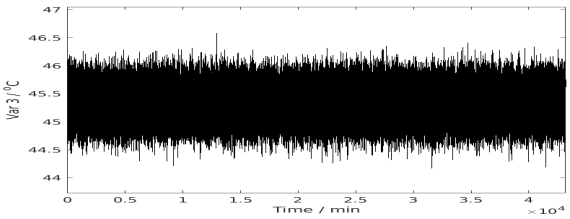
Observing Figure 5.15, the first three eigenvalues are the ones who explain more variability



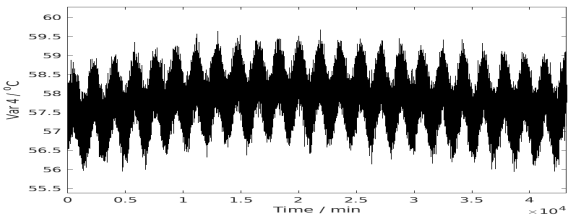
(a) Level measurements.



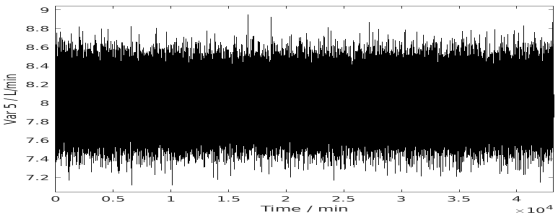
(b) Concentration of A measurements.



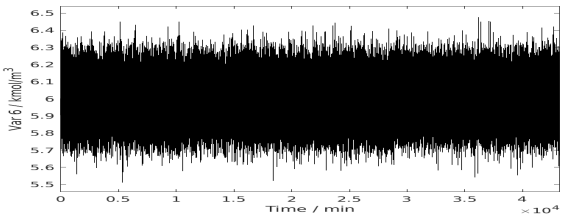
(c) Mixture temperature measurements.



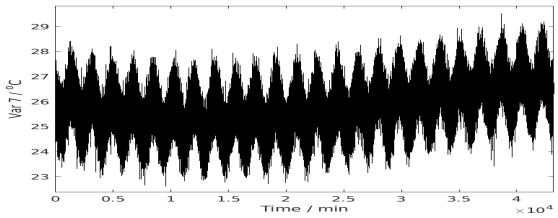
(d) Jacket temperature measurements.



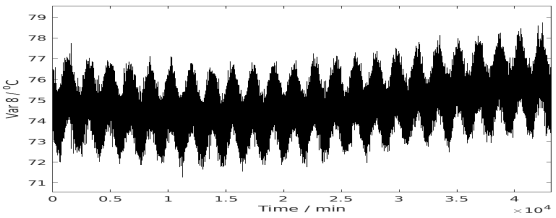
(e) Feed flow rate measurements.



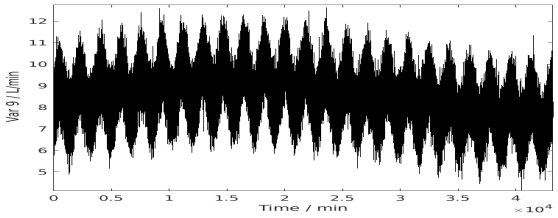
(f) Feed concentration of A measurements.



(g) Feed temperature measurements.



(h) Feed jacket temperature measurements.



(i) Thermal fluid flow rate measurements.

Figure 5.14: Measurement signals of variables over a month with process & equipment degradation.

of the data. For the fourth on, eigenvalues start to acquire a linear behavior in groups of two, probably representing different noise degrees. Therefore, three PCs are chosen to the PCA model which together represent 65.57% of data variability.

In Figure 5.16, the loadings of the first three PCs are presented. The first PC is dominated by temperature related variables, vars 3, 4, 7, 8 and 9. PC2 explains the variability of var 1, level, and var 5, feed flow rate. PC3 also explains variability in the same temperature related variables like PC1, but with a dominance of var 3, mixture temperature. The analysis of loadings indicates that the variables influencing data variability are temperature

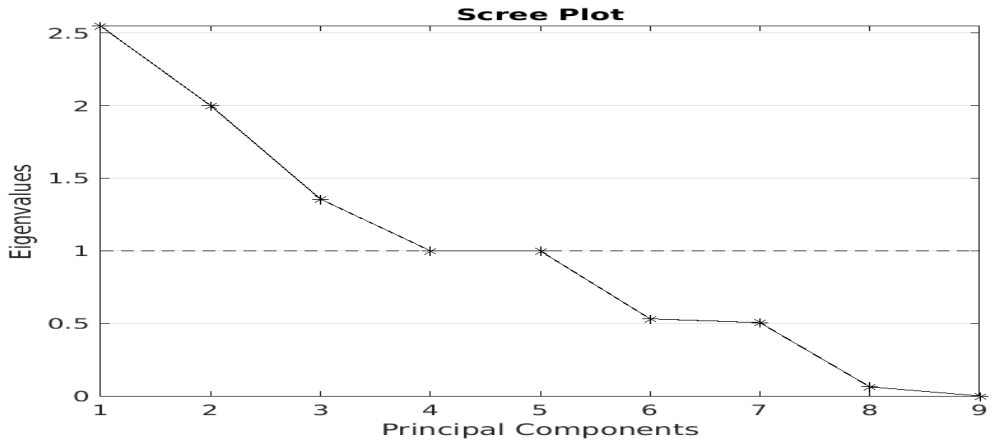


Figure 5.15: Scree plot with the eigenvalues from data matrix $\boldsymbol{X}$ during one month.

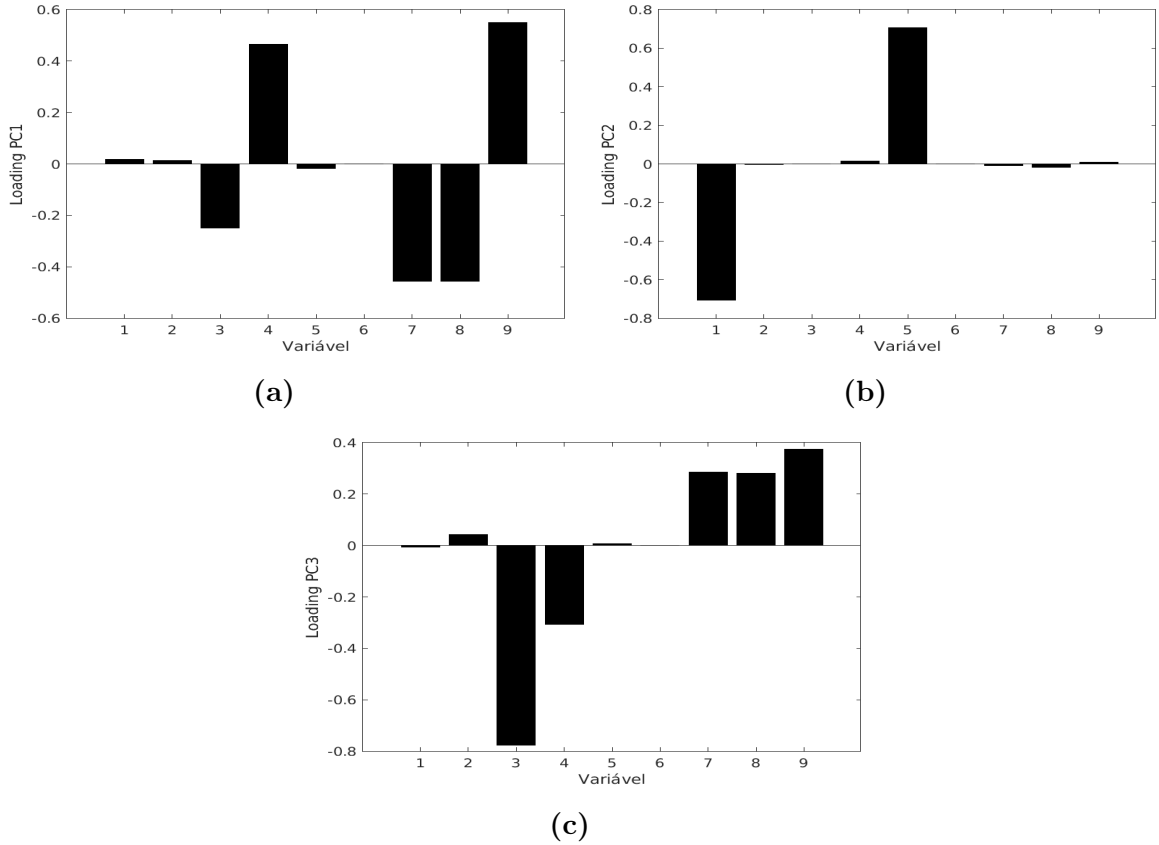


Figure 5.16: Loadings of a) PC1, b) PC2 and c) PC3.

related which points to possible issues in heat transfer or a strong influence of external temperature variation.

Considering the validation set and an $\alpha = 0.01$, the UCLs of the MSPC control charts are computed giving T^2 UCL equal to 9.241 and Q-statistic UCL equal to 14.4375.

Now, considering the test set, one begins a data exploratory analysis in search of data dynamic trends. Plotting the scores against time shows that there are slower trends present in the data.

In figure 5.17a, scores behave in a sinusoidal pattern. Given that PC1 is dominated

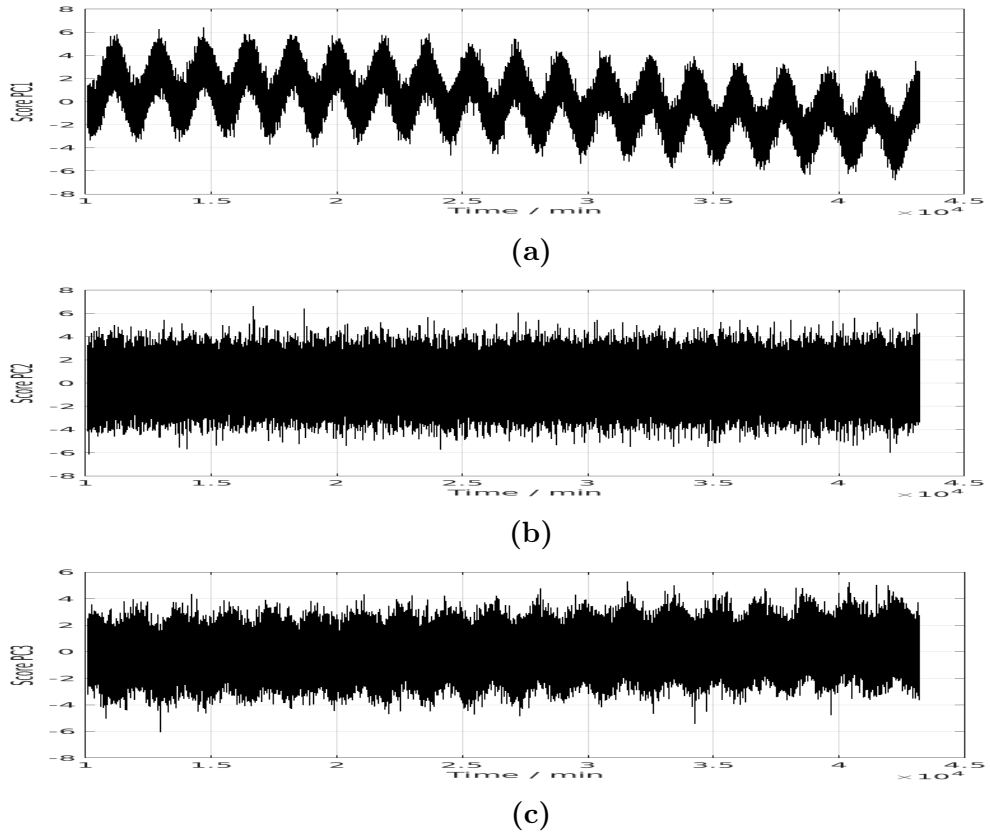


Figure 5.17: Dynamics of the scores of a) PC1, b) PC2 and c) PC3.

by temperature related variables, it appears that external temperature variations are responsible by the data variability. Figure 5.17b shows the scores of PC2 apparently with no visible longer term dynamics. In Figure 5.17c, scores show a light increase over time.

5.2.1 Window PCA

To get more insight in the changes suffered by the system over time, window PCA is applied to the 30-day dataset. This dataset is divided in 5-day windows and PCA is performed in each window. A scree test is made giving each window a 2 PCs PCA model. Then the minimum angle between the first window's PCA subspace and each of the next window's PCA subspaces is computed.

The computed angles are very small (less than 1^o), as seen in Figure 5.18, which means the different window PCA models are similar to each other. Therefore, to make sure that it was the method that was failing and not simulation parameters, another simulation was performed also representing 30 days of operation, but with higher m_d and k_d equal to 3.33×10^{-10} m^2 °C (J/min)$^{-1}$ min^{-1} and 5.77×10^{-6} min^{-1}, respectively. The computed angles, nevertheless, remained almost unchanged with the exception of the angle between the first and the second windows. This leads to the belief that window PCA fails to detect process and equipment degradation.

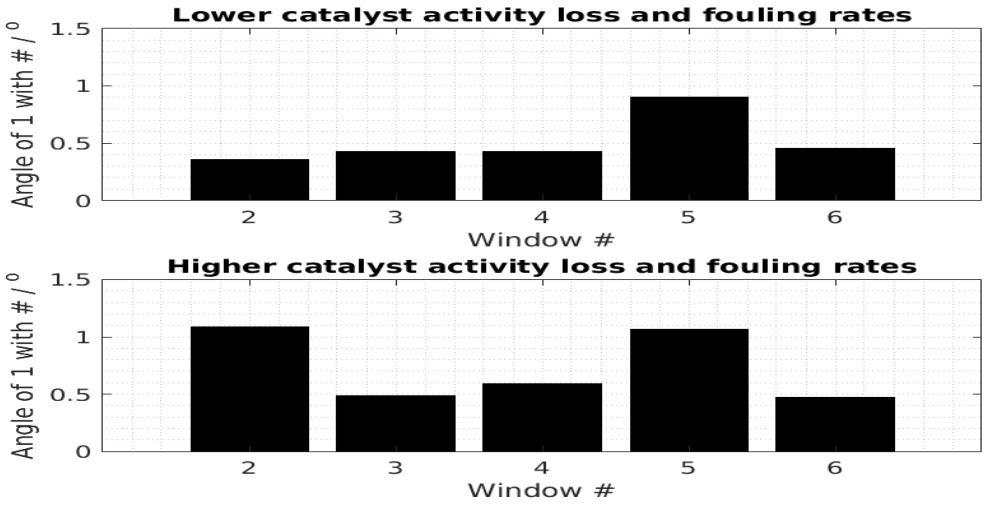


Figure 5.18: Angles formed between the subspace of the first window and all next five.

5.2.2 MRA

Considering the scores shown in Figure 5.17 for the methodology presented in Figure 3.4, wavelet transforms are applied to each PC scores (See appendix C). Various maximum decomposition levels, J_{max}, filters and reconstruction frameworks are tested until decide on the $J_{max} = 14$, the symmlet filter (Buckheit et al. 2005) and a three way reconstruction where the first way, from d_1 to d_5, has mostly noise, the second way, from d_6 to d_{13}, has daily dynamics and the third way, with d_{14} and a_{14}, that has long term dynamics (See Figures 5.19. The same reconstruction for the other scores can be consulted in appendix C, Figures C.4 and C.5).

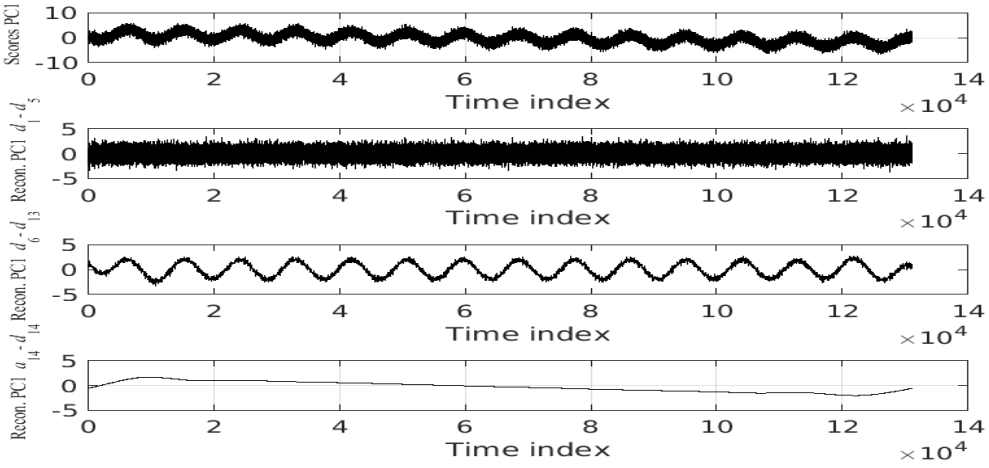


Figure 5.19: From top to bottom: Original PC1 scores signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

In Figure 5.19, the bottom decomposition shows a linear dynamic trend descending throughout the time index. The beginning and the end of the plot tend to curve, but that

is to the lack of data on the left and on the right, respectively. It is a wavelet transform numerical issue and therefore should be discarded from this analysis.

