CRANFIELD UNIVERSITY

T. SEDIGHI

DIAGNOSTIC AND PROGNOSTIC OF INTERMITTENT FAULTS (BY USE OF MACHINE LEARNING)

SCHOOL OF AEROSPACE, TRANSPORT AND MANUFACTURING
Through-Life Engineering Services Centre

PhD
Academic Year: 2020–2021

Supervisor: Prof. P. D. Foote and Prof. A. Starr
February 2020

CRANFIELD UNIVERSITY

SCHOOL OF AEROSPACE, TRANSPORT AND MANUFACTURING
Through-Life Engineering Services Centre

PhD

Academic Year: 2020–2021

T. SEDIGHI

Diagnostic and Prognostic of Intermittent Faults (by Use of Machine Learning)

Supervisor: Prof. P. D. Foote and Prof. A. Starr
February 2020

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

Abstract

This thesis investigates novel intermittent fault detection and prediction techniques for complex nonlinear systems.

Aerospace and defence systems are becoming progressively more complex, with greater component numbers and increasingly complicated components and subcomponents. At the same time, faults and failures are becoming more challenging to detect and isolate, and the time that operators and maintenance technicians spend on faults is rising.

Moreover, a serious problem has recently attracted a lot of attention in health diagnostics of these complex systems. Detecting intermittent faults that persist for very short durations and manifest themselves intermittently have become troublesome and sometimes impossible (also known as "no fault found").

In response to the above challenges, this thesis focuses on the development of a novel methodology to detect intermittent faults of these complex systems. It further investigates various probabilistic approaches to develop efficient fault diagnostic and prognostic methods.

In the first stage of this thesis, a novel model (observer)-based intermittent fault detection filter is presented that relies on the creation of a mathematical model of a laboratory-scale aircraft fuel system test rig to predict the output of the system at any given time. Comparison between this prediction of output and actual output reveals the presence of a fault. Later, the simulation results demonstrate that the performance of the model (observer)-based fault detection techniques decrease significantly as system complexity increases.

In the second stage of this research, a probabilistic data-driven method known as a Bayesian network is presented. This is particularly useful for diverse problems of varying size and complexity, where uncertainties are inherent in the system. Bayesian networks that model sequences of variables are called dynamic Bayesian networks. To introduce the time variable in the framework of probabilistic models while dealing with both discrete and continuous variables in the fuel rig system, a hybrid dynamic Bayesian network is proposed.

The presented results of data-driven fault detection show that the hybrid dynamic Bayesian network is more effective than the static Bayesian network or model (observer)-based methods for detecting intermittent faults.

Furthermore, the second stage of the research uses all the information captured from the fault diagnostic techniques for intermittent fault prediction by using a probabilistic non-parametric Bayesian method called Gaussian process regression, which is an aid for decision-making using uncertain information.

Keywords

No Fault Found; Intermittent fault; Model-based fault detection; Nonlinear systems, Nonlinear observer; Bayesian methods; Hybrid dynamic Bayesian network, Gaussian process regression.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors Prof. Peter D. Foote and Prof. Andrew Star for the continuous support of my Ph.D study and related research, for their patience, motivations, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis.

Besides my supervisors, I would like to thank the rest of my thesis committee for their insightful comments and encouragement, during my Ph.D study.

My sincere thanks also goes to Dr. Octavian Niculita who gave access to the IVHM laboratory and research facilities. Without his precious support it would not be possible to conduct this research.

Last but not the least, I would like to thank my family: my Mum, Parvaneh, my dad, Assadollah, my husband, Dr Alireza Daneshkhah, my daughters, Parinaz and Darya Khatoon and to my brothers, Amir and Salar for supporting me spiritually throughout writing this thesis and my life in general.

Moreover, the author gratefully acknowledges the support of EPSRC and BAE Systems for primarily funded this PhD project.

The project is part of the No Fault Found (NFF) research group's work at the Through-life Engineering Services (TES) Center, Cranfield University.

Contents

List of Figures

List of Tables

list of Abbreviations

ARE	Algebraic Riccati Equation
AI	Artificial Intelligece
ANN	Artificial Neural Network
BBN	Bayesian Belief Network
BFS	Best Fit Straight Line
BITE	Built-in Test Equipment
BN	Bayesian Network
BNFD	Bayesian Network Fault Detection
BNT	Bayesian Network Toolbox
CND	Can Not Duplicate
CPD	Conditional Probability Distribution
CPT	Conditional Probability Table
DT	Decision Trees
DAG	Direct Acyclic Graph
DBN	Dynamic Bayesian Network
DPV	Direct-acting Proportional Valve
EPSRC	Engineering Physical Science Research Council
ER	Erroneous Removal
FD	Fault Detection
FDI	Fault Detection and Isolation
FP	Fault Prediction
GUI	Graphical User Interface
GP	Gaussian Process
GPML	Gaussian Process Machine Learning
GPR	Gaussian Process Regression
HDBN	Hybrid Dynamic Bayesian Network
HMM	Hidden Markov Model
IFDP	Intermittent Fault Detection and Prediction
IVHM	Integrated Vehicle Health Management
JPD	Joint Probability Distribution
ML	Machine Learning
MSD	Mass Spring Damper
MI	Matrix Inequality
MCMC	Markov Chain Monte Carlo
MPLA	Maximum Posterior Likelihood Algorithm
MIMO	Multi Input Multi Output

MLP	Multilayer Perception
NDF	No Defect Found
NN	Neural Network
NFF	No Fault Found
NPF	No Problem Found
NTF	No Trouble Found
NUIO	Nonlinear Unknown Input Observer
NUI	Nonlinear Unknown Input
PCA	Principal Component Analysis
PLS	Partial Least Squares
RTOK	Re-test OK
RUL	Remaining Useful Life
SOV	Shut Off Valve
SIMO	Single Input Multi Output
SVM	Support Vector Machine
SPD	Symmetric Positive Definite
TES	Through-life Engineering Services
UAV	Unmanned Aerial Vehicle
UIO	Unknown Input Observer
VEA	Variable Elimination Algorithm

Chapter 1

Introduction

1.1 Overview

A system is made serviceable, functional, and worthy when its fault detection process is successful meaning that the fault in the system is detected and its main root causes are precisely identified.

When the symptoms of a fault are consistent (hard fault), it is usually not difficult to detect, isolate, and repair. However, a fault that persists for a very short duration and manifests itself intermittently and only during a particular set of operational stresses can be extremely difficult to identify and isolate (Khan *et al.*, 2014a). A substantial portion of malfunctions attributed to these type of faults is categorized as No Fault Found (NFF) (Syed *et al.*, 2016; Khan *et al.*, 2014b).

- **What is no fault found?**

NFF is a term referring to a system or component that has been returned to the manufacturer or distributor for a warranty replacement or service repair but operates properly when tested. This situation is also referred to as No Defect Found (NDF) and/or No Trouble Found (NTF)(Khan *et al.*, 2014b).

Understanding failure from a technical and non-technical perspective has begun to become increasingly important when considering a class of system faults which cannot be easily located, diagnosed or even reproduced under standard maintenance testing. This has been commonly termed the NFF phenomena (Khan *et al.*, 2014a).

The NFF phenomenon has a negative impact upon critical system stakeholder, requirement, which at the top level include systems safety, dependability and life-cycle costs (Khan *et al.*, 2014b). As a result, it is essential to prevent the causes of NFF, or at the very least to reduce the impacts and consequences. However achieving this requires the ability to understand, identify and correct the root causes of NFF events while the main causes of NFF can be categorized into the following set of classes (Khan *et al.*, 2014b):

- Technical,
- Organizational,

- Procedural,
- Behavioural.

In the context of aircraft engineering and maintenance, NFF is a chain of events that develops from a pilot experiencing a system malfunction with post-flight maintenance failing to reproduce the reported symptoms. Without any repair being undertaken, the malfunction may be experienced again on the subsequent flights. This present significant cost impacts to the industry that includes financial, reduced operational achievement, airworthiness challenges and potential flight safety issues (Khan *et al.*, 2014b).

One of the major causes identified for NFF occurrence within electronic, mechanical and hydraulic products are faults that are intermittent (Qi *et al.*, 2008). This makes it difficult to use systematic fault detection techniques effectively, as a system are subject to model uncertainties and unknown disturbances. The philosophy behind this criterion is that the designed fault detection and prediction filters should be robust to unknown inputs (disturbances) but sensitive to intermittent faults where the occupancy of intermittent faults can be alarmed by the use of an adaptive threshold. One of the aims of this thesis is to demonstrate the development of such methodologies and to examine their performances in a real-world testbed.

When a fault manifests itself intermittently, a big part of the system's failure is assigned to these Intermittent Faults (IF)s, and therefore they are considered as one of the main root causes of NFF issues. Other causes for NFF after the IFs are technician experience, troubleshooting manuals, and NFF diagnostics training. Likewise, Built-in Test Equipment (BITE) and software are the lowest-ranked causes for NFF.

- **What is intermittent fault?**

In literature, these type of faults is usually expressed as NFF phenomena (Mashkov *et al.*, 2019; Anderson & Synaptics, 2015). Other names used across the literature include: Erroneous Removal (ER), No Problem Found (NPF), Can Not Duplicate (CND), and Re-Test is OK (RTOK) (Khan *et al.*, 2014a).

Such faults (discontinuities) start with a very short duration and mostly of low amplitude. Therefore, to diagnose such faults, an on-line and, in many cases, real-time diagnostics system is required. Moreover, the intermittent fault is concerned with the dynamic behaviour of discrete events (intermittent failure) and reset events continuously along any path of the system's evolution. An IF is a loss of some function or performance characteristic in a product for a limited period and subsequent recovery of the function (Syed *et al.*, 2016). IFs can manifest themselves in any system, mechanical or electronic, in an unpredictable manner. The assumed unpredictability of an IF means that it cannot be easily detected or predicted and that it is not necessarily repeatable during maintenance testing. Faults of this nature raise many concerns in the realm of through-life engineering of products (Khan *et al.*, 2014b) .