With the knowledge and characterization gathered from the scores MRA, the method presented in Figure 3.4 proceeds by applying wavelet transforms to the raw data (See appendix C). The following Figures show the reconstruction of the signals of variables that presented long term dynamics. From this series of plots are excluded the following variables: tank level, feed flow rate, feed reactant concentration and the mixture temperature. The first three are not influenced by the slower dynamics simulated and effect in the mixture temperature is masked by the temperature controller. Nevertheless, the reconstruction plots of this variables can be consulted in appendix C.

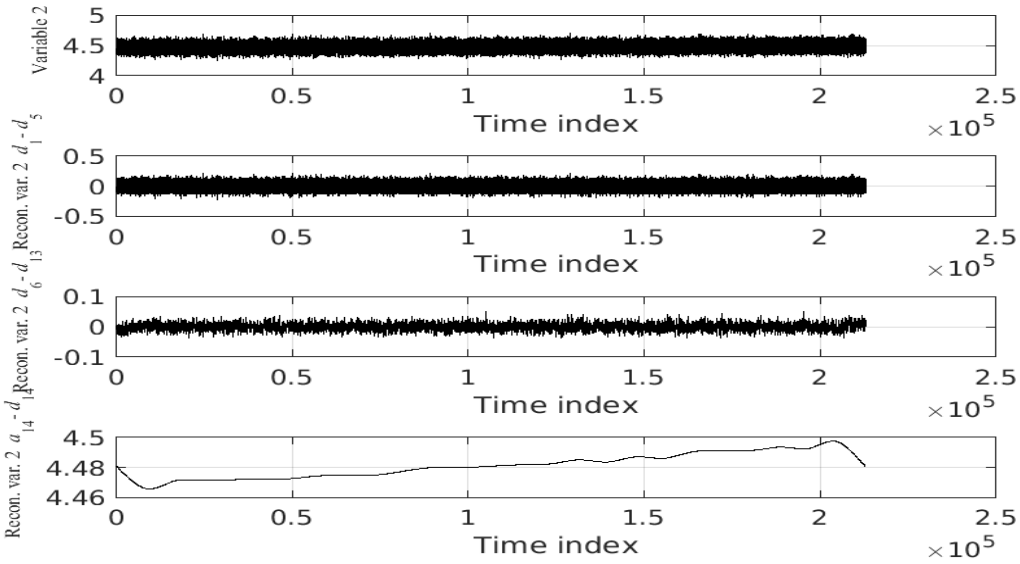


Figure 5.20: From top to bottom: Original concentration of A variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

Figure 5.20 shows that for the first two reconstructions the signal is constituted of stochastic noise. The last reconstruction, however, shows that the concentration of A is rising. This rise is due to catalyst activity loss, that reduces the reaction rate, and fouling that reduces heat transfer to the mixture which also reduces reaction rate due to the feed stream's low temperature, therefore increasing the reactant's concentration in the mixture.

The variables feed temperature and feed jacket temperature show identical daily sinusoidal dynamics in the second reconstruction and a growing curve in the last reconstruction (See Figures 5.21 and 5.22). These dynamics are caused by external temperature variations and the method succeeded in detecting them.

The remaining variables, jacket temperature and thermal fluid flow rate, also show a daily sinusoidal dynamics in the second reconstruction caused by external temperature variations. In the last reconstruction, the long term dynamics forms a concave curve (See Figures 5.23 and 5.24). In the case of the thermal fluid flow rate, these dynamics result of the controller's actions to counter external temperature variations and fouling. Regarding

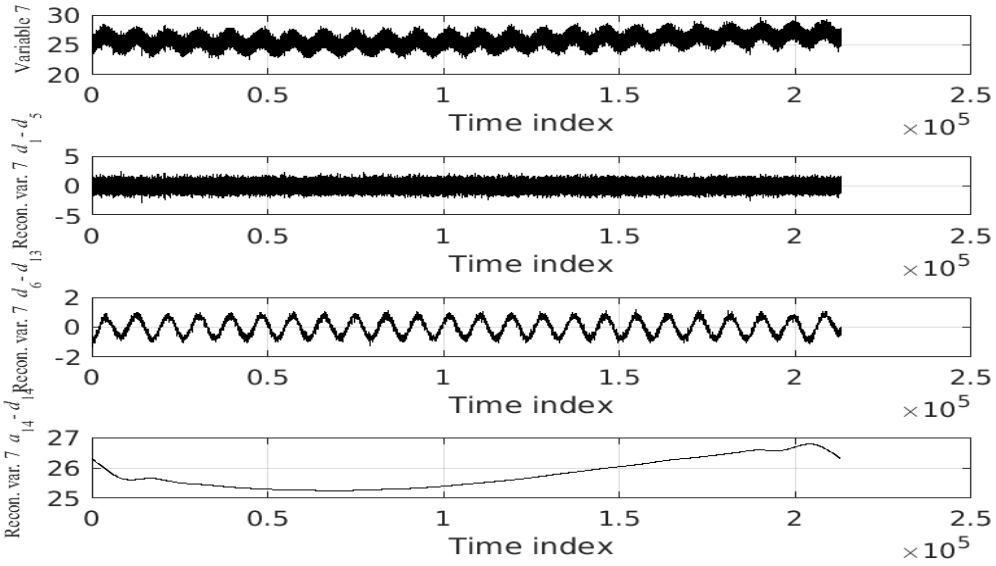


Figure 5.21: From top to bottom: Original feed temperature variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

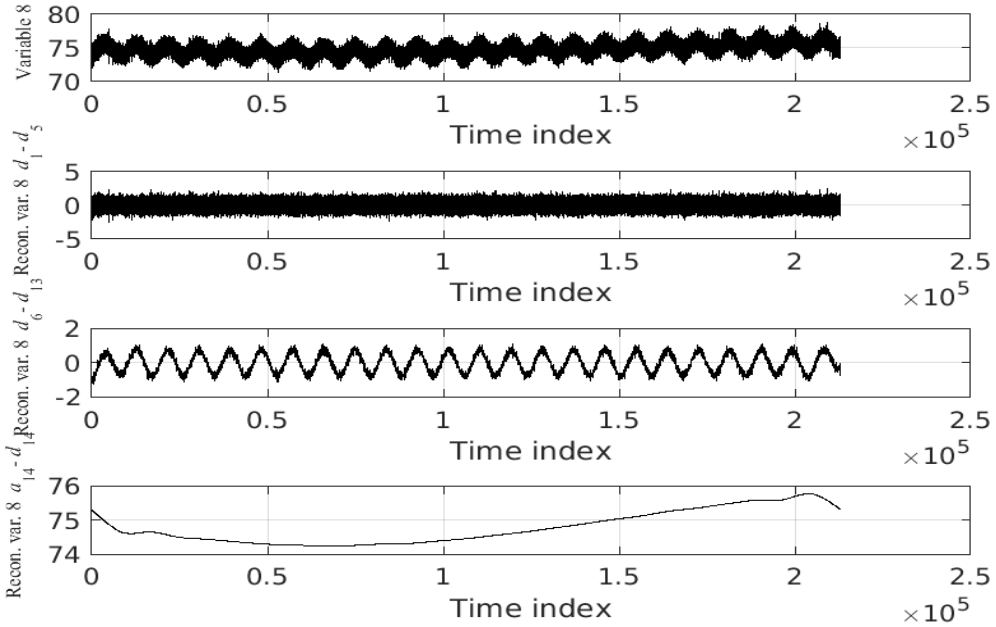


Figure 5.22: From top to bottom: Original feed jacket temperature variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

jacket temperature, fouling prevents heat transfer from the jacket to the liquid which in turn does not cool the thermal fluid.

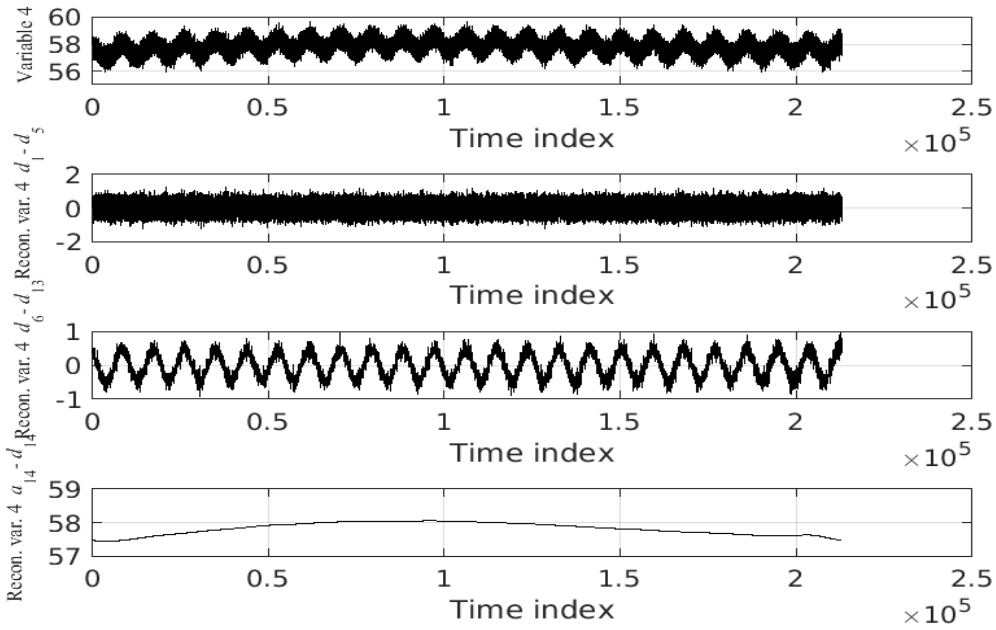


Figure 5.23: From top to bottom: Original jacket temperature variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

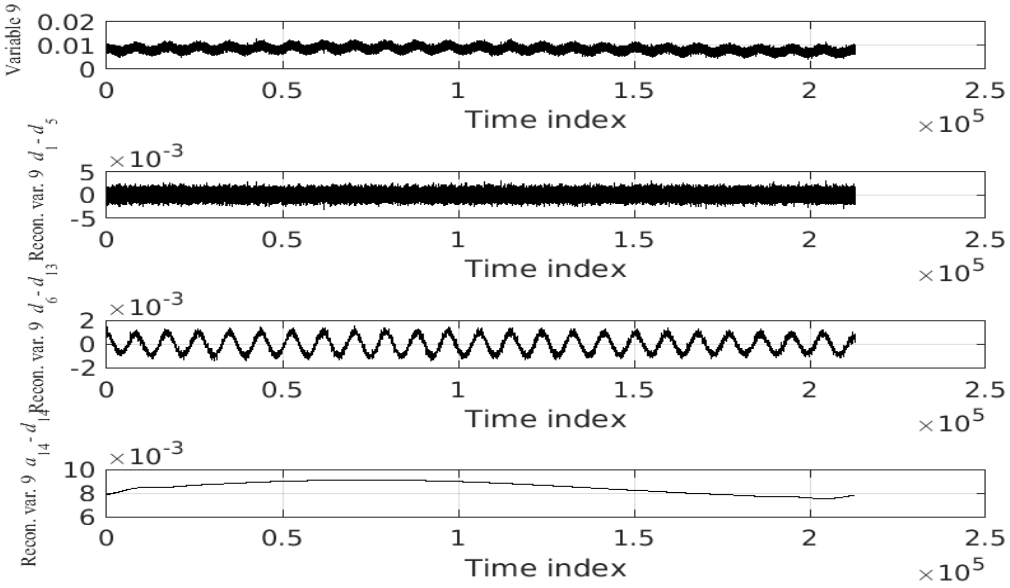


Figure 5.24: From top to bottom: Original thermal fluid flow rate variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

After reconstructing the three data matrices $\hat{\boldsymbol{X}}_{hf}$ (high frequency), $\hat{\boldsymbol{X}}_{mf}$ (medium fre-

quency) and $\hat{\boldsymbol{X}}_{lf}$ (low frequency), respectively reconstruction of the signal with details coefficients from d_1 to d_5, reconstruction of the signal with details coefficients from d_6 to d_{13} and reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}, one performs PCA to each data matrix.

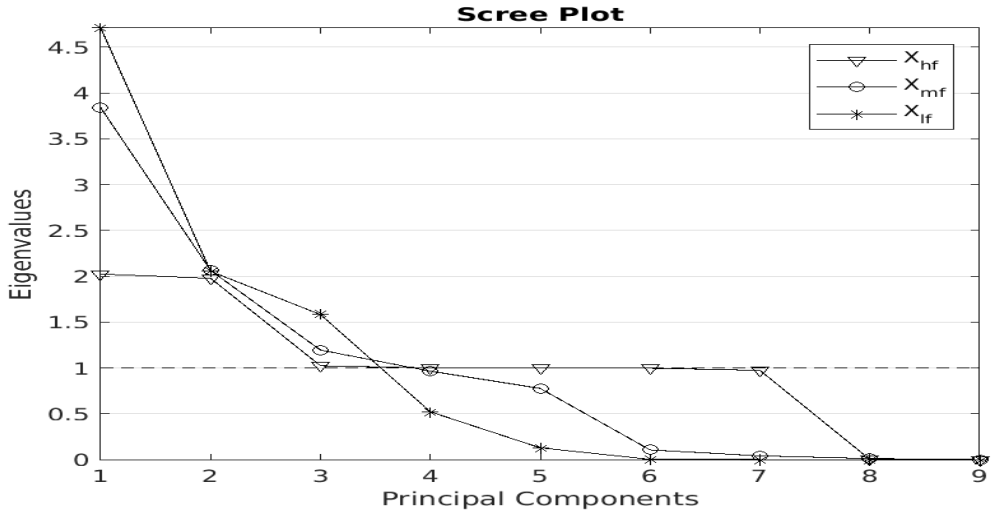
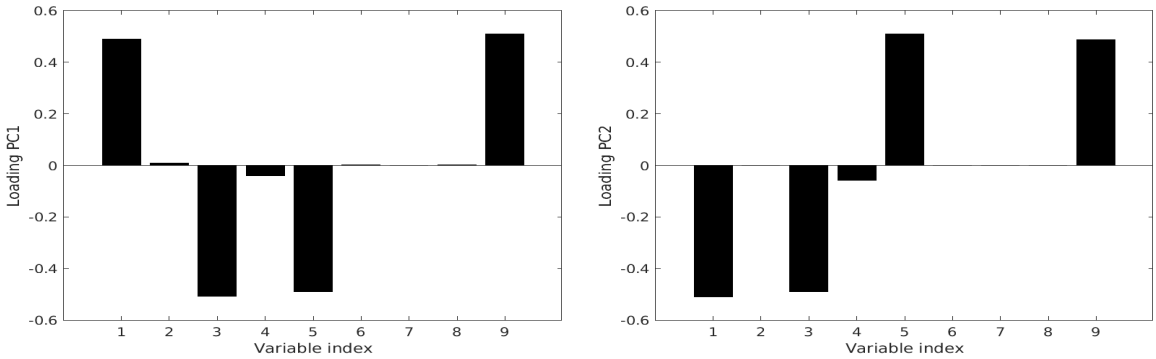


Figure 5.25: Scree plot with the eigenvalues from the reconstructed $\hat{\boldsymbol{X}}_{hf}$, $\hat{\boldsymbol{X}}_{mf}$ and $\hat{\boldsymbol{X}}_{lf}$ matrices.

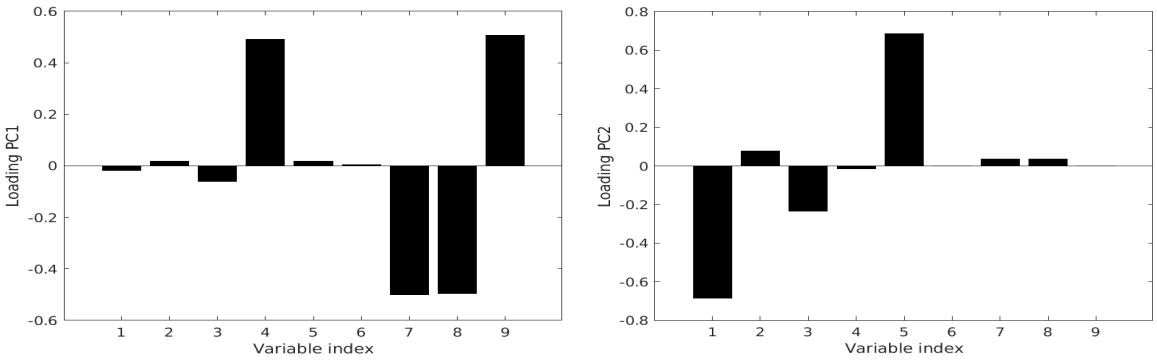


(a) **(b)**

Figure 5.26: From matrix $\hat{\boldsymbol{X}}_{hf}$, loadings of a) PC1 and PC2.

The three scree plots, in Figure 5.25, must be studied to decide which PC loadings are to be considered in the final analysis. Considering $\hat{\boldsymbol{X}}_{hf}$, two PCs are chosen because the first two eigenvalues have similar value and together represent 44.47% of data variability. For $\hat{\boldsymbol{X}}_{mf}$, from the third eigenvalue onward there is a linear trend which usually represent data noise. Therefore, two PCs are chosen for model which represent 65.60% of data variability. Regarding $\hat{\boldsymbol{X}}_{lf}$, one PC is selected for the PCA model because the second eigenvalue is less than half than the first while it is close to the third eigenvalue. Therefore, to keep to PCA model as simple as possible one PC representing 52.33% of data variability is selected.

Figure 5.26 the variables with more weight in both PCs are vars 1, 3, 5 and 9 or tank level, mixture temperature, feed flow rate and thermal fluid flow rate. This implies that the high frequency dynamics are dominated these four variables present in the controllers



(a) (b)

Figure 5.27: From matrix $\hat{\boldsymbol{X}}_{mf}$, loadings of a) PC1 and PC2.

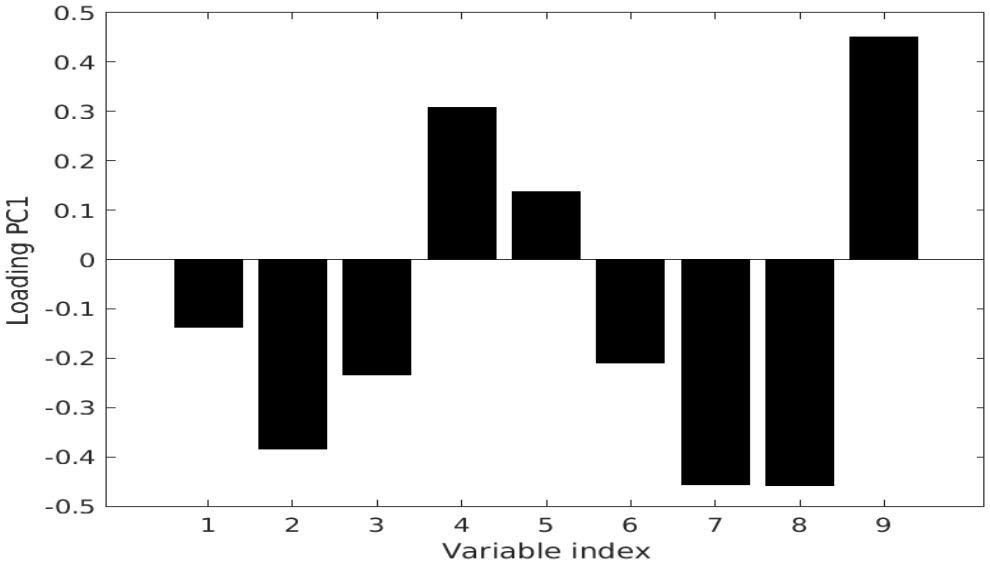


Figure 5.28: Loadings of PC1 from matrix $\hat{\boldsymbol{X}}_{lf}$.

loops. Therefore, the actions of control that keep the process stable and the controlled variables on target are responsible for high frequency data variability.

Figure 5.27 shows the loadings from daily dynamics. In first PC, the more important variables are vars 4, 7, 8 and 9. These variables were identified as having a sinusoidal daily dynamics. In the second PC, level and feed flow rate are the variables with more weight. Therefore, medium frequency dynamics are explained by the effect of external temperature variations and the level control loop.

Figure 5.28 shows the loadings relative to the slower dynamics. The variables with more weight are temperature related - vars 4 , 7, 8 and 9 - and the reactant concentration - var 2. The last one suffers the action of catalyst activity loss and fouling. The reactor needs a inlet heat stream to counter the effect of the colder feed stream which causes the reaction rate to decrease. As for the effect on temperature, the responsibility falls on the monthly external temperature variation, regarding the feed temperature and feed thermal fluid temperature, and heat transfer issues caused by fouling, regarding jacket temperature and thermal fluid flow rate.

Chapter 6

Conclusions and future works

«Finally, in conclusion,
let me say just this.»
*(*Peter Sellers*)*

Remembering that this thesis goals consists of proposing and testing procedures to monitor and diagnose abrupt process failures and to develop methodologies to detect long term equipment degradation, two sets of general problems have been addressed:

1. Detecting and diagnosis of abrupt faults;
2. Detecting process & equipment degradation.

The first topic was tackled with MSPC-PCA methodologies and contribution plots. MSPC-PCA methodologies are well established, easy to implement and have a good detection performance, highlighting the Q-statistic control chart as a sensitive chart that should be monitored carefully and calibrated with care. The introduction of a validation set to compute UCLs contributed to the method's robustness.

Contribution plots did not excel in fault diagnosis. They lead to fairly good results to sensor faults, but it is always needed some interpretation of the plot and knowledge of the process to get to the root cause. It is a helpful fault diagnosis tool that points to where a problem might be, but can not be fully trusted.

The second item was approached with window PCA and multivariate & multiresolution analysis. The window PCA was not able to detect the slower dynamics simulated. On the other hand, MRA was able to detect all three process & equipment degradation that were simulated. Moreover, the subsequent PCA analysis points out the causes of the slower dynamics, functioning as a pre-diagnose methodology.

Overall, the thesis goals were met.

In terms of future works, it is proposed to repeat the multivariate & multiresolution analysis under sets of different process conditions, changing the external temperature variations, catalyst deactivation rate and foulant deposit mass rate. With the resulting data the objective is to create a regression model able to predict the state of degradation of a given process based on the data measured in a chemical process. A model like this would be a handful tool to plan maintenance interventions.

Chapter 7

Bibliography

Aspuru-Guzik, A., Lindh, R. & Reiher, M. (2018), 'The matter simulation (r)evolution.', *ACS Central Science* **4**(2), 144–152.

Åström, K. J. & Hägglund, T. (2001), 'The future of pid control', *Control Engineering Practice* **9**(11), 1163–1175.

Bott, T. R. (1995), *Fouling of Heat Exchangers*, first edn, Elsevier Science & Technology Books.

Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. (1994), *Times series analysis*, third edn, Prentice-Hall, Inc, Upper Saddle River, New Jersey.

Bro, R. & Smilde, A. K. (2014), 'Principal component analysis', *Anal. Methods* **6**(9), 2812–2831.

Buckheit, J., S., C., Donoho, D., Johnstone, I. & Scargle, J. (2005), 'Wavelab version 0.850', http://statweb.stanford.edu/~wavelab/.

Choudhury, A. A. S., Shah, S. L. & Thornhill, N. F. (2008), *Diagnosis of Process Nonlinearities and Valve Stiction: Data Driven Approaches*, first edn, Springer-Verlag, Berlin.

Choudhury, A. A. S., Thornhill, N. F. & Shah, S. L. (2005), 'Modelling valve stiction', *Control Engineering Practice* **13**(5), 641–658.

Desborough, L. & Miller, R. (2002), 'Increasing customer value of industrial control performance monitoring — honeywell's experience', *AIChE Symposium Series 2001* **326**, 172–192.

Jackson, J. E. (1991), *A User's Guide To Principal Components*, first edn, John Wiley & Sons, Inc, New York, USA.

Johnson, R. A. & Wichern, D. W. (2007), *Applied Multivariate Statistical Analysis*, sixth edn, Pearson Prentice Hall, Upper Sadle River, New Jersey.

Jollife, I. T. (2002), *Principal Component Analysis*, second edn, Springer, New York, USA.

Kerkhof, P. V. D., Vanlaer, J., Gins, G. & Impe, J. F. V. (2013), 'Analysis of smearing-out in contribution plot based fault isolation for Statistical Process Control', *Chemical Engineering Science* **104**, 285–293.

Kourti, T. (2005), 'Application of latent variable methods to process control and multivariate statistical process control in industry', *International Journal of Adaptive Control and Signal Processing* **19**(4), 213–246.

Kourti, T. & MacGregor, J. (1995), 'Process analysis, monitoring and diagnosis, using multivariate projection methods', *Chemometrics And Intelligent Laboratory Systems* **28**, 3–21.

Kresta, J. V., Macgregor, J. F. & Marlin, T. E. (1991), 'Multivariate statistical monitoring of process operating performance', *The Canadian Journal of Chemical Engineerging* **69**, 35–47.

Krzanowski, W. J. (1979), 'Between-Groups Comparison of Principal Components', *Journal of the American Statistical Association* **74**(367), 703–707.

Levenspiel, O. (1999), *Chemical Reaction Engineering*, third edn, John Wiley & Sons, Inc.

Luyben, W. L. (1990), *Process Modeling, Simulation and Control for Chemical Engineers*, second edn, McGraw-Hill, Inc, United States of America.

MacGregor, J. F. & Kourti, T. (1995), 'Statistical process control of multivariate processes', *Control Eng. Practise* **3**(3), 403–414.

Melo, L. F., Bott, T. R. & Bernardo, C. A., eds (1988), *Fouling Science and Technology*, Springer Netherlands, NATO ASI Series 145.

Montgomery, D. C. (2013), *Introduction to Statistical Quality Control*, seventh edn, John Wiley & Sons, Inc., United States of America.

Montgomery, D. C., Peck, E. A. & Vining, G. G. (2012), *Introduction to Linear Regression Analysis*, fifth edn, John Wiley & Sons, Inc., Hoboken, New Jersey.

Muller-Steinhagen, H. (2000), *Heat Exchanger Fouling: Mitigation and Cleaning Techniques*, first edn, PUBLICO Publications, Essen, Germany.

OpenWeatherMap (2018), 'Current weather and forecasts in your city', https://openweathermap.org/city/2740637.

Palmer, J. C. & Debenedetti, P. G. (2015), 'Recent advances in molecular simulation: A chemical engineering perspective.', *AIChE Journal* **61**, 370–383.

Peyghambarzadeh, S. M., Vatani, A. & Jamialahmadi, M. (2012), 'Application of asymptotic model for the prediction of fouling rate of calcium sulfate under subcooled flow boiling', *Applied Thermal Engineering* **39**, 105–113.

Rahmani, S., McCaffrey, W. C., Dettman, H. D. & Gray, M. R. (2003), 'Coking kinetics of asphaltenes as a function of chemical structure', *Energy and Fuels* **17**(4), 1048–1056.

Rase, H. F. (1990), *Fixed-Bed Reactor Design and Diagnostics: Gas-Phase Reactions*, Butterworth Publishers, Stoneham, MA.

Ratcliff, L. E., Mohr, S., Huhs, G., Deutsch, T., Masella, M. & Genovese, L. (2017), 'Challenges in large scale quantum mechanical calculations.', *Wiley Interdisciplinary Reviews: Computational Molecular Science* **7**(1), 1–24.

Rato, T. J. & Reis, M. S. (2013*a*), 'Advantage of using decorrelated residuals in dynamic principal component analysis for monitoring large-scale systems', *Industrial and Engineering Chemistry Research* **52**(38), 13685–13698.

Rato, T. J. & Reis, M. S. (2013*b*), 'Fault detection in the Tennessee Eastman benchmark process using dynamic principal components analysis based on decorrelated residuals (DPCA-DR)', *Chemometrics and Intelligent Laboratory Systems* **125**, 101–108.

Rato, T. J., Rendall, R., Gomes, V., Chin, S. T., Chiang, L. H., Saraiva, P. M. & Reis, M. S. (2016), 'A Systematic Methodology for Comparing Batch Process Monitoring Methods: Part I-Assessing Detection Strength', *Industrial and Engineering Chemistry Research* **55**(18), 5342–5358.

Reis, M. S. (2005), Data-Driven Multiscale Monitoring, Modelling and Improvement of Chemical Processes, PhD thesis, University of Coimbra.

Reis, M. S. & Saraiva, P. M. (2000), Introdução à análise multiresolução e suas aplicações no contexto da engenharia química, Technical report, Departament of Chemical Engineering, Faculty of Sciences and Technology of the University of Coimbra.

Reis, M. S. & Saraiva, P. M. (2006), 'Heteroscedastic latent variable modelling with applications to multivariate statistical process control', *Chemometrics and Intelligent Laboratory Systems* **80**(1), 57–66.

Reis, M. S. & Saraiva, P. M. (2008), *Statistical Practice in Business and Industry*, first edn, John Wiley & Sons, Ltd, chapter 13 *Multivariate and multiscale data analysis*, pp. 337–371.

Ruel, M. (2000), 'Stiction: The hidden menace.', *Control Magazine* .

Schreier, P. J. & Fryer, P. J. (1995), 'Heat exchanger fouling: A model study of the scaleup of laboratory data', *Chemical Engineering Science* **50**(8), 1311–1321.

Seborg, D. E., Edgar, T. F. & Mellichamp, D. A. (2004), *Process Dynamics and Control*, second edn, John Wiley & Sons, Inc.

Shewhart, W. A. (1931), *Economic Control of Quality of Manufactured Product*, Van Nostrand Company, Inc., Princeton, NJ.

Smith, C. A. & Corripio, A. B. (1985), *Principles and Practice of Automatic Process Control*, first edn, John Wiley & Sons, Inc, New York.

Stephanopoulos, G. (1984), *Chemical Process Control: An Introduction to Theory and Practise*, first edn, Prentice Hall, Englewood Cliffs, NJ.

Westerhuis, J. A., Gurden, S. P. & Smilde, A. K. (2000), 'Generalized contribution plots in multivariate statistical process monitoring', *Chemometrics and Intelligent Laboratory Systems* **51**(1), 95–114.

Wise, B. M. & Gallagher, N. B. (1996), 'The process chemometrics approach to process monitoring and fault detection', *Journal of Process Control* **6**(6), 329–348.

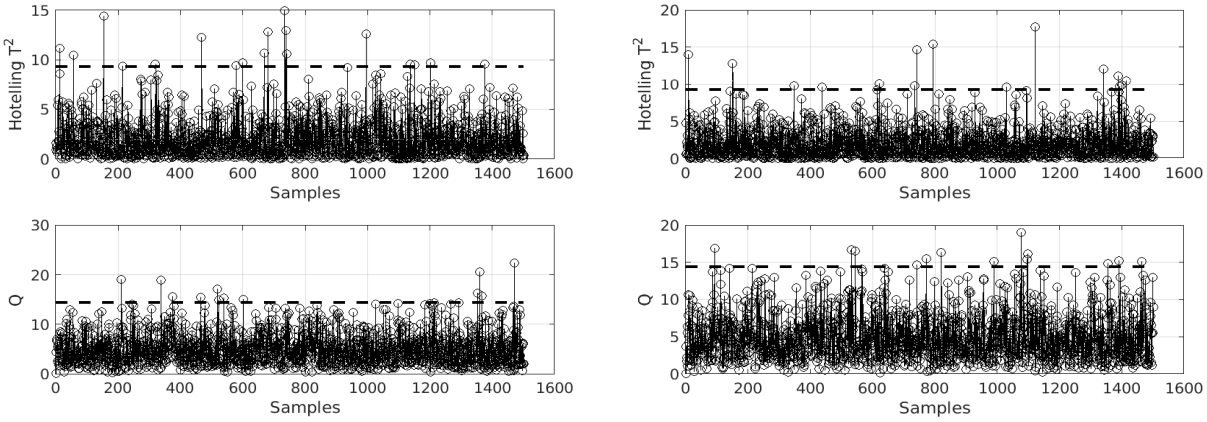
Wold, S., Sjöström, M. & Eriksson, L. (2001), 'PLS-regression: A basic tool of chemometrics', *Chemometrics and Intelligent Laboratory Systems* **58**(2), 109–130.

Xu, L., Liu, J., Wang, Q., Liu, S., Xin, W. & Xu, Y. (2004), 'Coking kinetics on the catalyst during alkylation of fcc off-gas with benzene to ethylbenzene', *Applied Catalysis A: General* **258**(1), 47–53.

Yeap, B. L., Wilson, D. I., Polley, G. T. & Pugh, S. J. (2004), 'Mitigation of crude oil refinery heat exchanger fouling through retrofits based on thermo-hydraulic fouling models', *Chemical Engineering Research and Design* **82**(1), 53–71.