In (Zhang *et al.*, 2018a) the intermittent fault detection and the event-triggered filtering has been investigated for a class of uncertain continuous systems with sensor saturation. In their deigned intermittent fault detection filters they used the upper bound of the system errors as their threshold to detect the intermittent faults which may cause to miss detection or false detection of the intermittent faults. They also only considered the

noise and uncertainties in the parameters not any unmatched uncertainties in the system which is one of the limitations in designing a fault detection filters.

If an IF is left unattended or missed during standardized maintenance testing, by its very definition it will reoccur at some time in the future and typically will tend to worsen over time. It will eventually become substantial enough to be detected with conventional test equipment (Khan *et al.*, 2014b).

In aerospace and defence avionics, IFs are the result of maintenance disruption, degradation mechanisms and the performing environment, which cause the degradation and ageing of the avionics.

- **Where does NFF have the highest impact?**

NFF is a serious problem that commonly affects aerospace and defence systems.

currently, systems, particularly aerospace and defence systems, are becoming increasingly complex because of the increasing number of components/subcomponents, which increases the complexity of the inter-dependency structure between these components. Thus, faults and failures are becoming harder to detect and isolate, and consequently, the time that operators and maintenance technicians must spend on fault detection rises in direct relation to these complexities. A further serious problem that has recently attracted much attention is the health diagnostics of these complex systems consisting of many electronic components, where IFs have become troublesome or almost impossible to find, which as a consequence will increase the NFF rate in those systems.

Hence, with the increasing complexity of real-world avionics in aerospace and defence systems, the need to improve the reliability, maintainability and safety of these systems also increased.

- **How much does NFF affect the aerospace and defence industries?**

NFF costs the affected industries a great deal of money. However, it is exceptionally difficult to quantify the exact cost related to NFF in the aerospace and defence sectors.

Based on available statistics drawn from surveys (Huby, 2012), avionics creates 75% of aircraft NFF incidents and that the avionics NFF proportion is 40% or higher.

Moreover, according to the industry-wide survey conducted in 2012 by Copernicus Technology, 66% of participants could not assess the cost of NFF in their industry, and 34% were unaware of the effect of NFF cost on their serviceability (Figure 1.1).

However, NFF has the most significant impact on people because it can restrict operational results and therefore cause extra pressure for the people in charge. When an operation is abandoned unfinished because of an IF, the amount of pressure on the aircrew is equal to the amount of pressure felt by the maintenance technicians whose efforts in repairing a fault ended up with NFF.

- **How is it possible to decrease the NFF rate in this research?**

In response to the above challenges, the main subject of this thesis is the development, application and testing of novel Intermittent Fault Detection and Prediction (IFDP) techniques for nonlinear complex systems. The IFDP tool should reliably detect IFs when they occur (see the first block of Figure (1.2)) and predict future IFs, regardless of any uncertainties in the system (see the second block of Figure (1.2)).

This chapter sets the context for this research work.

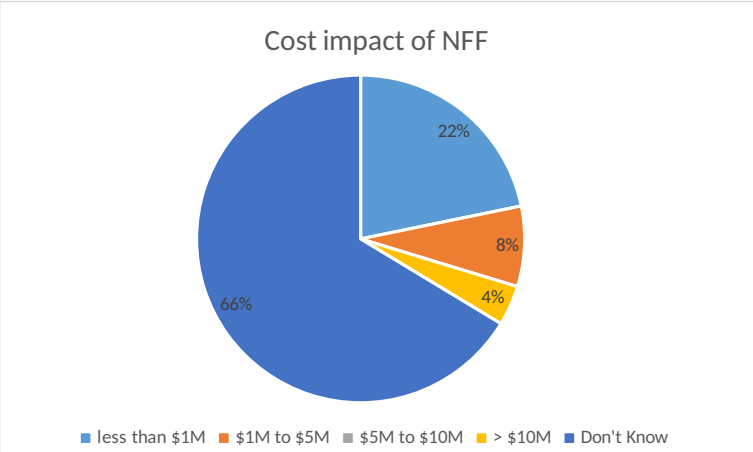

Figure 1.1: Cost impact of NFF according to the industry-wide survey conducted in 2012 by Copernicus Technology (Huby, 2012).According to this survey, 66% of participants could not assess the cost of NFF in their industry, 22% assess that the cost of NFF in their industry is less than 1M$ and 12% assess the NFF cost between 1M$ to 10M$.

1.1.1 The need for IF detection

Currently, the increasing number of components and subcomponents in aerospace and defence systems increases the complexity of the inter-dependency structure between these components. Thus, IFs are becoming harder to detect and isolate.

For years, in response to these demands, several methods have been introduced for detecting possible faults in complex systems to guarantee the normal functionality of the system (Wünnenberg, 1990). These conventional methods are useful in the detection of well-known faults such as hard faults in a single component system or systems with a limited number of building components.

In general, fault detection methods have several overlapping taxonomies. Some are more oriented toward a control engineering approach and others toward mathematical, statistical and Artificial Intelligence (AI) approaches. In practice, the designer selects one of several fault detection methods, based on the specifications of the system and the nature of possible faults.

The well-known fault detection methods are categorized as follows:

(a) **Physical redundancy methods:** A traditional approach to fault detection in the wider application context is based on hardware or physical redundancy methods that use multiple sensors, actuators or components to measure and control a particular variable (Ding, 2005).

(b) **Model-based methods:** In view of the conflict between reliability and the cost of adding more hardware, it is possible to use dissimilar measured values together to cross-compare them, rather than replicating each hardware component individually. Hence, model (observer)-based fault detection techniques were introduced (Ding, 2005).

Although among all the methods for online fault detection, one of the particularly interesting techniques is the model (observer)-based fault detection approach,

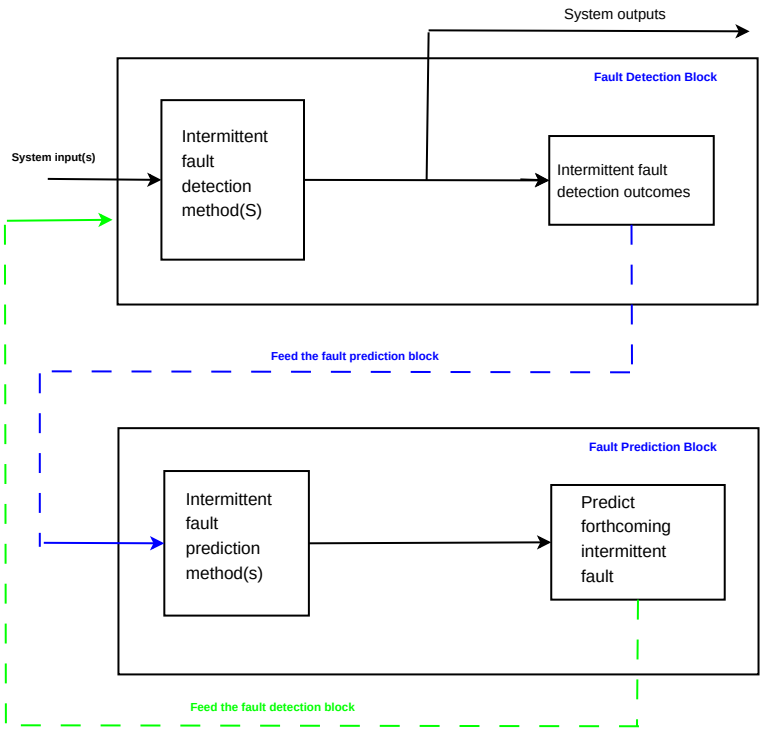

Figure 1.2: Thesis roadway

whose effectiveness has been shown in detecting sensor, actuator and system component faults (Chen, 2011), it has a few drawbacks. The most serious one is that, in the case of noise, input variations and change in the operating point of the monitored process, false alarms are possible.

(c) **Experience-based methods:**

Systems in industrial plants are becoming increasingly complex, so modelling them has become extremely time-consuming and costly and may involve numerous measuring errors. Hence, for better detection, substantial domain-specific knowledge is required.

However, for new or extremely complex systems, monitoring experience/knowledge is becoming increasingly difficult to obtain by ordinary field operators (Shen & Chouchoulas, 2000). Thus, there is a pressing need for computer-based solutions or data-driven methods.

(d) **Data driven-based or data-driven methods:**

Data-driven fault diagnostic techniques rely on historical data about previous operations performed for specific fault symptoms. This method uses captured information to solve new problems (Jiang *et al.*, 2014; Naik, 2010):

(i) Deterministic data-driven methods:

One important characteristic that a deterministic data-driven reasoning system relies on is the experience obtained. Experience should be complete and

adequate to cover all possible failure statuses. This can be crucial for new systems. Furthermore, for systems that need safety, such as nuclear plants or aircraft, a dearth of knowledge regarding some extraordinary events because of their rare occurrence may be an issue. Hence, deterministic data-driven reasoning can be beneficial when the understanding of the system is poor but knowledge of previous cases and actions taken is acceptable (Magoun, 2000).

(ii) Probabilistic data-driven methods:

However, to detect faults in complex systems in the presence of uncertainty, probabilistic data-driven fault detection methods are more appropriate. These techniques are particularly useful for diverse problems of varying size and complexity, where uncertainties are inherent in the system (Deng, 2002).

1.1.2 Intermittent fault prediction

As explained earlier, the increasingly complex and expensive systems built for use in intense environments have resulted in new challenges in maintenance, planning, decision-making and monitoring that make fault prediction a challenging task (Lorton *et al.*, 2013; **?**).