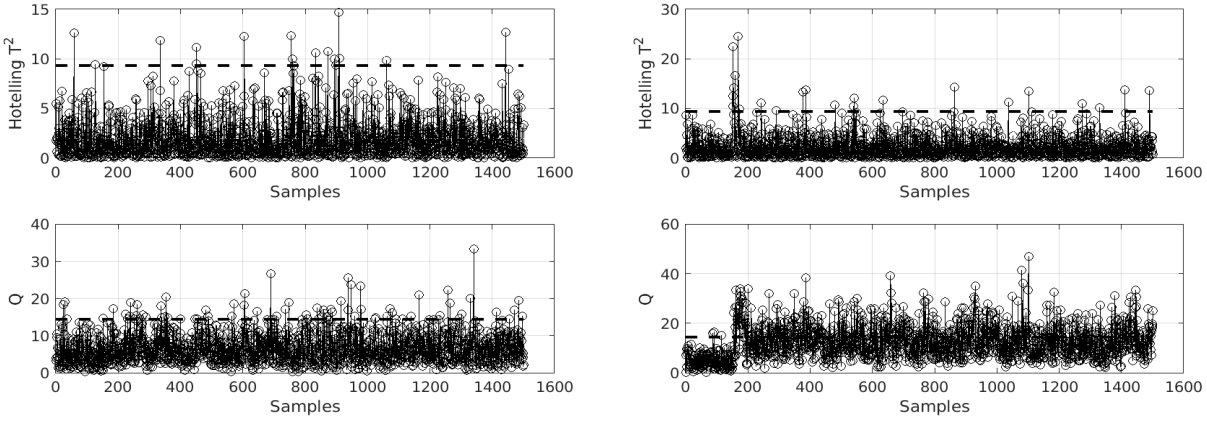
Appendix A

MSPC control charts and contribution plots



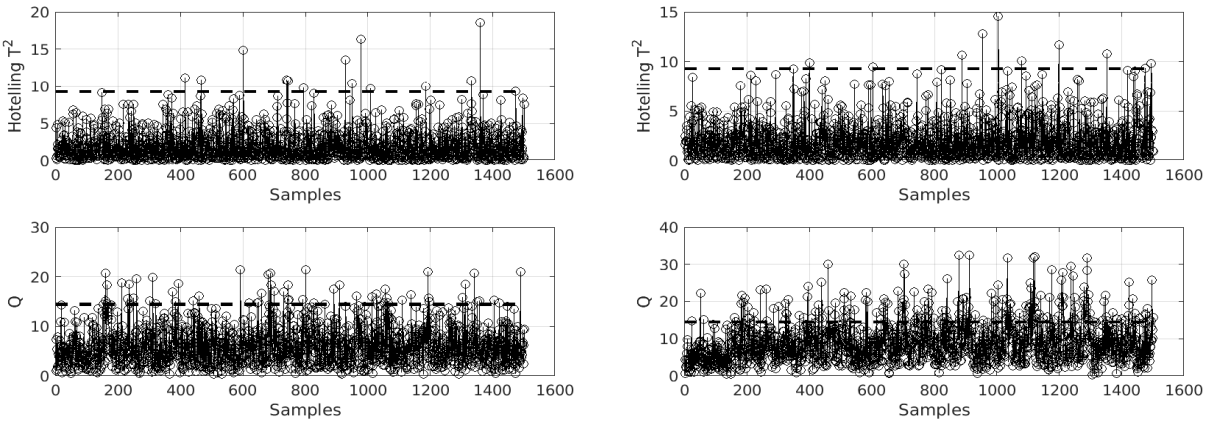
(a) **(b)**

Figure A.1: MSPC control charts level sensor fault a) intensity 1 and b) intensity 2



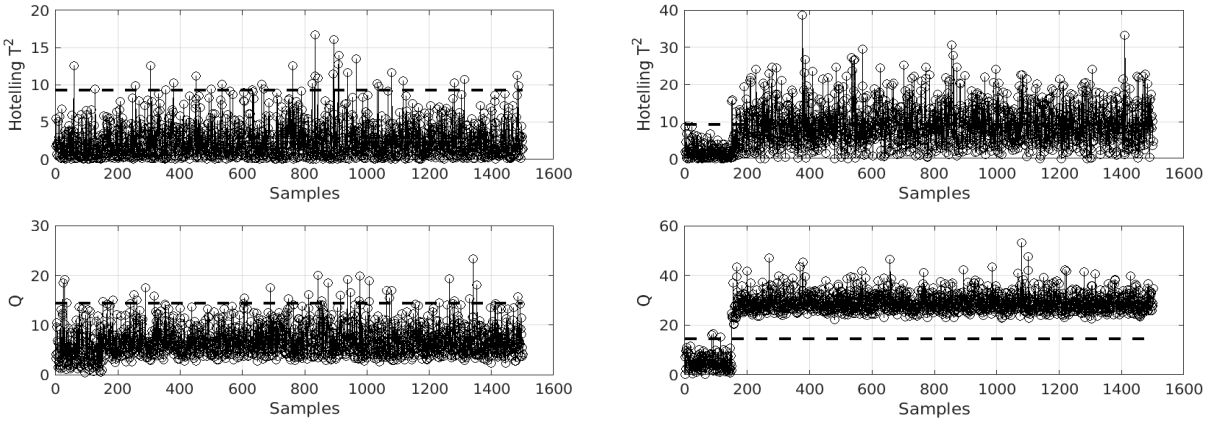
(a) **(b)**

Figure A.2: MSPC control charts mixture temperature sensor fault a) intensity 1 and b) intensity 2



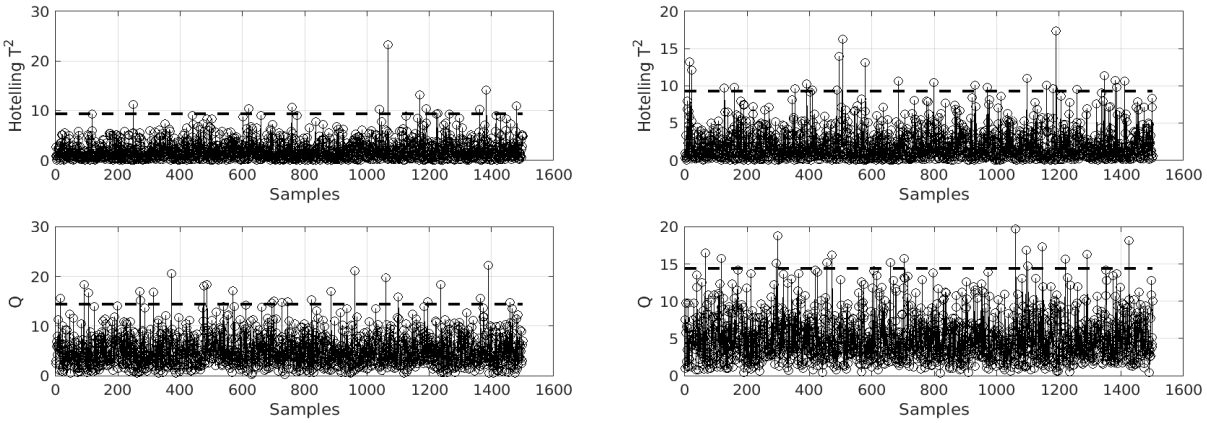
(a) **(b)**

Figure A.3: MSPC control charts jacket temperature sensor fault a) intensity 1 and b) intensity 2



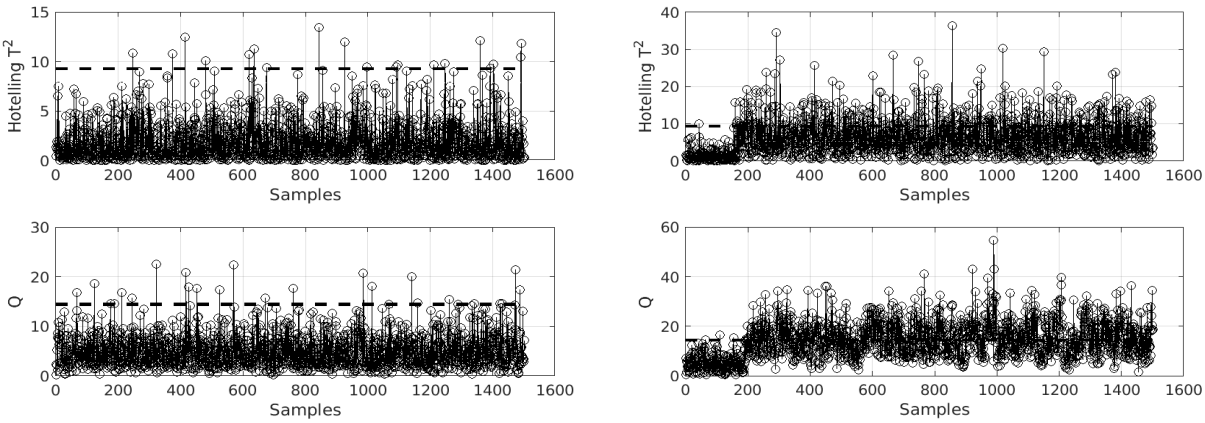
(a) **(b)**

Figure A.4: MSPC control charts pipe rupture fault a) intensity 1 and b) intensity 2



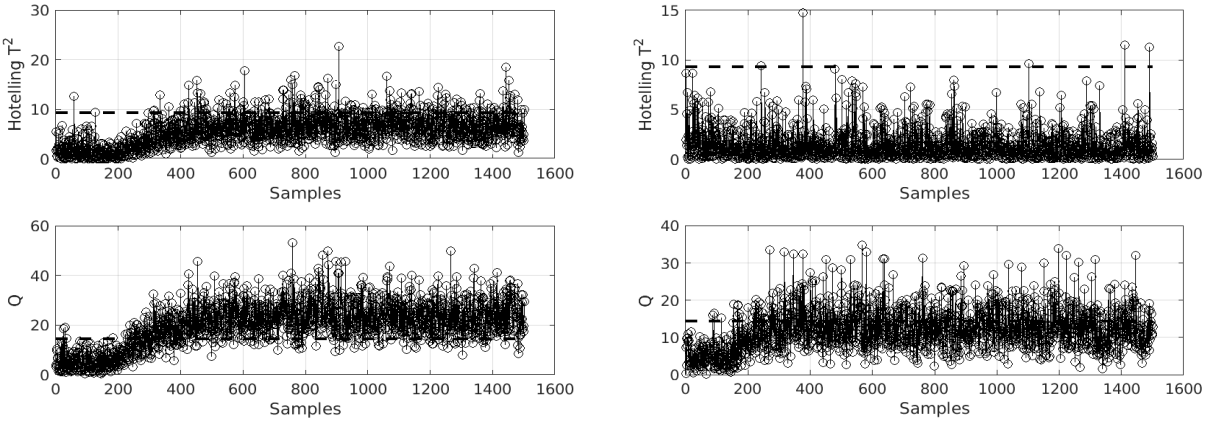
(a) **(b)**

Figure A.5: MSPC control charts jacket hole fault a) intensity 1 and b) intensity 2



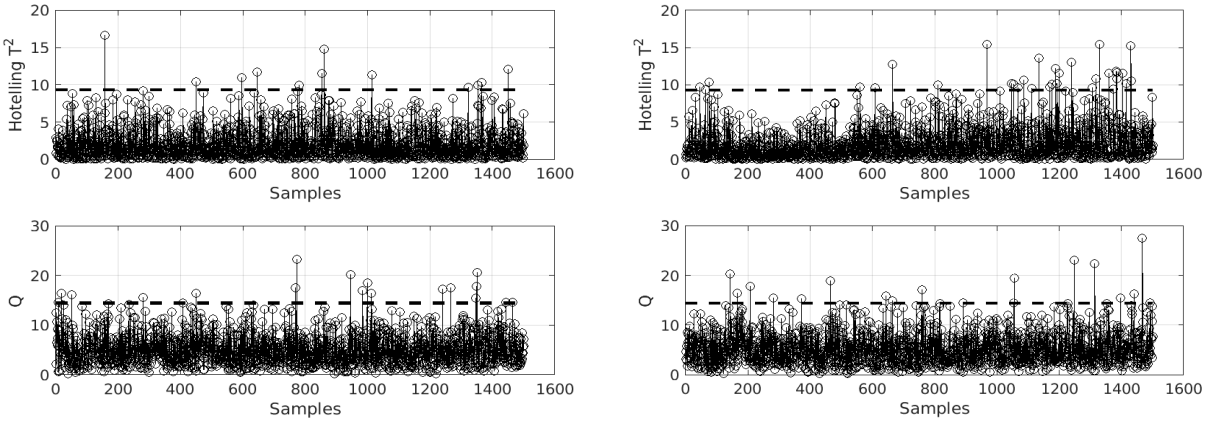
(a) **(b)**

Figure A.6: MSPC control charts tank and jacket holes fault a) intensity 1 and b) intensity 2



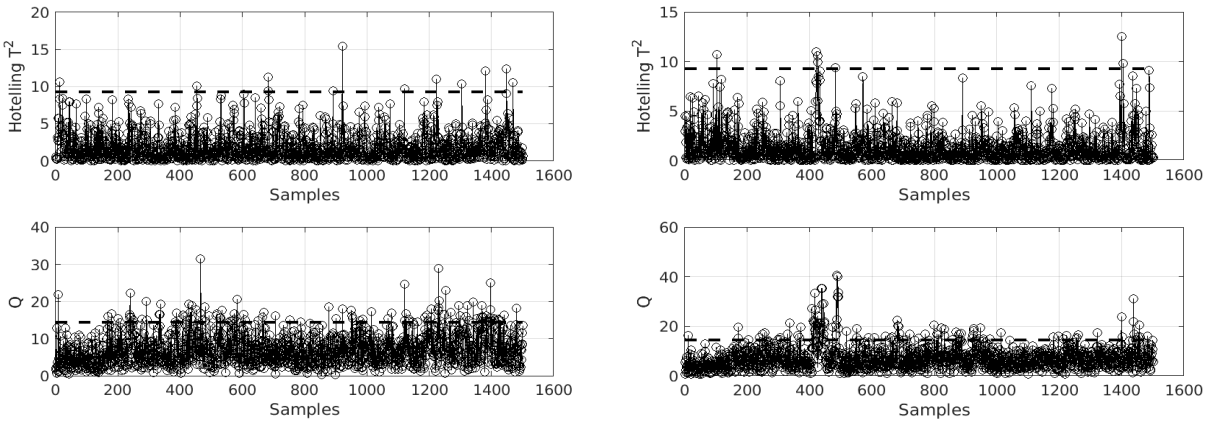
(a) **(b)**

Figure A.7: MSPC control charts valve lock fault a) intensity 1 and b) intensity 2



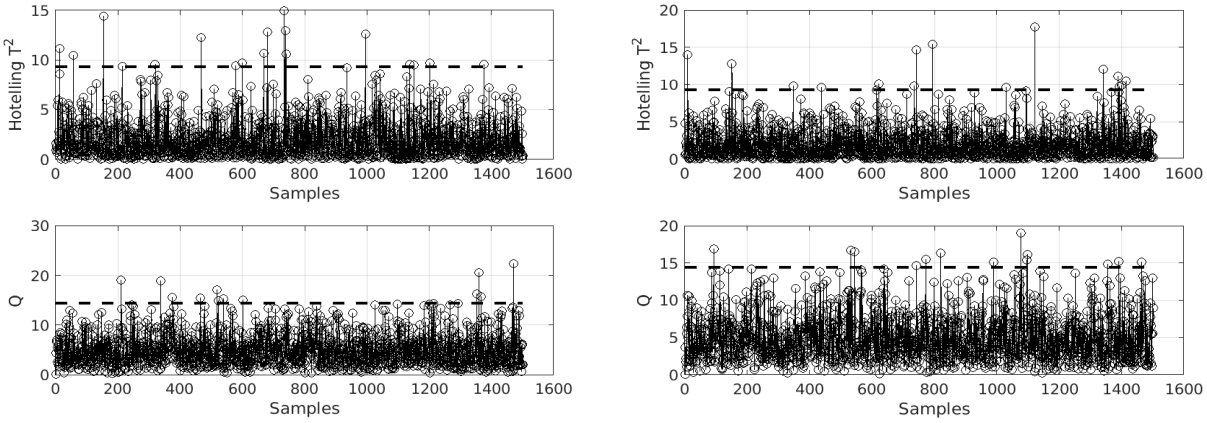
(a) **(b)**

Figure A.8: MSPC control charts valve offset fault a) intensity 1 and b) intensity 2



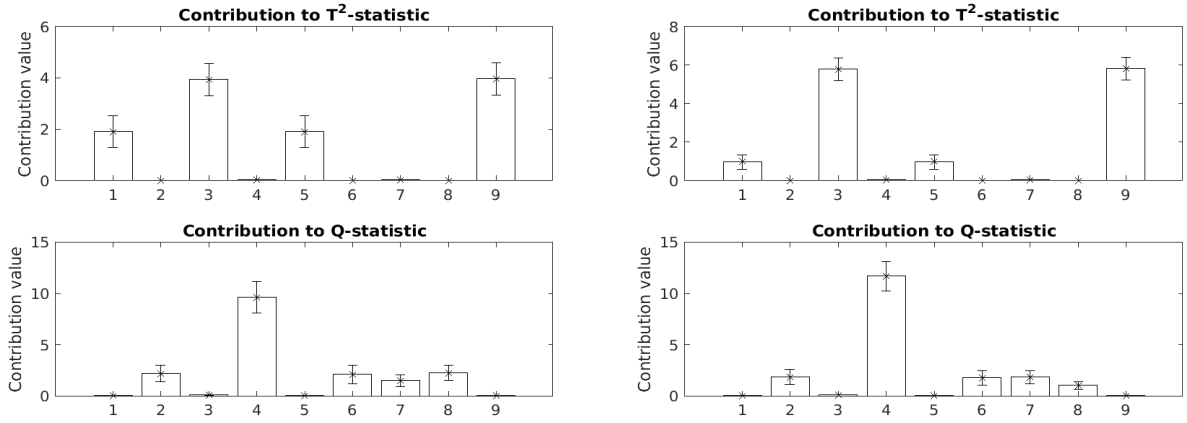
(a) **(b)**

Figure A.9: MSPC control charts valve stiction fault a) intensity 1 and b) intensity 2



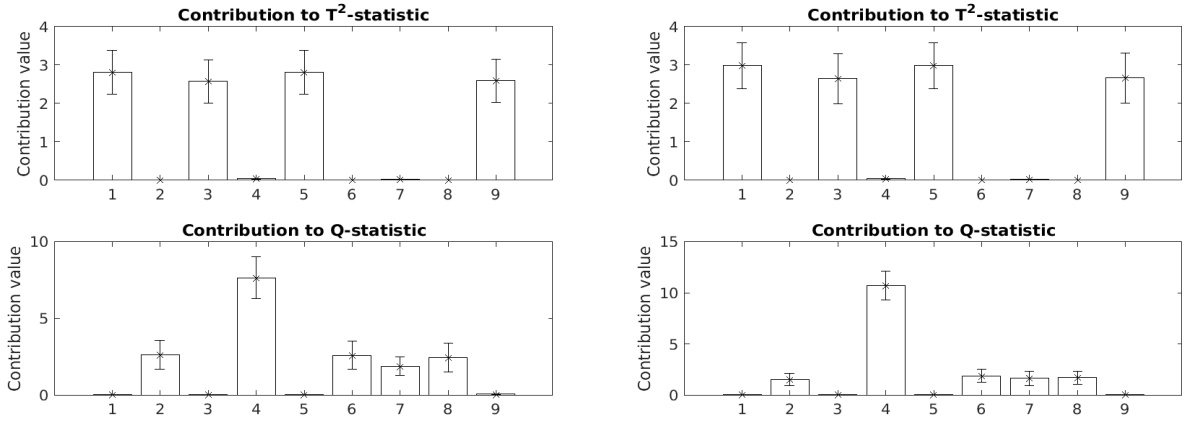
(a) **(b)**

Figure A.10: Contribution plots of level sensor fault a) intensity 1 and b) intensity 2



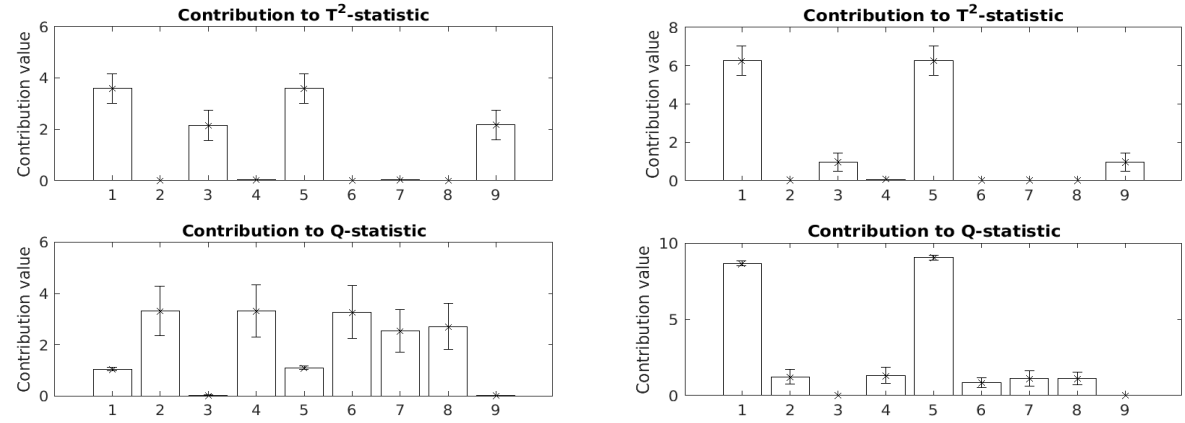
(a) **(b)**

Figure A.11: Contribution plots of mixture temperature sensor fault a) intensity 1 and b) intensity 2



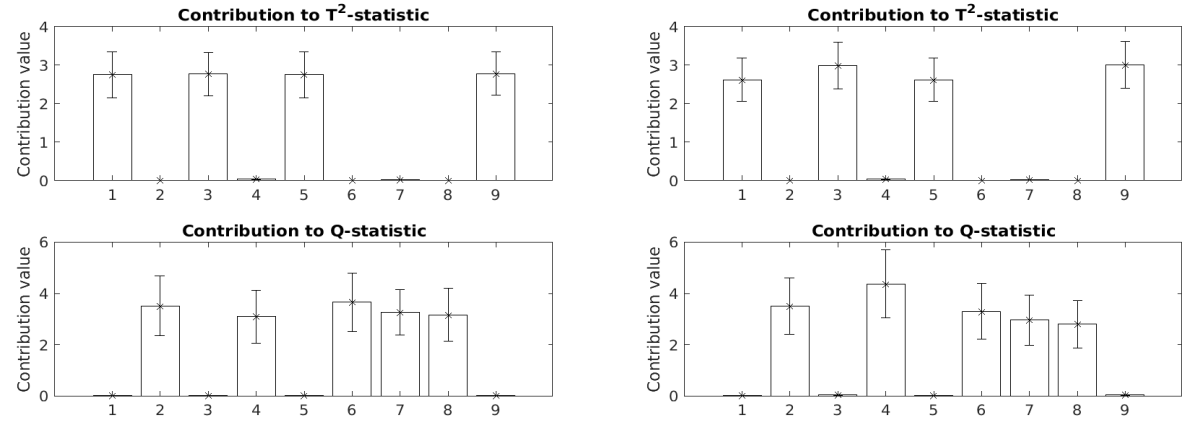
(a) **(b)**

Figure A.12: Contribution plots of jacket temperature sensor fault a) intensity 1 and b) intensity 2



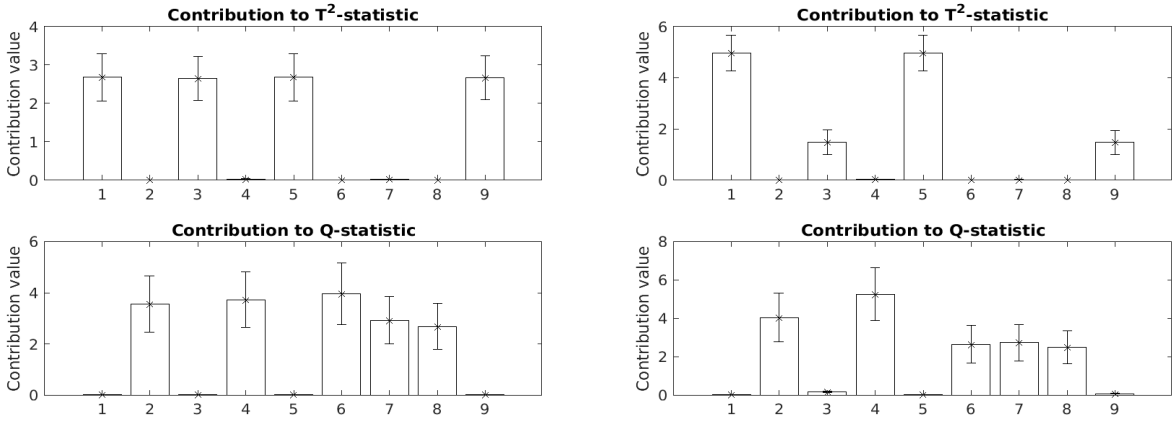
(a) **(b)**

Figure A.13: Contribution plots of pipe rupture fault a) intensity 1 and b) intensity 2



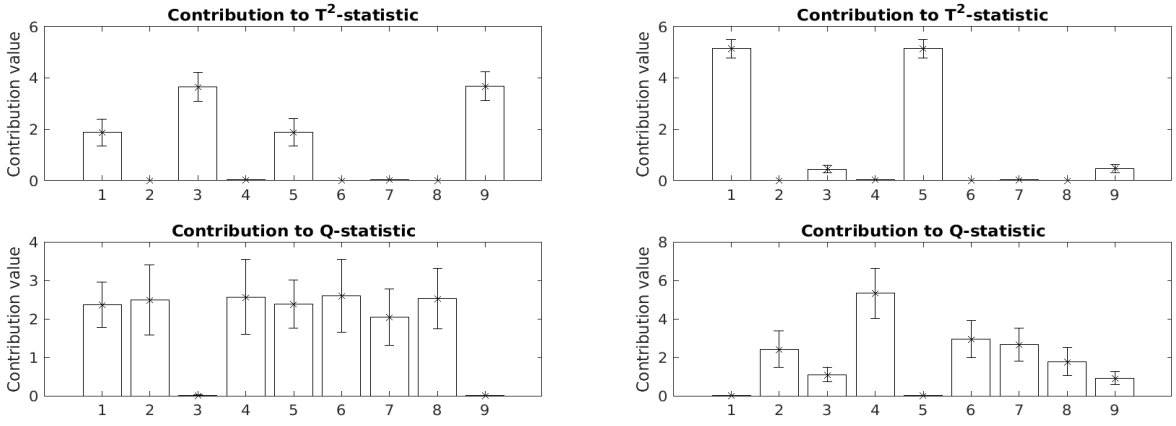
(a) **(b)**

Figure A.14: Contribution plots of jacket hole fault a) intensity 1 and b) intensity 2



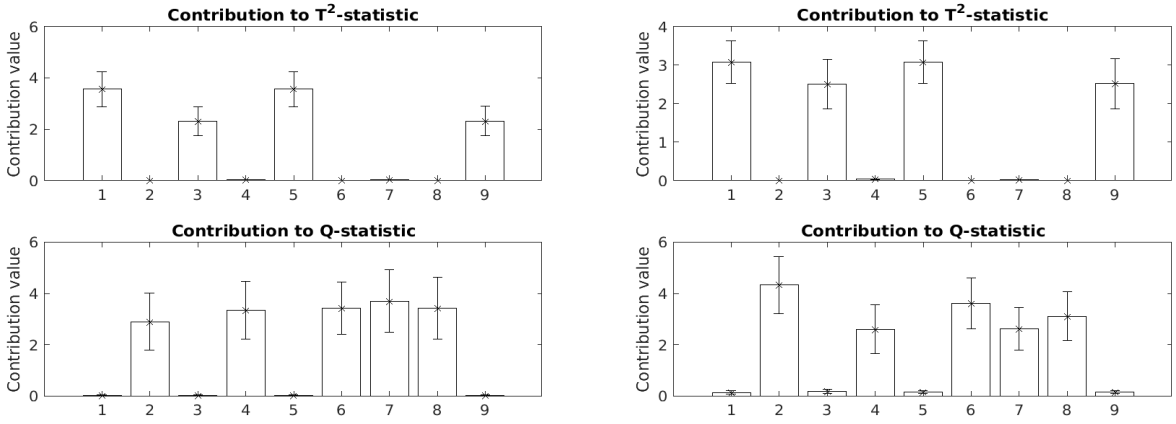
(a) **(b)**

Figure A.15: Contribution plots of tank and jacket holes fault a) intensity 1 and b) intensity 2



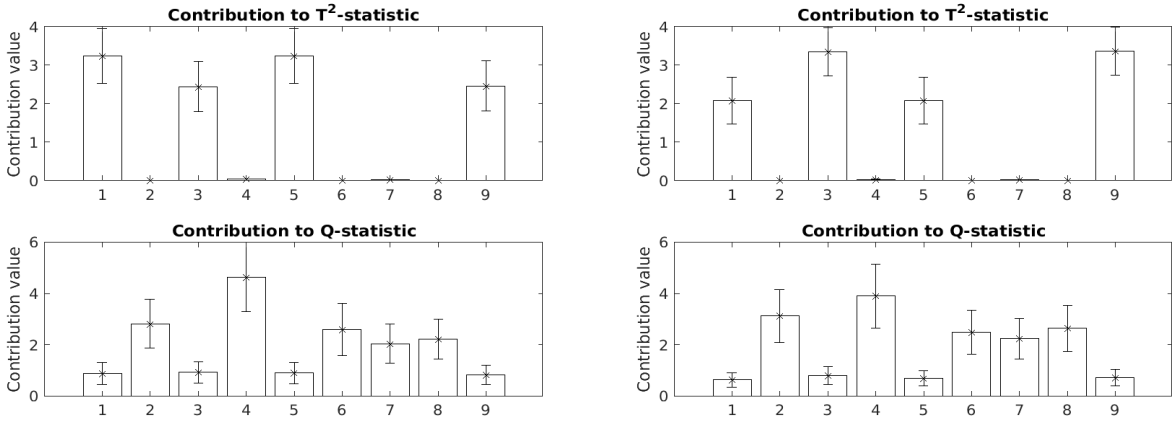
(a) **(b)**

Figure A.16: Contribution plots of valve lock fault a) intensity 1 and b) intensity 2



(a) **(b)**

Figure A.17: Contribution plots of valve offset fault a) intensity 1 and b) intensity 2



(a) **(b)**

Figure A.18: Contribution plots of valve stiction fault a) intensity 1 and b) intensity 2

Appendix B

Detection evaluation

Detection strength and speed are two predominantly important aspects to characterize the performance of monitoring techniques. Detection strength is defined as the methods ability to accurately detect unusual situations, without excessive false alarms. Detection speed regards the speed to signal said unusual situation, after it occurs. The considered aspects can be measured with indicators. For strength FAR,

$$\mathrm{FAR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} = \frac{\mathrm{FP}}{\mathrm{N}}\,, \tag{B.1}$$

where FP is the total false positives, TN is the total true negatives and N represents the total number of negatives. TDR which is defined as

$$\mathrm{TDR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} = \frac{\mathrm{TP}}{\mathrm{P}}\,, \tag{B.2}$$

TP is the sum of true positives, FN is the number of false negatives and P the total number of positives.

In regards to speed, the ARL is the number of runs, on average, the method takes to detect unusual situations. The ARL can be related to the ATS through the step size which gives

$$\mathrm{ATS} = \mathrm{ARL} \cdot \Delta t\,, \tag{B.3}$$

as the ATS is the time the method takes to signal an unusual situation (Rato et al. 2016).

Table B.1: Detection strength and speed parameters for level sensor fault.

Chart/Intensity	**TDR** (σ)	**FAR** (σ)	**ARL** (σ)	**ATS** (σ)
$T^2/1$	0.0102 (0.0028)	0.0110 (0.0077)	79.2300 (97.2617)	15.8460 (19.4523)
$T^2/2$	0.0124 (0.0035)	0.0109 (0.0084)	31.4300 (61.3162)	6.2860 (12.2634)
$T^2/3$	0.0146 (0.0033)	0.0122 (0.0093)	4.1200 (17.4508)	0.8240 (3.4902)
Q/1	0.0136 (0.0032)	0.0125 (0.0079)	68.3900 (60.1463)	13.6780 (12.0293)
Q/2	0.0136 (0.0035)	0.0123 (0.0092)	92.9400 (81.3364)	18.5880 (16.2673)
Q/3	0.0142 (0.0032)	0.0135 (0.0093)	72.3100 (80.5049)	14.4620 (16.1010)

Table B.2: Detection strength and speed parameters for mixture temperature sensor fault.

Chart/Intensity	TDR (σ)	FAR (σ)	ARL (σ)	ATS (σ)
$T^2/1$	0.0121 (0.0028)	0.0119 (0.0082)	61.3400 (85.4871)	12.2680 (17.0974)
$T^2/2$	0.0190 (0.0044)	0.0107 (0.0082)	6.1400 (13.2909)	1.2280 (2.6582)
$T^2/3$	0.0348 (0.0047)	0.0120 (0.0082)	1.8100 (1.1951)	0.3620 (0.2390)
Q/1	0.0599 (0.0083)	0.0150 (0.0094)	19.3800 (14.9205)	3.8760 (2.9841)
Q/2	0.3747 (0.0229)	0.0136 (0.0098)	8.5700 (3.3128)	1.7140 (0.6626)
Q/3	0.8539 (0.0131)	0.0139 (0.0111)	5.8300 (1.8094)	1.1660 (0.3620)

Table B.3: Detection strength and speed parameters for jacket temperature sensor fault.

Chart/Intensity	TDR (σ)	FAR (σ)	ARL (σ)	ATS (σ)
$T^2/1$	0.0103 (0.0023)	0.0109 (0.0084)	101.5900 (92.6753)	20.3180 (18.5351)
$T^2/2$	0.0107 (0.0031)	0.0119 (0.0091)	103.4700 (117.0923)	20.6940 (23.4185)
$T^2/3$	0.0108 (0.0026)	0.0108 (0.0080)	94.0100 (91.4385)	18.8020 (18.2877)
Q/1	0.0349 (0.0055)	0.0131 (0.0098)	35.1400 (37.9963)	7.0280 (7.5993)
Q/2	0.1494 (0.0125)	0.0149 (0.0099)	9.3900 (8.2436)	1.8780 (1.6487)
Q/3	0.4398 (0.0189)	0.0120 (0.0079)	2.7300 (2.7883)	0.5460 (0.5577)

Table B.4: Detection strength and speed parameters for jacket hole fault.