Fault prediction methods can be classified as follows:

(a) Model-based (Wende *et al.*, 2014)

(b) Statistical (Suresh *et al.*, 2014)

(c) Deterministic data-driven (Williams & Rasmussen, 2006)

(d) Probabilsitic data-driven (Williams & Rasmussen, 2006)

(e) Hybrid approaches (combinations of multiple approaches, i.e., a combination of data-driven and model-based approaches) (Li *et al.*, 2013; bar, n.d.)

Some recent studies have attempted to use probabilistic data-driven methods to provide a unified framework for fault prediction in complex systems. Dealing with uncertainty in prognosis and balancing decision-making strategy are major benefits of the chosen methods (Khodakarami, 2009; Deng, 2002).

Most probabilistic data-driven fault prediction methods can integrate information from different sources, including experimental data, historical data, and prior expert opinion. This is particularly useful for the reliability assessment of fault-tolerant systems, where failures in test or field operations are prohibitively expensive, or even impossible (such as IFs).

1.2 Research aim and objectives

1.2.1 Principal aim

In response to the above challenges, the main aim studied in this thesis is as follows

- **To achieve a robust intermittent fault detection and prediction approach for nonlinear complex systems.**

The work described in this research aims to first develop a novel model-based IF detection method in the presence of unknown inputs (uncertainties). Then, appropriate model-free detection and prediction methods are employed to detect, isolate and predict IFs when uncertainty is inherent in the system.

1.2.2 Objectives

To achieve the principal aim, a series of objectives must be attained:

1) Identify, test and validate novel model-based methods that can detect faults in the presence of uncertainties, with a special focus on intermittency.

2) Employ, compare and validate model-free approaches to detect and isolate faults in the presence of uncertainties, with a special focus on intermittency.

3) Propose, test and validate model-free approaches to predict faults in the presence of uncertainties, with a special focus on intermittency.

1.3 Thesis structure

The thesis chapters are organized as set out in the list below and shown in Figure (1.3):

- Chapter 2 provides a literature review identifying other important or relevant publications and background in fault detection and isolation methods.

- Chapter 3 describes the methodology and solutions used in this research.

- Chapter 4 describes a fully functioning experimental fuel rig, the main case study used for the validation and evaluation of methods and algorithms.

- Chapter 5 studies the novel model-based IF detection methods. This chapter focuses mainly on the novel development of the model-based fault detection techniques for a wider range of continuous nonlinear systems. The designed methods can be used for a wider range of nonlinear systems with fewer limitations.

- Chapter 6 introduces the application of the designed model-based IF detection filter for detecting the IFs of an experimental fuel rig system with unknown inputs. To do that, first, the mathematical model of the experimental fuel rig is presented. Next, an appropriate model-based fault detection filter is designed for the system. Finally, the effectiveness of the designed fault detection filter is demonstrated with the presented simulation results.

- Chapter 7 investigates how the specific form of probabilistic data-driven methods, Bayesian Networks (BN)s, known as a Hybrid Dynamic Bayesian Network (HDBN) can be applied to detect IFs in the experimental fuel rig system under consideration. Initially, the knowledge of the model-based fault detection method from Chapter 6 is used to construct the static BN. Next, a static BN is used to detect and isolate IFs in the experimental fuel rig system. Then, the posterior probabilities of the component failures give an indication of which components have caused the symptoms observed (isolation). After that, the HDBN is designed for the same system to detect and isolate IFs. HDBN can handle the temporal dynamics and uncertainties in the system. The performance of the two networks is then compared using sensitivity analysis.
- Chapter 8 introduces the Bayesian-based prediction method known as Gaussian Process Regression (GPR) for IF prediction. The GPR model is an example of a flexible, probabilistic, non-parametric model for prediction with uncertainty quantification. It offers a range of advantages for modelling from data and has therefore been used for dynamic systems and time-series modelling. This chapter also deals with the issue of covariance function selection and presents a two-step process to detect, isolate and predict IFs in the experimental fuel rig system. The proposed method is then verified by using a simulation case study.
- Chapter 9 contains a summary, discussions and conclusions on the overall thesis results along with ideas for possible further work.
- Appendix A: This appendix covers definition of Lipschitz nonlinearity.
- Appendix B: This appendix covers the Bayesian probability theory terms and definitions.
- Appendix C: This appendix covers the Gaussian and uniform distributions.
- Appendix D: This appendix covers Beta distribution.
- Appendix E: This appendix covers the limitation of the model (observer)-based fault detection techniques as system complexity increases.

1.4 List of Publications

- Papers from Chapter 5:
  - **Sedighi, T.**, Foote, P. D., Sydor, P. (2017). Feed-forward Observer-based Intermittent Fault Detection, *CIRP Journal of Manufacturing Science and Technology*, **17**, 10-17.
  - **Sedighi, T.**, Foote P. D., and Khan, S. (2014). The Performance of Observer-based Residuals for Detecting Intermittent Faults: the Limitations, *Procedia CIRP*, **22**, 65–70.
  - **Sedighi, T.**, Phillips P. and Foote P. D. (2013). Model-based Intermittent Fault Detection, *Procedia CIRP*,**11**, 68 – 73.

- Paper from Chapter 6:
  - **Sedighi, T.**, Foote P. D., and Khan, S. (2015). Intermittent Fault Detection on an Experimental Aircraft Fuel Rig: Reduce the NFF Rate, *Proceeding of the International Conference on System and Control, ICSC'2015*
- Papers from Chapter 7:
  - Naghshband, S. N., Varga, L., Purvis, A., McWilliam, R., Minisci, E., Vasile, M., Troffaes, M., **Sedighi, T.**, Weisi, G., Manley, E., Jones, D. H. (2020). A Review of Methods to Study Resilience of Complex Engineering and Engineered systems, *IEEE Access*.
  - **Sedighi, T.** (2019). Using Dynamic and Hybrid Bayesian Network for Decision Making, *International Journal of Strategic Engineering (IJOSE)*, **2**, 13.
  - Daneshkhah, A., Hosseinian-Far, A., Chatrabgoun, O., **Sedighi, T.**, and Farsi, M. (2018) . Probabilistic Modeling of Financial Uncertainties, *International Journal of Organizational and Collective Intelligence*, **8**, 1-11.
  - Daneshkhah, A., Hosseinian-Far, A., **Sedighi, T.**, and Farsi, M. (2017). Prior Elicitation and Evaluation of Imprecise Judgements for Bayesian Analysis of System Reliability, *Strategic Systems Engineering for Cloud Computing and Big Data Analytic*, 63-79.
  - Farsi, M., Hosseinian-Far, A., Daneshkhah, A., and **Sedighi, T.** (2017). Mathematical and Computational Modelling Frameworks for Integrated Sustainability Assessment, *Strategic Systems Engineering for Cloud Computing and Big Data Analytic*, 3-27.

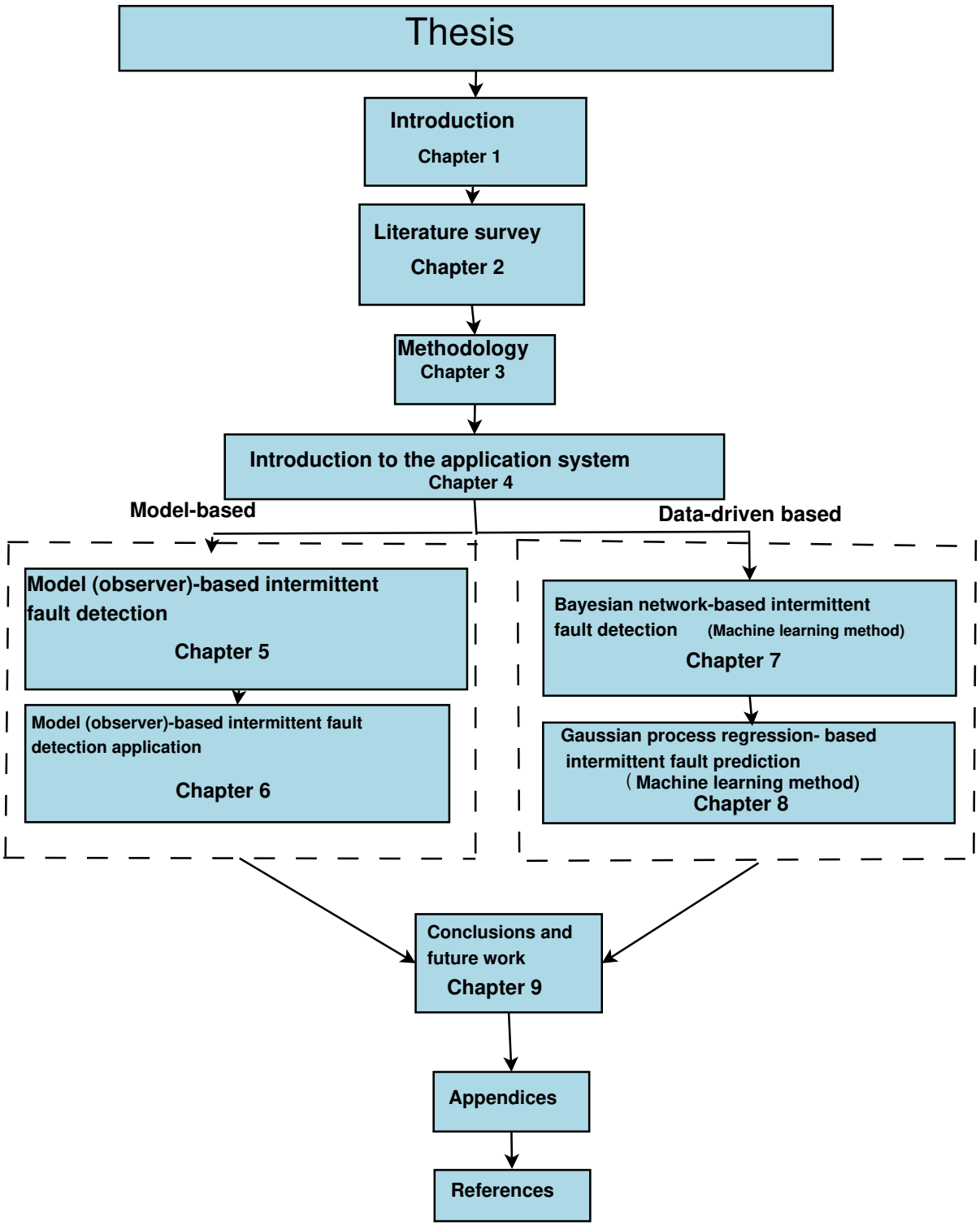

Figure 1.3: Thesis Structure.