Chart/Intensity	TDR (σ)	FAR (σ)	ARL (σ)	ATS (σ)
$T^2/1$	0.0096 (0.0027)	0.0103 (0.0077)	103.2300 (96.1483)	20.6460 (19.2297)
$T^2/2$	0.0117 (0.0030)	0.0111 (0.0087)	87.4400 (74.6533)	17.4880 (14.9307)
$T^2/3$	0.4267 (0.0100)	0.0106 (0.0080)	13.4100 (4.7887)	2.6820 (0.9577)
Q/1	0.0129 (0.0058)	0.0136 (0.0091)	85.9400 (79.7462)	17.1880 (15.9492)
Q/2	0.0138 (0.0030)	0.0131 (0.0093)	64.1600 (78.1157)	12.8320 (15.6231)
Q/3	0.9608 (0.0034)	0.0123 (0.0091)	3.9400 (0.9301)	0.7880 (0.1860)

Table B.5: Detection strength and speed parameters for jacket and tank holes fault.

Chart/Intensity	TDR (σ)	FAR (σ)	ARL (σ)	ATS (σ)
$T^2/1$	0.0118 (0.0027)	0.0112 (0.0088)	83.6000 (70.2432)	16.7200 (14.0489)
$T^2/2$	0.2443 (0.0096)	0.0120 (0.0092)	15.4600 (6.6415)	3.0920 (1.3283)
$T^2/3$	0.9984 (0.0004)	0.0101 (0.0073)	3.0400 (0.6655)	0.6080 (0.1331)
Q/1	0.0138 (0.0034)	0.0137 (0.0094)	69.2900 (68.3468)	13.8580 (13.6694)
Q/2	0.4436 (0.0207)	0.0137 (0.0097)	28.8400 (14.3975)	5.7680 (2.8795)
Q/3	0.9962 (0.0011)	0.0131 (0.0092)	4.2800 (1.0644)	0.8560 (0.2129)

Table B.6: Detection strength and speed parameters for pipe rupture fault.

Chart/Intensity	TDR (σ)	FAR (σ)	ARL (σ)	ATS (σ)
$T^2/1$	0.0270 (0.0042)	0.0119 (0.0082)	44.5700 (30.2546)	8.9140 (6.0509)
$T^2/2$	0.4355 (0.0086)	0.0107 (0.0082)	3.8100 (6.2695)	0.7920 (1.2539)
$T^2/3$	0.9827 (0.0027)	0.0120 (0.0084)	1.0000 (0.0000)	0.2000 (0.0000)
Q/1	0.0919 (0.0052)	0.0150 (0.0094)	30.1300 (28.1273)	6.0260 (5.6255)
Q/2	1.0000 (0.0000)	0.0136 (0.0098)	1.0000 (0.0000)	0.2000 (0.0000)
Q/3	1.0000 (0.0000)	0.0139 (0.0111)	1.0000 (0.0000)	0.2000 (0.0000)

Table B.7: Detection strength and speed parameters for valve lock fault. There are ARLs and ATSs with a minor than sign because some of the runs never detect the fault. In these cases, the last point is considered as a detection.

Chart/Intensity	**TDR** (σ)	**FAR** (σ)	**ARL** (σ)	**ATS** (σ)
$T^2/1$	0.4800 (0.4042)	0.0119 (0.0082)	116.0600 (145.1190)	23.2120 (29.0238)
$T^2/2$	0.0046 (0.0048)	0.0107 (0.0082)	280.2100 (303.5997)	56.0420 (60.7199)
$T^2/3$	0.3693 (0.3991)	0.0120 (0.0084)	<497.8000 (555.0649)	<99.5600 (111.0130)
Q/1	0.7082 (0.3755)	0.0150 (0.0094)	24.1200 (22.8010)	4.8240 (4.5602)
Q/2	0.4496 (0.3894)	0.0136 (0.0098)	33.4600 (48.7775)	6.6920 (9.7555)
Q/3	0.8212 (0.2775)	0.0139 (0.0111)	16.8500 (21.9979)	3.3700 (4.3996)

Table B.8: Detection strength and speed parameters for valve offset fault. There are ARLs and ATSs with a minor than sign because some of the runs never detect the fault. In these cases, the last point is considered as a detection.

Chart/Intensity	**TDR** (σ)	**FAR** (σ)	**ARL** (σ)	**ATS** (σ)
$T^2/1$	0.0099 (0.0027)	0.0109 (0.0086)	51.3900 (81.0059)	10.2780 (16.2012)
$T^2/2$	0.01977 (0.0041)	0.0111 (0.0080)	373.9300 (130.0408)	74.7860 (26.0082)
$T^2/3$	0.0001 (0.0000)	0.0116 (0.0096)	<1349.7000 (13.0000)	<269.94 (2.6000)
Q/1	0.0129 (0.0033)	0.0121 (0.0098)	65.4200 (67.2508)	13.0840 (13.4502)
Q/2	0.0144 (0.0224)	0.0131 (0.0101)	71.4200 (75.2116)	14.2840 (15.0423)
Q/3	0.0224 (0.0079)	0.0130 (0.0095)	29.1100 (49.8093)	5.8220 (9.9619)

Table B.9: Detection strength and speed parameters for valve stiction fault.

Chart/Intensity	**TDR** (σ)	**FAR** (σ)	**ARL** (σ)	**ATS** (σ)
$T^2/1$	0.0108 (0.0033)	0.0120 (0.0082)	112.0800 (95.5398)	22.4160 (19.1080)
$T^2/2$	0.0041 (0.0021)	0.0116 (0.0076)	270.6300 (234.5675)	54.1260 (46.9135)
$T^2/3$	0.00065 (0.0024)	0.0115 (0.0085)	199.2100 (194.2100)	39.8420 (38.8109)
Q/1	0.0946 (0.0152)	0.0155 (0.0100)	18.1900 (15.7843)	3.6380 (3.1569)
Q/2	0.0644 (0.0129)	0.0128 (0.0082)	21.5500 (20.6918)	4.3100 (4.1384)
Q/3	0.0768 (0.0126)	0.0139 (0.0112)	21.5500 (19.3053)	4.3100 (3.8611)

Appendix C

Multiresolution analysis

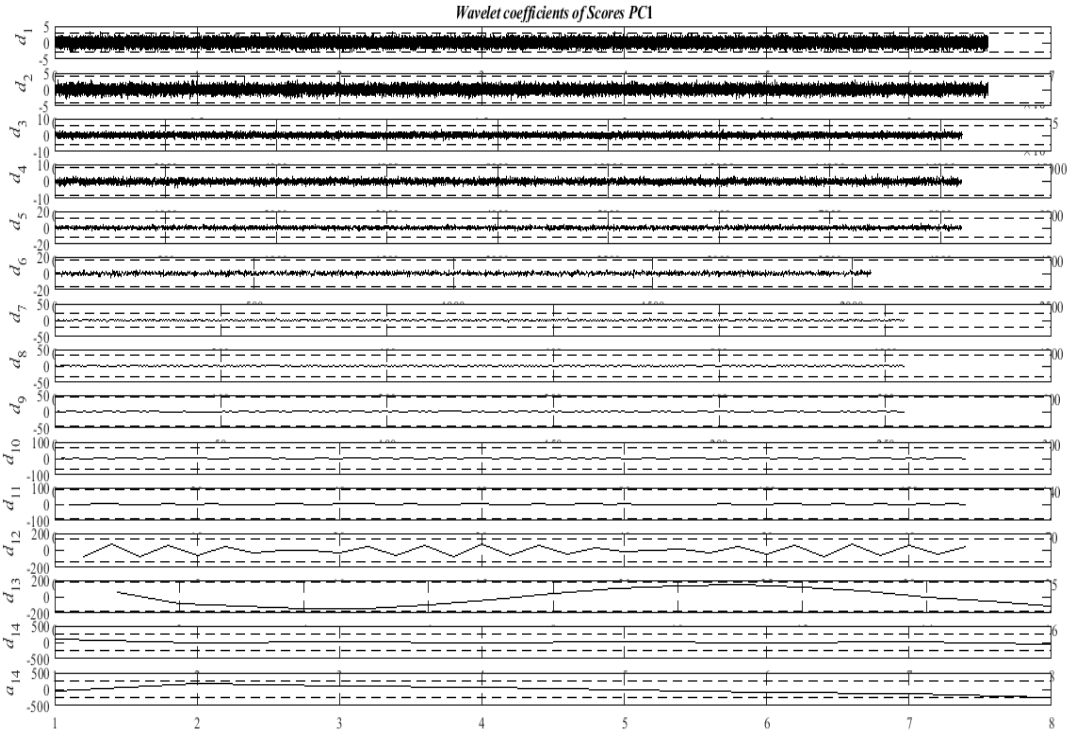


Figure C.1: Wavelet decompostion of the scores of PC1.

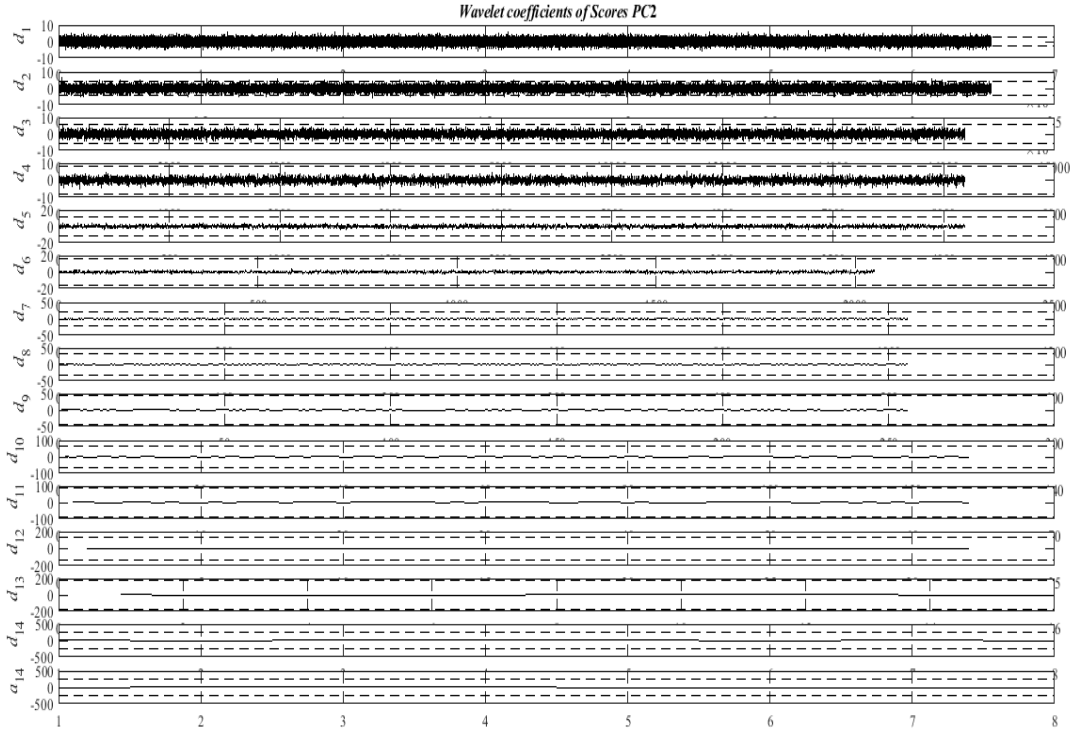


Figure C.2: Wavelet decompostion of the scores of PC2.

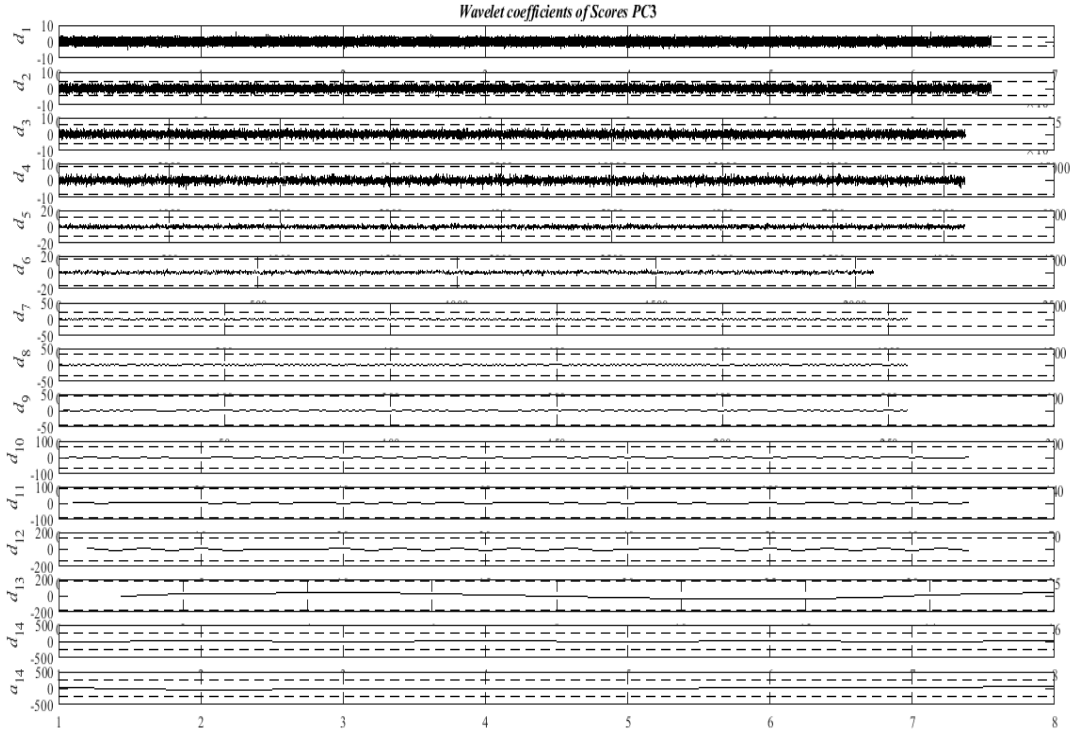


Figure C.3: Wavelet decompostion of the scores of PC3.

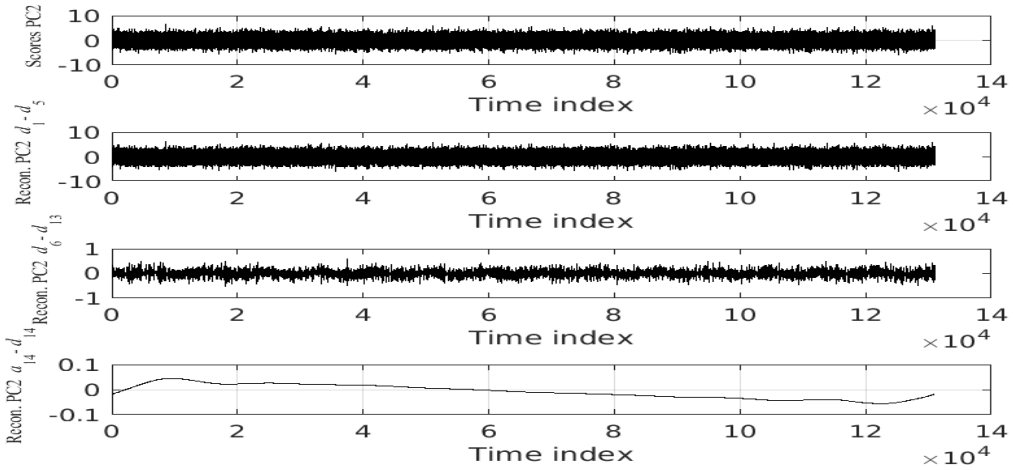


Figure C.4: From top to bottom: Original PC2 scores signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

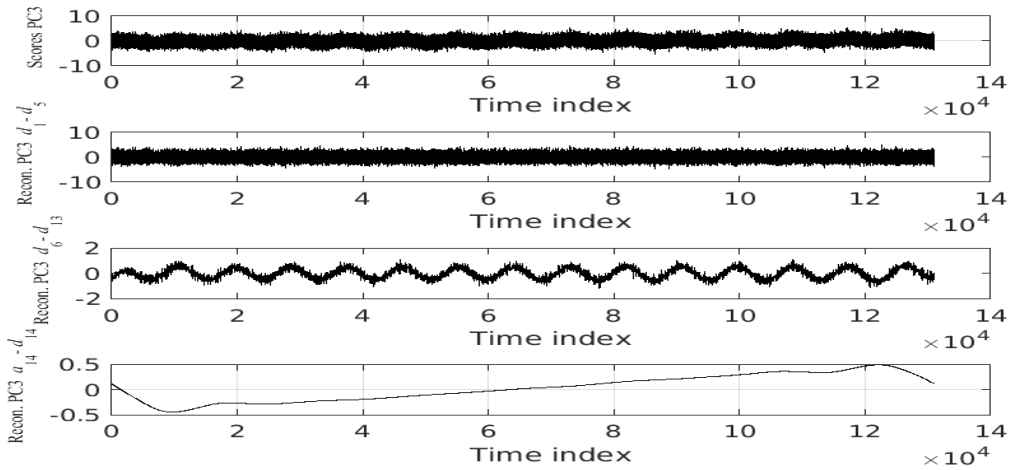


Figure C.5: From top to bottom: Original PC3 scores signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