Chapter 2

Literature Survey

2.1 Introduction

It is essential in the early stage of this thesis to explain and identify available techniques for the main purpose of this research, the intermittent fault detection, isolation and prediction to reduce the NFF rate, to substantiate the appropriateness of the chosen methods.

Hence, in this chapter, a broad overview of the current methodologies and techniques in the field of fault detection and prediction is presented instead of an in-depth study of a specific method.

This chapter is organized as follows: an introduction to the machine learning and artificial intelligent techniques in fault detection and prediction are given in 2.2 Intermittent fault is briefly introduced in Section 2.3. . Then the current fault detection and prediction methods are then addressed in Sections 2.4- 2.5 respectively while the discussions and conclusions are given in Section 2.6.

2.2 Machine learning and artificial intelligent

Machine learning (ML) tools are very important methods for fault detection and prediction in complex systems with huge amount of information and data because these methods are able to process many data and are engine of decision-making support for fault diagnosis. The most important machine learning techniques between the available methods are Support Vector Machine (SVM), Artificial Neural Network (ANN), Decision Trees (DT), Bayesian Network (BN) (Lo *et al.*, 2019).

2.2.1 Support vector machine

Support vector machine is a new computational learning technique on the statistical theory. Because of it's great generalization capability, SVM becomes more attractive than other traditional techniques such as neural network in machine learning. In the fault

diagnosis problem, this method is used for understanding and classification of specific patterns from the signal, based on the fault occurrence in the system (Widodo & Yang, 2007).

2.2.2 Artificial neural network

Artificial neural networks are inspired from biological neural networks of the human brains (Haykin & Network, 2004). In the modelling of artificial neural networks, knowledge are introduced as numeric values which are known as weights. Weights are then used to build the relationships between the input and the output parameters of the system under investigation. The most frequently used ANN is the Multilayer Perception (MLP) method. In this method which is a typical back propagation ANN, there are an input layer, hidden layer and an output layer. The MLP is normally trained by the back propagation of errors between targeted outputs and the obtained outputs from the ANN by using gradient descent or conjugate gradient algorithms (Mekki *et al.*, 2016). These errors are then used to detect the faults in the physical system.

2.2.3 Decision trees

A Decision Tree (DT) is a diagram or chart which is used to explain a set of actions with an statistical probability and is very papular in the online fault detection and isolation applications and is a rule based method. Each branch of the decision tree performs a possible decision, outcome, or reaction. The farthest branches on the tree shows the final outcomes (Samantaray, 2009).

2.2.4 Bayesian network

BN also is known as probability network or Bayesian Belief Network (BBN) is the tool to estimate certainties of events that are unobservable or costly to observe where evidence/information are given (Lou *et al.*, 2020). BN is explained in more details in section (2.4.4).

2.3 Intermittent fault

Fault in a system is an external input that causes a deviation from the normal behaviour of the system (Baghernezhad & Khorasani, 2016). Faults can occur in the actuators, process components or the sensors as shown in Figure (2.3), and are categorized accordingly (Witczak, 2003). Each of these faults and their effects is briefly described below.

Component faults

Component faults appear in the components of the system. These faults change the physical parameters of the system which results in a change of its dynamical properties

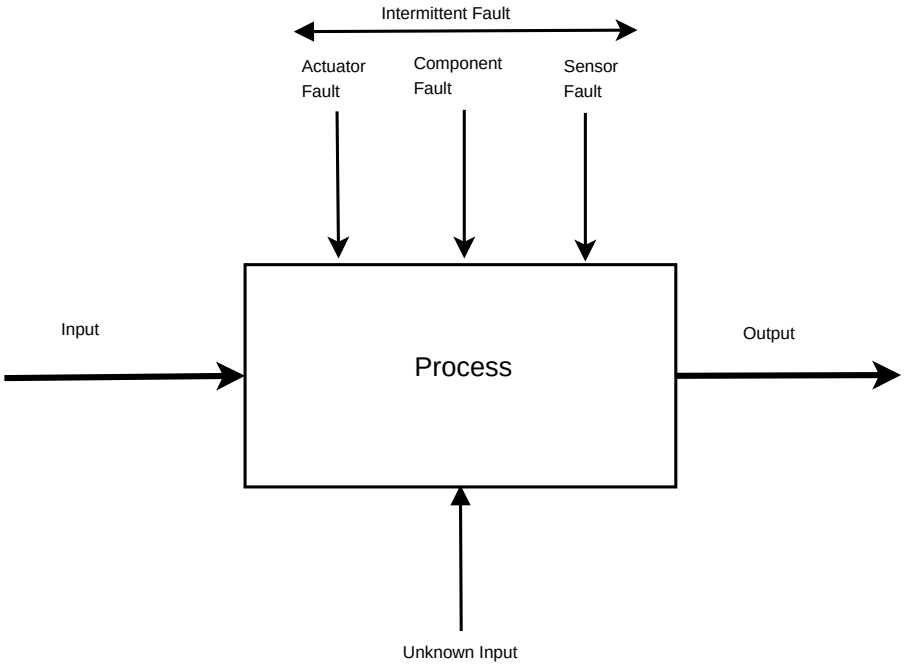

Figure 2.1: Fault types.

consecutively. Wear and tear, ageing of components, etc. are usually the ordinary reasons for component faults. Leakages in tanks, breakages or cracks in gearbox system, change in friction due to lubricant deterioration are some example of these faults. Since component faults may result in irregularity and instability of the operation, it is very valuable to detect these faults (Li & Xu, 2010; Armaou & Demetriou, 2008; Moraes *et al.*, 2006).

Actuator faults

To transform control signals into suitable actuation signals such as torques and forces to drive the system, actuators are needed and a fault in an actuator may cause greater energy usage. Hence, it is very important to detect actuator faults. Stuck-up of control valves, faults in pumps, motors etc. are examples of these faults (Li & Yang, 2016; Armaou & Demetriou, 2008; Theilliol *et al.*, 1998).

Sensor faults

In a system which is a closed loop, the measurements are usually obtained by sensors. These measurements are then used to set up the control inputs. Therefore, any fault in sensors may cause degradation in the performance (loss of accuracy) of the system. Hence, it is very valuable to detect these faults. Bias, drift, loss of accuracy in the performance, sensor freezing and calibration error are some examples of these faults (Wang & Song, 2014; Sharma *et al.*, 2010; Wang & Xiao, 2004).

Faults can also be categorized based on how these are modelled into:

- Additive faults: where the faults are defined as fictive inputs which act on the system the same as any unknown external unknown inputs (disturbances) (Zhang

& Basseville, 2014; Yang & Stoustrup, 2000).

- Multiplicative faults: often describe systems with parametric faults (i.e., abnormal variations of some model parameters)(Rotondo *et al.*, 2016; Ferdowsi & Jagannathan, 2013; Abdelghani & Friswell, 2007)

Actuator and sensor faults are more easily modelled as additive faults, whereas component faults are modelled as multiplicative faults (Boulkroune *et al.*, 2011; Ferdowsi & Jagannathan, 2011).

Furthermore, faults are also classified based on whether they have grown slowly during the operation of a system (incipient fault) (Parlangeli *et al.*, 2007); arisen suddenly like a step change as a result of a sudden breakage (abrupt faults) (Ukil & Zivanovic, 2015); or develop in discrete intervals assigning to component degradation or unknown system interactions (intermittent faults) (De Kleer, 2009).

Incipient faults develop gradually and causing degradation of equipment. Therefore, their gently changing performance makes it hard to detect them (Parlangeli *et al.*, 2007).

Abrupt faults have more serious influence and may cause damage to equipment. However, these faults are straightforward to detect (Ukil & Zivanovic, 2015).

Intermittent fault

An intermittent fault is defined as one which persists in a system for a limited time, and then a system recovers from the intermittent fault and can perform as before without going through any maintenance actions. Intermittent faults are difficult to diagnose and usually are repeating (De Kleer, 2009). Hence, intermittent fault puts an ever-increasing challenge in the maintenance of electronic, mechanical and hydraulic equipment (Ahmad, 2017). An electronic circuit with loose solder joints subjected to a vibration that causes intermittent fault (open circuit) is an example of these faults (Cai *et al.*, 2016; Kim, 2014; Incarbone *et al.*, 2014).

When the persistent fault appears in a system it will not disappear and may cause a major breakdown to the systems operation. Also, while the transient fault appears in the system it will be disappear shortly and it is very less likely to appear again. However, unlike these types of faults, the intermittent faults will appear as a set of discrete intervals in the system and if left unattached it may get longer in duration and larger in amplitude (Khan *et al.*, 2014a).

Based on these definitions, intermittent fault can be modelled as a combination of impulses at discrete intervals:

$$f_i(t) = \begin{cases} dd_1 & for \quad t_0 \leq t_1 \\ dd_2 & for \quad t_1 \leq t < t_2 \\ dd_3 & for \quad t_2 \leq t < t_3 \\ \vdots & \\ dd_{n-1} & for \quad t_{n-2} \leq t < t_{n-1} \\ dd_n & for \quad t_{n-1} \leq t < t_n \end{cases} \tag{2.1}$$

where $dd_i = A_i - A_{i-1}$ for $(i = 1, \cdots, n)$, are constants and t indicates the time (see Figure 2.2).