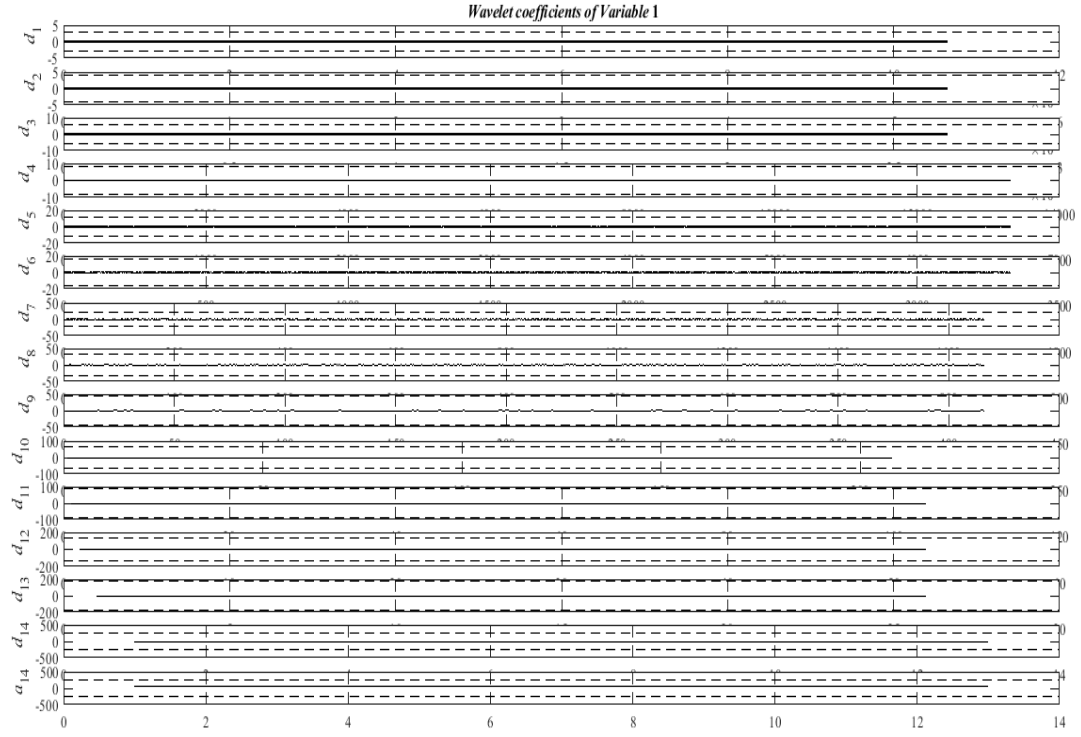


Figure C.6: Wavelet decompostion of the variable 1.

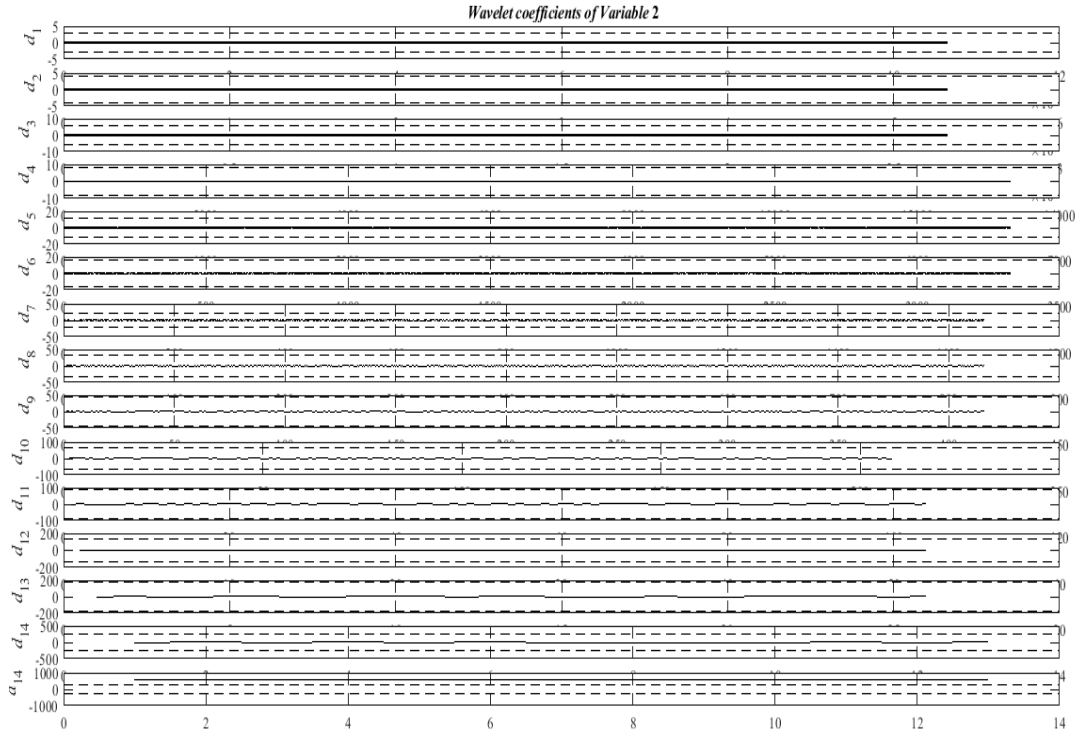


Figure C.7: Wavelet decompostion of the variable 2.

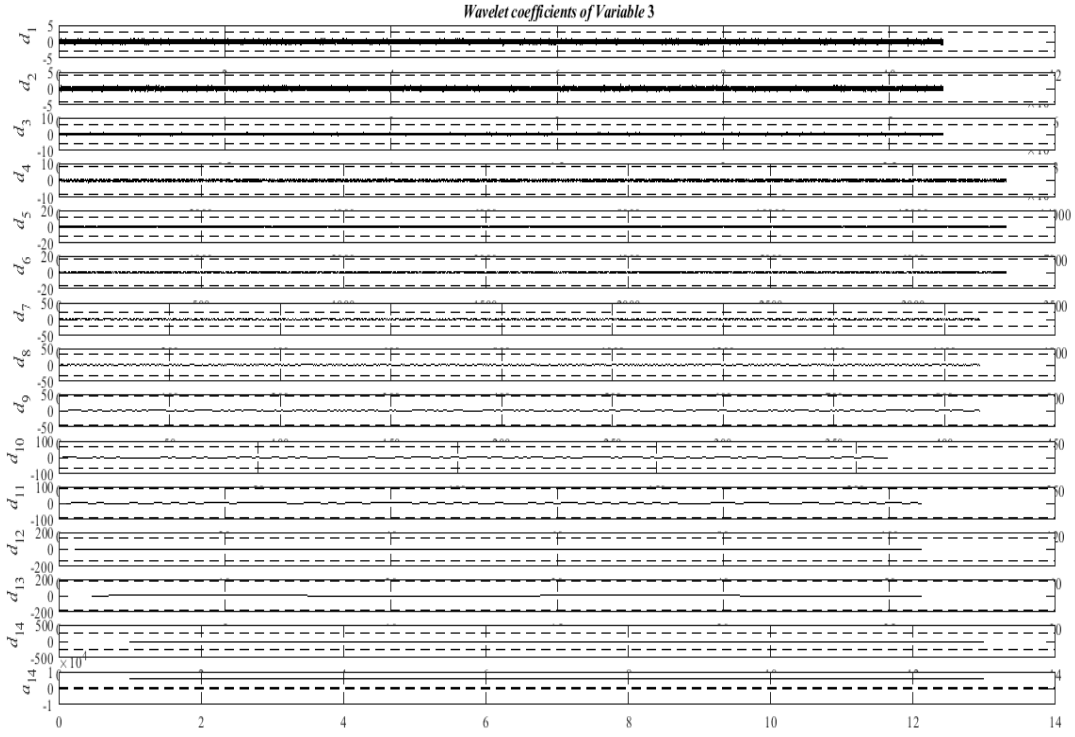


Figure C.8: Wavelet decompostion of the variable 3.

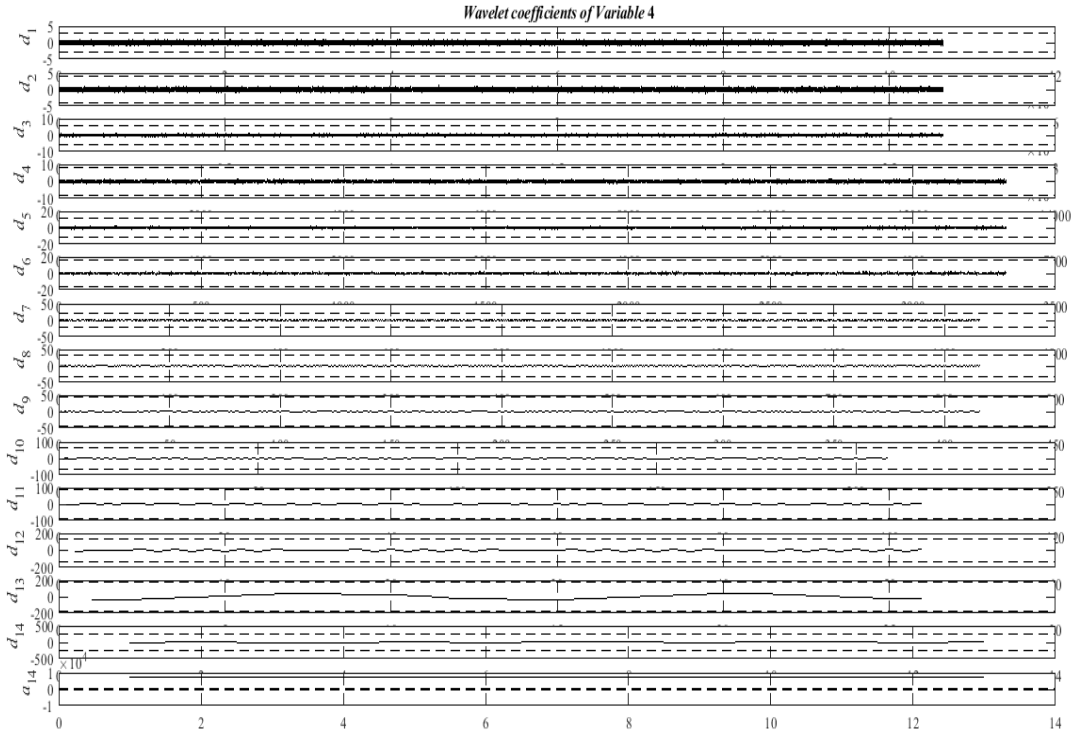


Figure C.9: Wavelet decompostion of the variable 4.

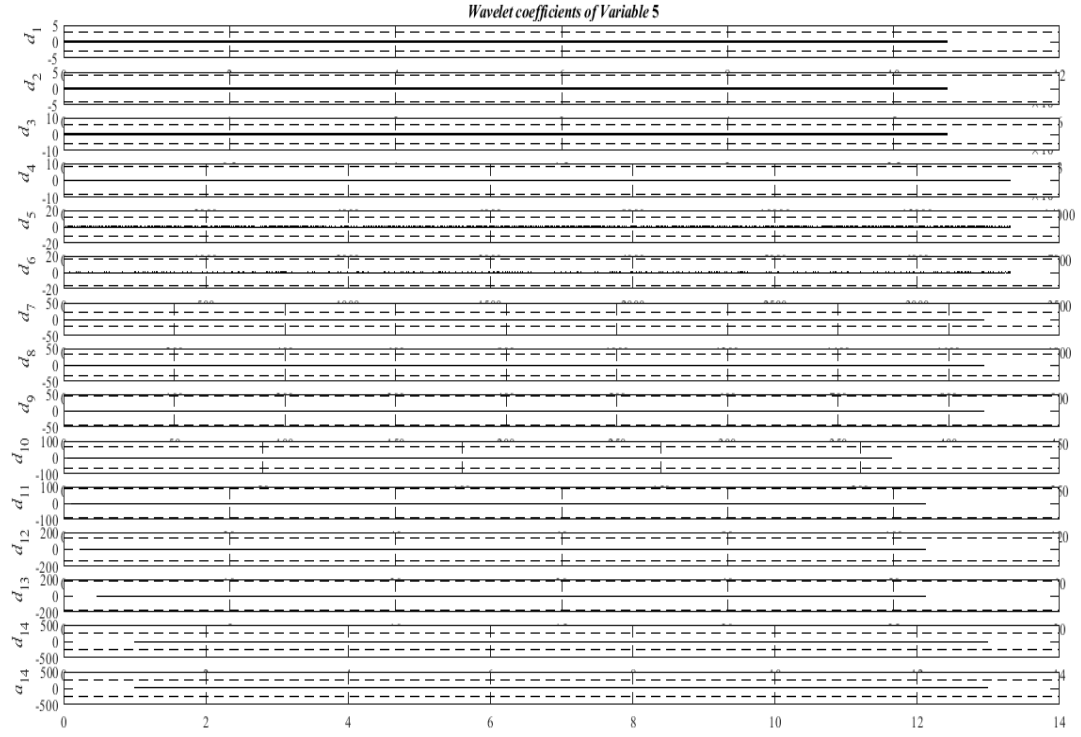


Figure C.10: Wavelet decompostion of the variable 5.

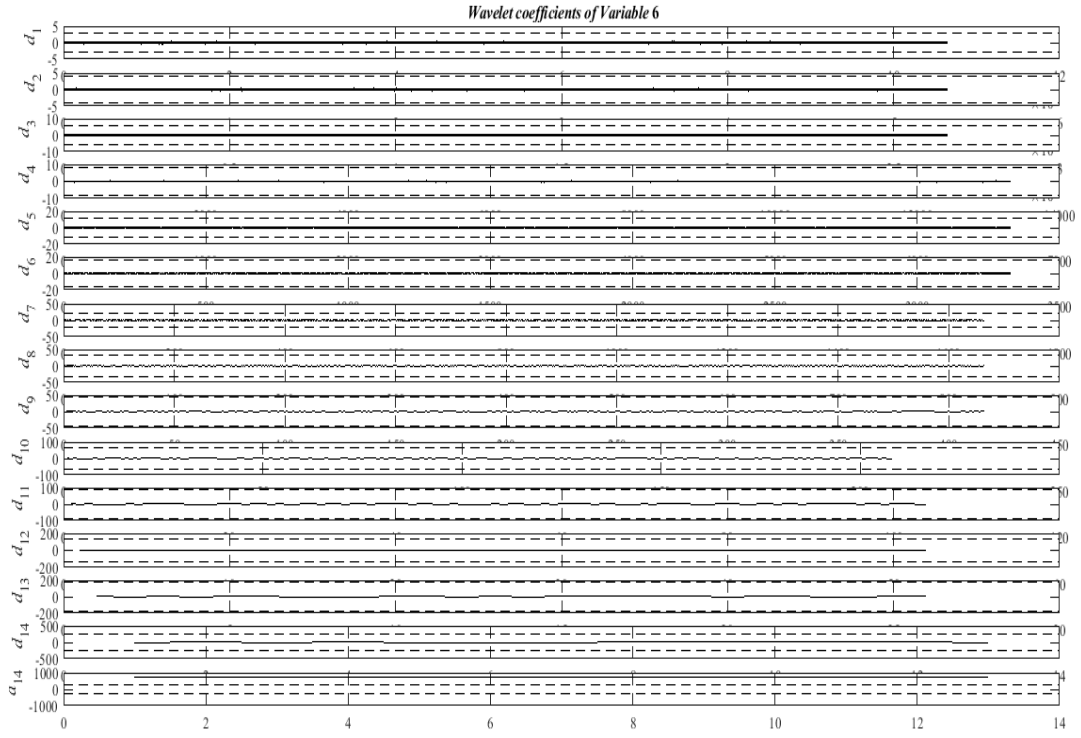


Figure C.11: Wavelet decompostion of the variable 6.