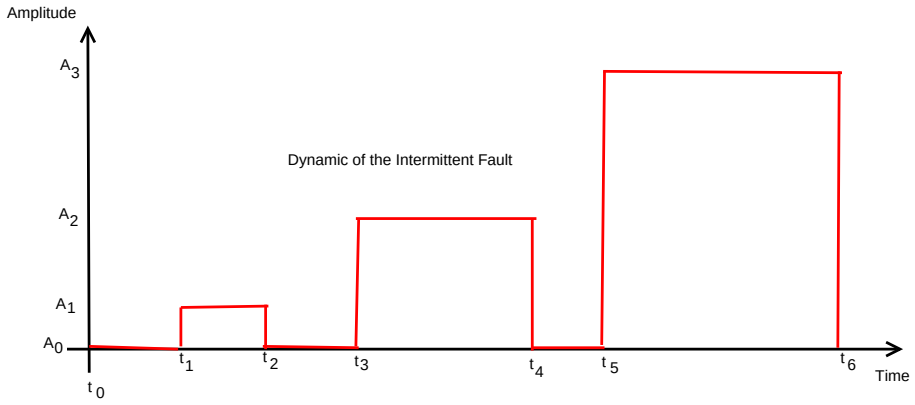

Figure 2.2: Intermittent fault dynamic.

In the real world, the intervals between the intermittent faults, the duration of each interval and the amplitudes are unknown like any other faults.

However, the intermittent fault typically tends to worsen with time, until it eventually becomes substantial enough that it can be detected with conventional test equipment's. Hence, developing the capability for early detection and isolation of the intermittent fault is very valuable (Zhao *et al.*, 2009).

2.4 Fault detection methods

For years, to assurance normal functionality of the system, different methods have been introduced for detecting possible issues in dynamic systems. The designer selects one out of several Fault Detection (FD) methods, based on the specifications of the system and the nature of possible faults (Basseville *et al.*, 2000). Some methods are more suitable for off-line FD test. One example is introduced in (Liberatore *et al.*, 2006; Chen & Patton, 1999) in which the off-line FD method is used for health monitoring of mechanical structures, such as bridges. Other methods aim at detecting faults on-line (Yan *et al.*, 2017).
Fault detection methods have several overlapping taxonomies. Some are more oriented toward control engineering approach, other to mathematical, statistical and AI approach.

Interesting divisions are described in (Miljkovic, 2011) and the references therein as follows Figure (2.3):

- Model-based methods
- Experience-based methods
- Data driven-based or data-driven methods

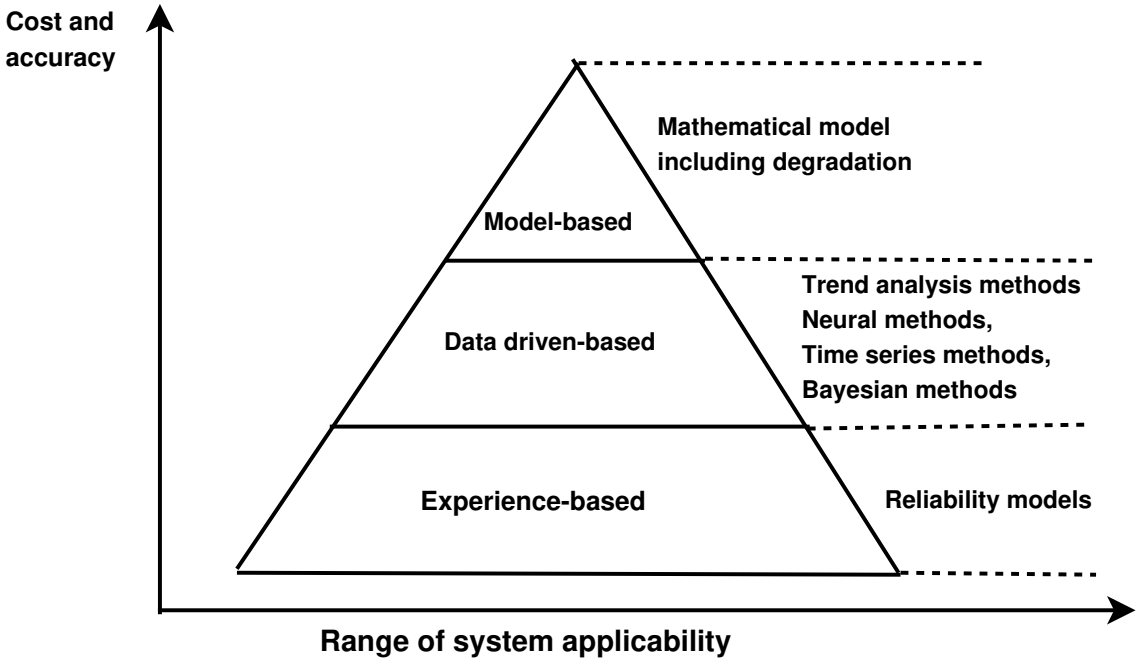

Figure 2.3: Diagnostic and prognostic main approach.

2.4.1 Model-based methods

Among all the methods for online fault diagnosis, one of the particular interesting techniques is the model-based FD approach. This method is very effective in detecting sensor, actuator and component faults in the system (Chen & Patton, 2012; Lee *et al.*, 2003; Venkatasubramanian *et al.*, 2003).
The considerable benefit of the model-based FD approaches is that additional hardware components are not needed to recognize the FD algorithm and a model-based FD algorithm can be achieved via software on a process control computer. In many cases, the measurements which are needed to control the action are also adequate for the FD algorithm, hence, additional sensors are not needed to be installed. Analytical redundancy makes use of a mathematical model of the system under investigation and it is therefore often referred to as model-based approach for fault diagnosis (Basseville, 1997).
Moreover, the model-based FD techniques, make use of analytical models for a residual generation. These can broadly be classified into (Basseville, 1997):

- Parity space FD
- Parameter identification based FD
- Observer-based FD.

Parity space approach

In this approach, a set of accurately modified system equations in terms of the transfer function (also called parity relations) is derived according to the measured signals from

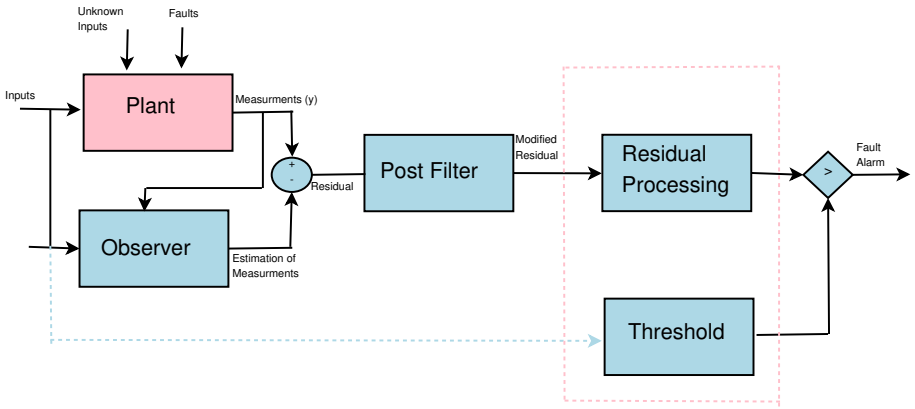

Figure 2.4: Summarized diagram for the model (observer)-based fault detection scheme.

the process. These parity relations decouple the residuals from the system states and also from each other. Then the inconsistency in the parity relations indicates the presence of fault (Ye *et al.*, 2004).

Parameter identification-based approach

In this approach, fault detection is performed according to the on-line parameter estimation. The information of the fault can be obtained by comparing the estimated parameter with the nominal process parameter. Any difference between the two indicates fault (Wang *et al.*, 2010).

Observer-based approach

The observer-based technique is one of the mostly applied model-based schemes for detecting the faults in a system (Abid, 2010). Observer-based methods rely on the ability to produce an exact model of the system without influence from unknown inputs (external disturbances). The difference between the outputs of the model and the real world system, often called the observer-based residual, which is used to assess the condition of the real world system as shown in Figure (2.4).

Unknown Input Observers (UIO)s (Zhang *et al.*, 2015), Luemberger/high gain observers (Prasov & Khalil, 2012), feed-forward observers (Ruderman, 2013), sliding mode observers (Ríos *et al.*, 2017) and Kalman filters (Barrau & Bonnabel, 2016) are a few of the many types of observers, also known as state estimators, and can be used as a residual generation function for detection and isolation of faults (Garcia-Alvarez *et al.*, 2011; Li & Jaimoukha, 2009).

When the system is performing exactly as modelled then the observer-based residual is equal to zero and any significant deviation from this will be due to a fault in the system. However, in practice, this will not be the case due to the noise, disturbances, and modelling errors which are accruing over time and the cumulative effect of modelling errors can lead to a large difference between the model and system.
Hence, the designed observer-based FD filter should be robust against disturbances and

measurement noises which led the researchers to the area of robust Fault Detection and Isolation (FDI) algorithms. First, Frank (Schneider & Frank, 1996) used the robust observer-based fault detection schemes for instrument faults. Later robust Unknown Input Observers (UIO) were introduced in (Guang-Ren & Patton, 2001) and then considerable contribution was made.

Note that in model-based FD, unknown inputs or disturbances are types of uncertainty to the system.

2.4.2 Unknown input observers:

Finding systematic design method for systems subjected to unknown disturbances and model uncertainties has been proven to be difficult, (Chen & Patton, 1999). In (Xuhui & Ziyi, 2012), the Nonlinear Unknown Input Observer (NUIO) is established by radial basis function neural network to guarantee the stability of model free adaptive control with the bounded disturbances where (Martinez-Gardea *et al.*, 2015) presents a nonlinear observer design for both linear and nonlinear systems subject to disturbances for a laboratory antilock braking system but not experiencing faults.
The main idea of UIO/NUIO was to make the residual signal decoupled of the external unknown inputs (Zhang *et al.*, 2018b; Yan *et al.*, 2016).

It was later observed that the existence condition for the unknown input observer are very hard, which were then relaxed by introducing the so-called matrix principle approach. In this approach, attempts were made to make the residual signal insensitive to unknown inputs instead of decoupling of estimated states from the unknown inputs (Ahmadizadeh *et al.*, 2013).

Some recent results aiming at this goal for linear time invariant systems are reported. See (Guo & Xu, 2015; Qning *et al.*, 2014; Liu & Yang, 2005; Ding & Ding, 2000) and the references therein.

Aldeen & Sharma (2008) by decoupling the faults and unknown disturbances through some state and output transformation, develop an observer which can be used to estimate the unknown input (disturbance) state and fault signals simultaneously.

Later, (Sharma & Aldeen, 2011) present a fault and unknown input reconstruction technique based on utilization of a network of two interconnected sliding-mode observers.

By using the inherent features of sliding-mode observers, (Yan & Edwards, 2008) propose a fault reconstruction scheme which can be implemented online for a class of nonlinear systems with uncertain parameters.

Moreover, (Castillo *et al.*, 2012) propose a robust FDI system by using a state estimator which is capable of dealing with both bounded uncertainties and parameters while (Bejarano *et al.*, 2011) address the state observation problem for a class of switched linear systems with unknown inputs.

Furthermore, (Gao & Lin, 2012) investigate the problem of active fault tolerant control for a reusable launch vehicle with an actuator fault using both adaptive and sliding-mode techniques and (Zhu & Yang, 2013) developed a new fault reconstruction method based on a high-order high-gain sliding-mode observer to achieve the FDI objective.

In (Eddine & Belkhiat, 2015) the FDI problem has been solved by minimization of the norm of the H_{inf} and maximization of the H_{index} for a class of switched linear systems

subject to sensor faults and bounded unknown inputs with selecting a suitable trade off between the robustness to disturbances and the robustness to sensor faults.
Moreover, for observer-based FDI, many design methods have been developed under the assumption that the observer matching condition is satisfied, (Chen & Chowdhury, 2010; Zhang *et al.*, 2010; Alwi *et al.*, 2009; Yan & Edwards, 2008; Edwards & Tan, 2006).

For instance, (Alwi *et al.*, 2009) propose a design method of sliding-mode observers for sensor fault reconstruction under the assumption that the observer matching condition is satisfied.

Then, for linear parameter varying systems, (Alwi & Marcos, 2012) present observer schemes for actuator sensor fault reconstruction by using a virtual system comprising the system matrix and a fixed input disturbance matrix while (Said *et al.*, 2013) presents structural information-based residual generation for industrial application.

However, in systems subject to faults and disturbances, since both disturbances and faults contribute to the residual generated by the FD observer, some small faults cannot be detected by a pre-designed threshold (Andrade *et al.*, 2016).

A perfect or ideal FD observer should minimize the maximal undetectable fault size in the worst case as its goal. However, this criterion is not adopted for FD observer design directly. The philosophy behind this criterion is that a FD observer should be robust to disturbances but sensitive to faults, (Liu & Yang, 2005; Ding & Ding, 2000). They demonstrate how the matching rank method can be used to select the most appropriate matching that leads to residual computational sequences.

Moreover, by transforming the system into two subsystems, (Chen & Chowdhury, 2010) achieve the purpose of early detection of incipient faults based on sliding-mode observers .

Recently, the issue of the observer matching condition was discussed in some existing literature. For example, (Raoufi & Zinober, 2010) present a scheme to design robust sliding-mode observers to satisfy the observer matching condition by using linear matrix inequality optimization.

Later (Ng *et al.*, 2012) present a disturbance decoupled fault reconstruction scheme using cascaded sliding-mode observers when the observer matching condition is not satisfied.

(Liu & Zhang, 2014) also deals with the estimation of states for a class of linear systems while the bounded uncertainty exists in the both states and outputs.

Furthermore, all the discussions in, (Alwi & Marcos, 2012; Chen & Chowdhury, 2010; Alwi *et al.*, 2009), are carried out under the assumption that the observer matching condition is satisfied.

Observer-based residual

.

The FD consists essentially of two steps (Wan *et al.*, 2016; Wang *et al.*, 2009),

- Residual generation
- Residual evaluation

The purpose of the first step is to generate a signal, the residual, which is supposed to be nonzero in the presence of a fault and zero otherwise. However, the residual is almost always nonzero due to unknown inputs (disturbances) and model perturbations, even if there is no fault (Li & Jaimoukha, 2009; Sedighi *et al.*, 2013).
The purpose of the second step of the FD algorithm is thus to evaluate the residual and draw conclusions on the presence of a fault. This is done by comparing some function of the residual to a threshold (Wan *et al.*, 2016; Khan *et al.*, 2008).

Threshold

The threshold is obtained based on the residual dynamics in the fault-free case. Hence the value of threshold gives an explicit bound on the fault-free case and thus provides a valuable guideline for robust threshold selection (Wang & Wang, 2007; Al-Salami *et al.*, 2010; Puig *et al.*, 2012).

There are different methods to define threshold such as fixed threshold design (Baligar *et al.*, 2006) and adaptive threshold design (Abdo *et al.*, 2011a). The adaptive threshold is related to the main factors including system input, output, unknown inputs (disturbance) and parameters drifting over time (Zhang *et al.*, 2018b; Abdo *et al.*, 2011b; Al-Salami *et al.*, 2010).

2.4.3 Experience-based methods

In recent time there is a trend towards experience-based and artificial intelligence methods. It presents an expert system that uses a combination of object-oriented modelling, rules, and semantic networks to deal with the most common faults, such as bias, drift, scaling, and dropout, as well as system faults, (Shen & Chouchoulas, 2000; Jiang & Maskell, 2015).

In (Cecati, 2015) the current experience/knowledge-based techniques for the fault detection has been demonstrated. Although, these approaches are very valuable and useful when the experiences are complete, but in the case of intermittent fault detection cannot be accurate because the knowledge about IF characteristics usually are not complete.

One important characteristic this method relies on is the experience obtained. Experience should be complete and adequate to cover all possible status of failures. This can be crucial to obtain for new systems.

Furthermore, for systems that need safety, such as nuclear plants or aircraft, there can be a dearth of knowledge regarding some extraordinary events due to their rare occurrence.

Hence, data-driven based reasoning can be beneficial when the understanding of the system is poor but knowledge of previous cases and actions taken is acceptable (Magoun, 2000).

2.4.4 Data driven-based methods (model-free based methods)

One of the promising method for fault detection in complex systems which are introduced by linear and nonlinear models is the model-free or data-driven method. One of the characteristic of this method is that the parameters of the system under investigation may be unknown (lack of experience or information).

In general, data-driven fault detection systems rely on historical data known about previous operations taken for specific fault symptoms. So this method uses captured experience to solve new problems (Zhao *et al.*, 2018; Jiang *et al.*, 2014; Naik, 2010).

Different data-driven methods has been introduced such as Principal Component Analysis (PCA) (Abdi & Williams, 2010), partial least squares (PLS)(Hair Jr *et al.*, 2016; Barker & Rayens, 2003), neural network (NN) (Kalchbrenner *et al.*, 2014) and BN (Cheng & Greiner, 2001).

Among all of these methods, BN is one of the data-driven based reasoning methods which have been successfully used to assist problem solving in a wide range of disciplines including information technology, engineering, medicine, and more recently biology and ecology. BNs are particularly useful for diverse problems of varying size and complexity, where uncertainties are inherent in the system (Nannapaneni *et al.*, 2016; Marcot, 2012; Deng, 2002). Such uncertainties may occur from:

- The natural variability,
- Measurement error,
- Model error,
- Hypothesis testing error,
- Error in inference or
- Any optimization approximation,

The established theory of probability makes BN an appropriate method (engine) for reasoning and decision making under such uncertainty (Hosseini *et al.*, 2019).

(Rohmer, 2020) presents the uncertainties in structure and parameters of BNs. The author suggested that for decreasing the uncertainty representation and propagation in BNs alternative frameworks can be compare including approaches using probabilities, or interval-valued or Dempster–Shafer structures, possibility distributions, Fuzzy sets.

Bayesian network

BN is well established as a representation of relations among a set of random variables that are connected by direct edges and given Conditional Probability Table (CPT) at each variable (Friedman & Koller, 2003).

The most suitable method used to solve the network can depend on the size of the network, the type and amount of data available to populate the network, and the type of information is needed to obtain from the network.

The choice of the BN representation depends on the type of variables which are dealing with. In general, BNs may categorize based on their variables dependency on time as (Cheng & Greiner, 2001):

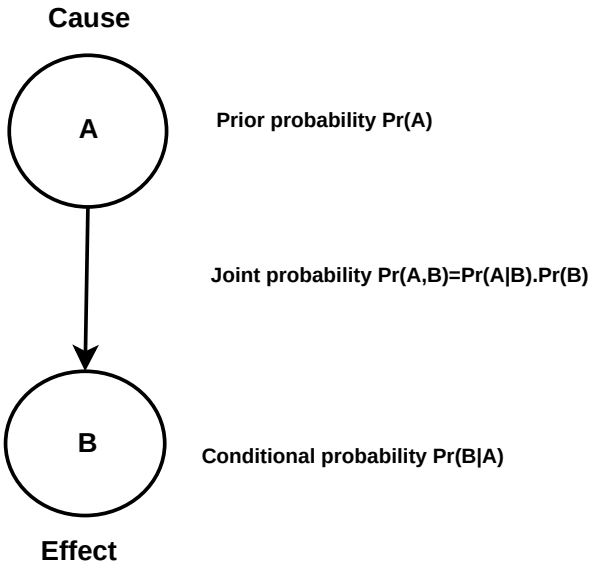

Figure 2.5: Bayesian network probabilities.