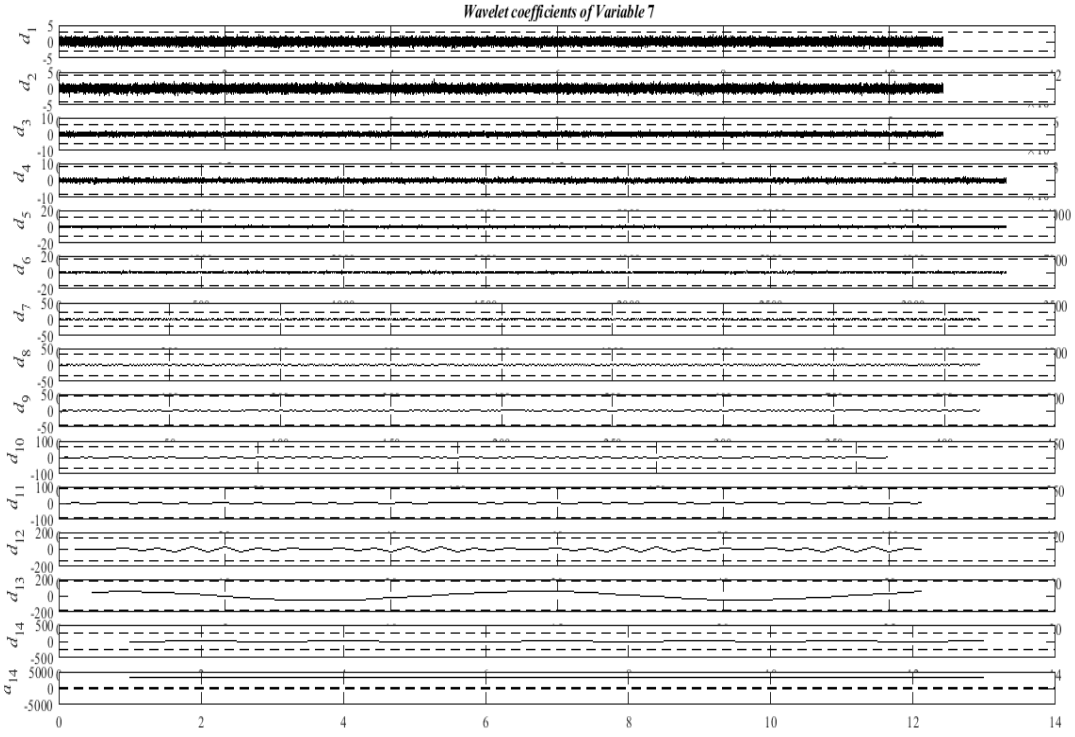


Figure C.12: Wavelet decompostion of the variable 7.

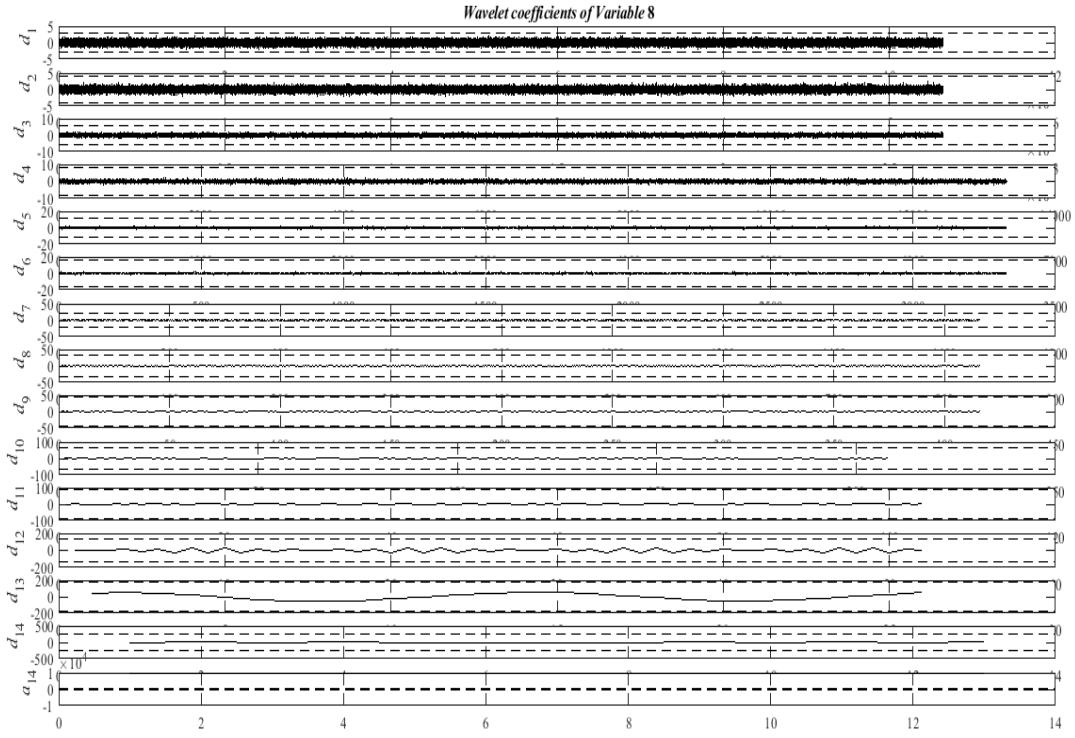


Figure C.13: Wavelet decompostion of the variable 8.

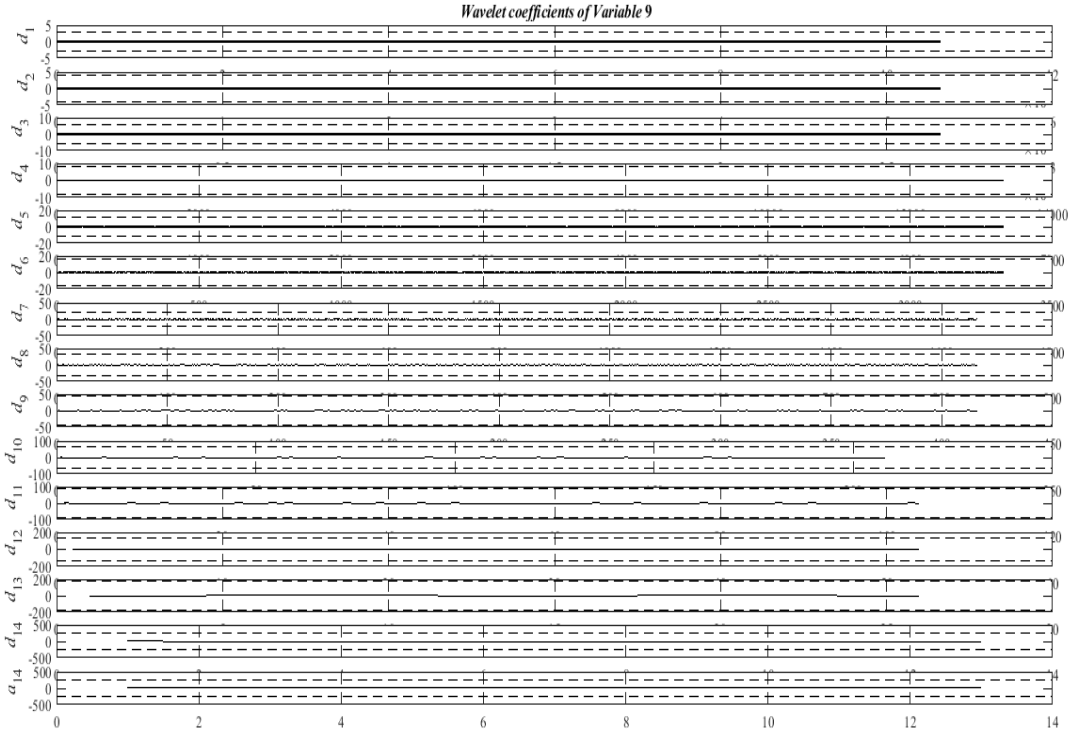


Figure C.14: Wavelet decompostion of the variable 9.

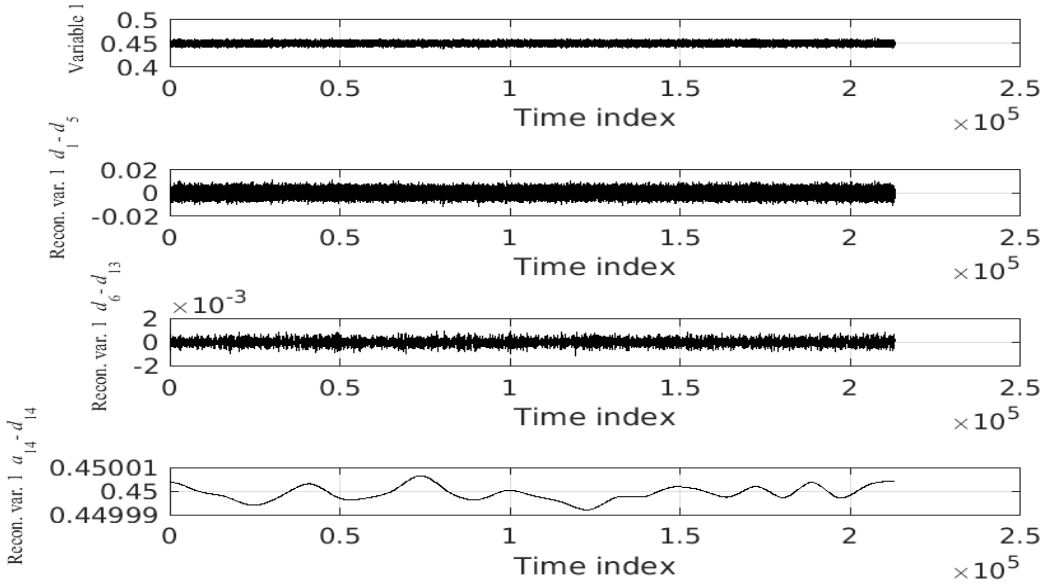


Figure C.15: From top to bottom: Original level variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

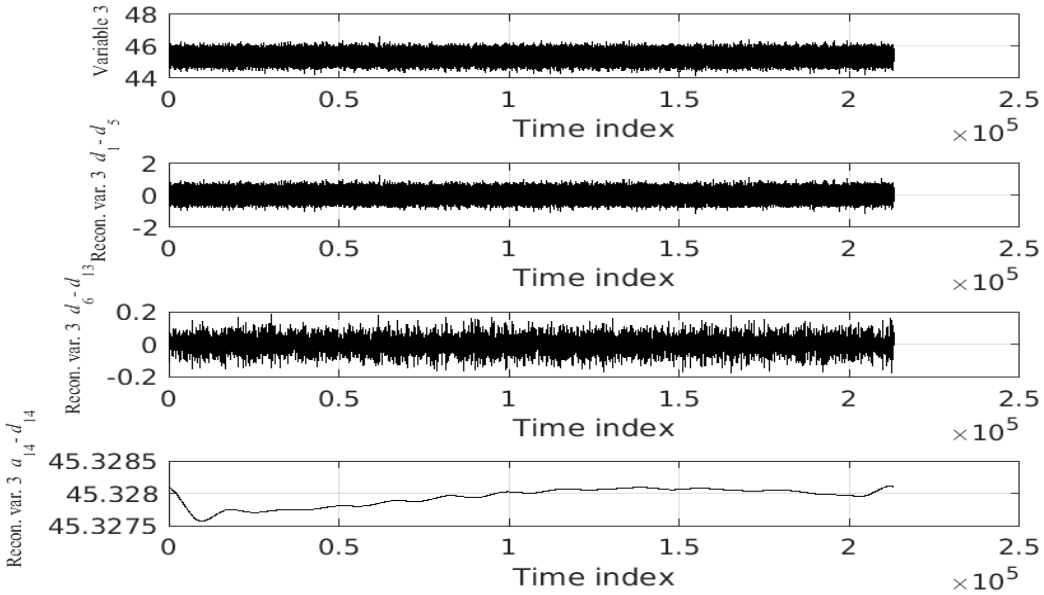


Figure C.16: From top to bottom: Original mixture temperature variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

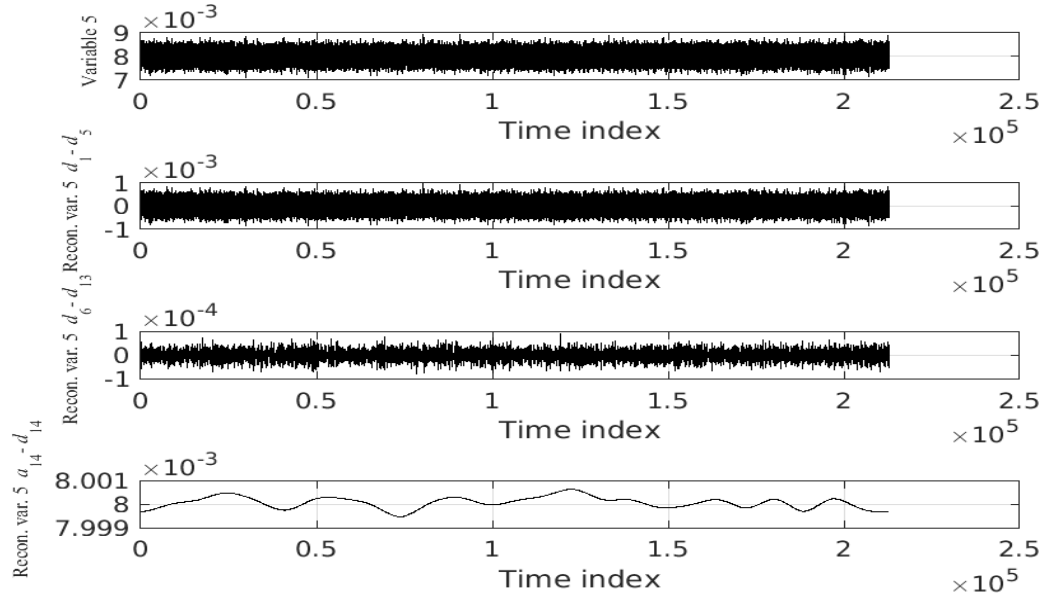


Figure C.17: From top to bottom: Original feed flow rate variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

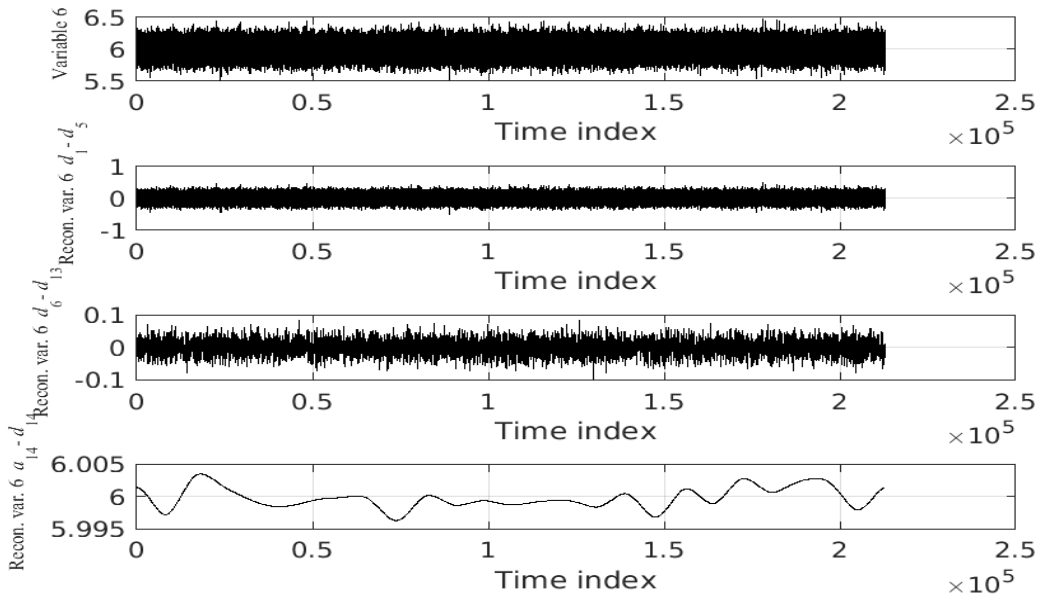


Figure C.18: From top to bottom: Original feed concentration of A variable signal; wavelet reconstruction of the signal with details coefficients from d_1 to d_5; wavelet reconstruction of the signal with details coefficients from d_6 to d_{13}; and wavelet reconstruction of the signal with detail coefficient d_{14} and approximation coefficient a_{14}.

Appendix D

Health and safety risk assessment

The physical and psychological stress to which a programmer is subjugated daily is predominantly related to the nature of his work, coupled with high hours of continuous work and excessive workload. In this way, it becomes clear the importance of the detailed study of the health problems that come from the practice of this type of activity.

There are five main health problems that can affect computer professionals:

Thrombosis - called the formation of a blood clot in a blood vessel, especially of the legs, which can travel to the lungs or brain causing strokes or pulmonary embolisms, affects mainly professionals who spend a great number of hours sitting. One form of prevention is to take breaks and simply get up for short periods of time.

Heart disease - several studies have already shown an increased risk of heart disease in those who spend a large number of hours sitting. As a prevention, short intervals are recommended in conjunction with periodic physical exercises.

Carpal Tunnel Syndrome - fairly common in computer users, this syndrome is caused by the compression of the median nerve, the main nerve of the wrist, after continuous physical stress. Regular preventative exercises should be done before and after activity along with the use of an ergonomically well-prepared workstation.

Anxiety, stress, and depression - programmers typically manage crises and disaster recovery, which, in addition to the inherent stress that can catapult into the onset of physical symptoms, can cause mood swings, anxiety, and depression. An effective way to combat stress and anxiety can be exercise. For depression it is recommended to restrict the number of hours to the computer, especially time spent on the internet, during rest hours.

Lower back pain - spending too much sitting hours combined with poor posture can cause irreparable damage to the spinal cord, leading to chronic pain. As a precaution, attention to posture and the suitability of the workplace are fundamental to physical health.

Neck and eye strain - improperly adjusted monitors and a regular computer usage can lead to neck and eye strain. The adequacy of working conditions are also fundamental for improving the health of professionals.