- Static BNs (SBN)s: all the variables are discrete.
- Dynamic BNs (DBN)s: all the variables are continuous (time-dependent).
- Hybrid DBNs (HDBN)s: variables are the combination of discrete and continuous.

However, the wrong set of variables can cause a smearing effect or give rise to an incorrect solution.

Causal relations in BNs are represented as a directed edge/arc/ link between variables (nodes), leading from the cause variable to the effect variable (Figure 2.5).

Nodes and links create the qualitative part of the network, i.e. its structure, while the quantitative part is represented by the probability associated with the variables.

To construct a BN two main tasks need to be done:

1) Qualitative task: find the Direct Acyclic Graphs (DAG)s with random variables known as nodes (La & Vuong, 2019; Sedighi & Varga, 2019),

2) Quantitative task: assign the prior probabilities to each random variable based on measurements, human experts, textbooks and any other reliable information (Sedighi, 2019a; Wan & Freitas, 2015).

Qualitative task: structure or graphs

BBNs or BNs are probabilistic graphical models represented as DAG, $G = (Y, L)$. These are used in many areas where reasoning under uncertainty is needed. The networks are constructed of nodes (Y), representing random variables X with n states, $(X_i = X_1, \cdots, X_n)$, of interest (i.e., the occurrence of an event or a component of a system), and links (L) joining the nodes, representing causal relations among the variables. The only constraint on the arcs allowed in a BN is that there must not be any directed cycles: it cannot return to a node simply by following directed arcs. Such networks are

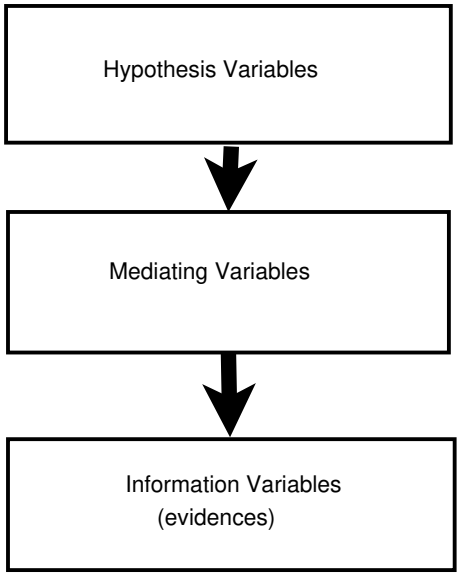

Figure 2.6: Bayesian network variables.

called directed acyclic graphs, or simply dags. The direct causal relation between two components shows that the corresponding components will have a great influence on the system rather than others (Cai *et al.*, 2018; Jensen & Nielsen, 2007).

There are several steps that a knowledge engineer must undertake when building a BN graph (bou, n.d.; Kwoh & Gillies, 1996) (Figure 2.6):

- Identify hypothesis variables/nodes
- Identify mediating variables (Mediation is introduced to have more refined network model of the domain)
- Identify information/evidence variables/nodes
- Identify the direct links between variables/edges
- Specify CPTs or Conditional Probability Distributions (CPD)s for each variable

The edges or casual represent a direct link between variables, leading from the parent (cause) variable to the child (effect) variable (Figure 2.5).

Quantitative task: assigning probabilities

Each node has a finite number of comprehensive and mutually exclusive states. Moreover, every node with a direct ancestor (parent) is combined with a CPT that hold the probability of each state of the node for any possible combination of the states of its parents (Cai *et al.*, 2018; Wang & Druzdzel, 2013).

For the nodes with no parents (root nodes), the CPT indicates the probability of being in each of the states of the associated variable (Wang & Druzdzel, 2013).

In general, BNs assigning prior probabilities to the variables based on the information, before any data is collected. In case of not enough information, all possible values of the parameters have equal prior probabilities (Nguyen-Trang & Vovan, 2017).

When the states of some of the variables in a network are known, it is possible to calculate the updated probability, of the remaining unknown variables, given the new evidence. Assessing this probability, known as posterior probability, is the main task of BN.

However, it is important to note that in the real-world their process is not so simple because assigning the prior probability is not straight-forward (Macci, 1996).

BN learning could be the structure learning (links), parameters learning or the combination of both.

Structural learning is much harder problem compared to the parameter learning since the number of candidate network grows super-exponentially when the number of variables increases (Chickering, 2013).
Note that to specify the graph structure it is needed to specify the size and type of the nodes. The nodes are either discrete where their size is the number of possible values each node can take on (binary) or continuous which may be presented as a vector and its size is the length of the vector.

The physical systems in the real world, commonly, comprises both continuous and discrete quantities, and to introduce the time variable in the framework of a probabilistic model there are two options:

- Disceretize the continuous variables into several discrete states. This approach can cause several problems including (Ghanmi *et al.*, 2011):
  - Increasing the number of defined intervals in discretization to achieve more accuracy which can cause a heavy cost of computational and complexity.
  - Characteristic of the continuous variable is fundamentally different from those of a discrete variable (loss of generality).
- Using DBN or HDBN:
  - DBNs/HDBNs extend BNs from static domains into dynamic domains which are more reasonable in describing the random changes of a dynamic process.

Dynamic Bayesian network

BNs that models sequences of variables are called DBNs which consists of a limited number of BNs, each of which corresponds to a particular time intervals. The connections between adjacent BNs represent how the state of domain evolve over time (Brandherm & Jameson, 2004).
For example, (Brandherm & Jameson, 2004) assumed that the system states can be represented by a set of variables, X_i, each of these variables can be either continuous or discrete. The continuous variables are partitioned into two subsets: one of the subsets is the variables that are measurements, i.e., their value is known to us; the remaining subset is unobserved variables.

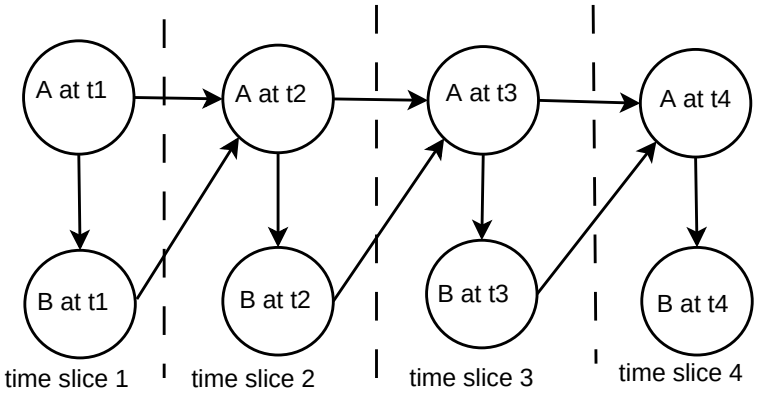

Figure 2.7: Dynamic Bayesian network representing a feedback loop.

Hence, the system is modelled as evolving in discrete time steps.

In general, there are two main assumptions for constructing DBNs (Murphy & Russell, 2002):

1) Network structure will not change over time, thus a DBN is constructed by an unrolling two-slice temporal BN (2TBN).

2) The DBN satisfies the Markov property (see Appendix B): given state at time t, the state of a variable at time $t+1$ is independent of the states prior to time t (see Appendix B, equation B.5).

In Figure (2.7), node A at $t1$ affects node B at current time slice $t1$, but is in turn affected by node B in the next time slice, $t2$. This presents a feedback loop, since A affects B and B affects A (at time $t2$)(Zhang *et al.*, 2006). In general the DBNs evolves (Iamsumang, 2015):

- Prior probability, $Pr(X_0)$, at $t = 0$.
- Transition probability distribution, $Pr(X_{i_t}|X_{i_{t-1}})$. The transition probabilities for any variable are determined completely by the value of the variables in the current, t, and previous, $t-1$, time steps, Monte Carlo Markov Chain (MCMC), (see Appendix B). This Markov assumption requires to model explicitly any variables, such as failures, that induce long term correlations on the system state.
- Observed probability likelihood, $Pr(X_{E_t}|X_t)$.

where t indicates the time.

To build a DBN it is necessary to insert a certain number of regression nodes representing the values of given variables at previous time instances where nodes are repeated at each time-slice. The number of regressions for each variable depends on the particular dynamic system considered (Iamsumang, 2015).

In the presence of discrete parents, the model is known as HDBN.

In HDBN if a given variable is continuous the node is also continuous and the probability distribution is usually Gaussian. If the variable is discrete then the node is also

a discrete and the probability is uniformly distributed, unless other information is available (Li *et al.*, 2012).

In case of failure presence, variables that represent failure nodes of the system are typically boolean, the two states representing the normal and fault behaviour respectively (Li *et al.*, 2012).

Bayesian network approaches for intermittent fault detection and isolation

In (Cai *et al.*, 2018) and Cai *et al.* (2017), an specific data-driven methods, BNs, has been reviewed for fault diagnosis. In their work they have presented the fault diagnosis modelling using BNs and reviewed different methods for each step. The general algorithms for fault detection presented in this paper are organized as: BN constructing including structure and parameter modelling, BN inference, fault identification, validation and verification.

Fault diagnosis in a qualitative sense is the reasoning of the cause-effect or fault-symptom relations and in almost all cases single symptom will be caused by several faults, while a single fault will exhibit several symptoms (Azarian, 2009).

Every fault and even symptom is modelled by a random variable in the network with a probability distribution.

In fault diagnosis application, variable, X, may be interpreted as the hypotheses of fault and evidence, X_E, is the observed symptoms.

When observed symptoms (evidence) are input to the network, probabilities of every fault are computed according to the Baye's rule equations in Appendix B, (B.3 and B.4). So, the ranking of different faults with the given symptoms and the eliminating of potential fault candidates is possible (Cobb & Shenoy, 2012; Mori & Mahalec, 2016).

In general, there are two methods for Bayesian Network Fault Detection (BNFD) (Cai *et al.*, 2016) as follow:

- Model-based BNFD: This method is an alternative approach to the model-based FD, where BN is adopted to diagnose the failure. The goal is to detect and localize faulty components in the system. Hence, the model should incorporate structural information about the system and mathematical modelling is such a representation. The performance of these methods will be greatly impaired when a poor model is used. In model-based BNFD the main structure of BN comes from the structure of the preliminary process model (Cai *et al.*, 2016).

- Data-driven BNFD: In this method, the structure of BN learns from the data. However, the accuracy of the learned BN is largely affected by the "richness" of the data and the prior knowledge of the network order. The crucial step in data-driven based BNFD process is to find the DAG (Yamaguchi *et al.*, 2012).

A fault variable in belief state takes into consideration all the evidence available up to the present time to determine a probability distribution of the considered fault. However, in a simplistic network, the fault node influences all measurements.

The BNFD can be applied to localise multiple faulty components that are correlated to exert a single symptom since it is the strength of BN (Gasse *et al.*, 2014; Deng, 2002).

2.5 Fault prediction methods

The increase of complex and expensive systems for use in intense environments has resulted in new challenges in maintenance, planning, decision-making and monitoring which results in increasing the uncertainties in the system fault prediction techniques.

Moreover, online system measurements may occur on various time scales from each other or only be available in specific system configurations (Lorton *et al.*, 2013; Ebden *et al.*, 2008; Williams & Rasmussen, 2006).

Hence, it is important that a prognosis methodology considering the uncertainty coming from many sources such as variability, information uncertainty, and model uncertainty.

In most cases, prognosis about the future is based on the diagnosis of the current state; therefore it should account for uncertainty in diagnosis (i.e. uncertainty in damage probability, detection, isolation, and quantification). (González *et al.*, 2013; Ebden *et al.*, 2008; Williams & Rasmussen, 2006). It must be known that a key distinction between a system model capable of diagnosis and one capable of prognosis is that a prognostic model can estimate the evolution of damage in the future while a diagnosis model only needs the ability to infer the current state of damage. Diagnostic procedures based on fault signatures or pattern recognition are examples of this. While they may be able to detect and isolate damage, they do not necessarily have any ability to model progressive damage mechanisms such as crack growth, wear, and corrosion (Lefebvre, 2014; Ebden *et al.*, 2008; Williams & Rasmussen, 2006).

Hence, on-line fault prognosis of a system is a very essential part of modern control and supervision system (Lin & Li, 2006), and one of the challenges of prognosis is developing accurate and comprehensive physics of failure models (Chookah *et al.*, 2011). These damage mechanisms are complex, varying with system design and dynamics, and can interact in many ways.

There are different prognosis techniques which can be classified as:

- Statistical methods include statistical process control (Vlasselaer & Meert, 2012), logistic regression (Pampel, 2000), survival models (Martinussen & Scheike, 2006), and stochastic process models (Klimenko *et al.*, 2015).,

- Model-based methods using system models to estimate Remaining Useful Life (RUL) or other relevant metrics (Ebden *et al.*, 2008; Williams & Rasmussen, 2006). Such methods rely on accurate physics-based models for prediction (Lorton *et al.*, 2013). These include physical failure models (Tinga, 2010), filtering models (Nguang & Lin, 1999), and statistical models (Jablonski, 1985).

- Data-driven methods or data driven-based methods consist of machine learning methods (support vector machines (Djeffal *et al.*, 2017), relevant vector machines (Joachims, 2005), Neural Networks (Wang & Huang, 2001)) and graphical models such as DBNs (Cai *et al.*, 2016), Hidden Markov Models (HMM)s (Elliott *et al.*, 2008) and Gaussian Process (GP) (Wang *et al.*, 2005).

- Hybrid methods , such as a combination of data-driven and model-based approaches (Skima *et al.*, 2016; Ebden *et al.*, 2008; Williams & Rasmussen, 2006)).

In the machine learning context, supervised learning is concerned with inferring the values of one or more outputs, or response variables, for a given set of inputs that have not yet been observed, or predictor variables (Lewis & Catlett, 1994).

These predictions are based on the training samples of previously solved cases depending on whether the output is continuous or discrete.

Traditional approaches to solving this kind of problem usually consist of parametric models, on which the behaviour of data is described by a previously defined model and the parameters of this model are learned from the training data (Vapnik, 1998). Moreover, many of the classical machine learning algorithms fit the following pattern while a training set of sample from some unknown distribution are given (Murphy, 2012),

- To find the single best-fit model for the data, solve a convex optimization problem, and
- To make the best-guess predictions for future test input points, use this estimated model.

where a convex optimization is a problem of minimizing a real-valued function defined on an interval (convex) over the convex set (Murphy, 2012).
By adjusting these parameters, it is possible to fit the model to the data. Once this is done, it should be straightforward to use the model and predict the output if new inputs are provided.

Both linear (Hahne *et al.*, 2014) and nonlinear (Sadler, 1975) regression techniques have been extensively used for this purpose, using different estimation techniques to fit the data, namely several different flavors of the least-squares algorithms (Zhang *et al.*, 2011), ridge regression (Ming, 2014), etc.

Despite all the advantages of these traditional regression techniques, in all of them, it is necessary to make assumptions about the smoothness of the model.

While incorporating prior knowledge in the model that correctly describes the evolution of the available data can be of great value, sometimes this information is just not available. And using a model that does not correctly characterizes the data is likely to lead to poor results (Ostrom, 2010).

2.5.1 Gaussian process regression for prognosis

Unlike classical learning algorithm, Bayesian algorithms do not attempt to identify best-fit models of the data (or similarly, make best-guess predictions for new test inputs). Instead, they compute a posterior distribution over models (or similarly, compute posterior predictive distributions for new test inputs). These distributions provide a useful way to quantify the uncertainty in model estimates, and to exploit the knowledge of this uncertainty to make more robust predictions on new test points (Doshi-Velez, 2012; Ebden *et al.*, 2008; Williams & Rasmussen, 2006).

One of the Bayesian algorithms, Gaussian Process (GP), is by definition, a collection of random variables with the property that the joint distribution of any of its sub-set is joint Gaussian distribution. GPs are the powerful non-parametric technique with explicit uncertainty models, that are mainly used in regression and classification problems (Wang & Neal, 2012; Vanhatalo *et al.*, 2009).

Note that a clear distinction between a Gaussian distribution and a GP is that multivariate Gaussian distributions are useful for modelling finite collections of real-valued variables because of their nice analytical properties (Ebden *et al.*, 2008).

However, GPs are the extension of multivariate Gaussian to infinite-sized collections of real valued variable (Wang *et al.*, 2005). In particular, this extension allows to think of GPs as distributions not just over random vectors but in fact distributions over random functions.

In order to get an intuition for how GP work, consider a simple zero-mean GP (Anilkumar, 1994; Ebden *et al.*, 2008; Williams & Rasmussen, 2006),

$$f(x) \sim gp(0, k(x, x')) \tag{2.2}$$

defined for functions $f : X \to \mathbb{R}$. $k(x, x')$ represents the selected covariance function which is explained in more details in Chapter 8.

2.6 Conclusions and discussions

This chapter introduced a short survey of literature in the field of intermittent fault detection, isolation and prediction.

Research shows that early fault detection and prediction can minimize plant downtime, extended equipment life, increase the safety and reduce manufacturing costs and NFF. A number of issues must be considered when choosing particular fault detection method. Most important are

- Type of failures,
- Description of process structure,
- Process dynamics,
- Available process signals,
- Process complexity,
- Available amount of process input-output data,
- Process suitability for description in terms of rules.

Among the available fault detection methods, model-based fault detection includes process dynamics and nonmeasurable state variables (a set of the variable to describe the mathematical states of dynamic systems), however (Zhu & Jin, 2016),

- Requires accurate models and is easier to apply for well-defined processes such as electrical and mechanical than for thermal and chemical processes.
- These approaches are not well suited to handle uncertain systems. One solution is to model the uncertainties as unknown inputs and used the prescribed approaches.

- Another problem in these approaches is if the fault lies in the same subspace of the disturbances, then upon decoupling the disturbances, the residual will also be insensitive to a fault, which is not the objective of a detection filter. Moreover, the existence conditions for decoupling the unknown inputs are quite strict, which limits the use of these approaches.

- Moreover, designing an appropriate threshold to improve the fault detection mechanism is important. One author mentioned in this research presented the sensor's intermittent fault detection method for the class of uncertain linear systems. The considered system is subjected to the parameter uncertainty and limited resolution (Tao *et al.*, 2015). They used the upper bound of the minimized error at each step to detect the intermittent faults. However, the designed threshold cannot distinguish between the transient faults, measuring errors and intermittent faults that may raise false/miss alarms.

How to deal with these problems is explained in details in Chapters 5 and 6 .

Furthermore, if the basic relationship between faults and symptoms is known in the form of rules, experience (knowledge)-based methods are then the choice for the successful fault detection. However, when a large number of process input-output data can be obtained, but the process structure is unknown or too complex to be modelled, data-driven-based methods are more appropriate.

In this chapter, several studies have been introduced that used data-driven-based FD. One of the presented literature used the data-driven hoteling T2 statistic (T2 control chart) method to detect the intermittent faults (Zhao *et al.*, 2018). In their work to detect permanent faults by using the historical data, they were able to improve the detectability, however, for the intermittent fault detection, they need to derive sufficient conditions to guarantee the detectability of intermittent faults. Moreover, in their designed method, to detect the intermittent faults with small magnitude and short duration the optimal window length should be equal to the time duration of the intermittent fault, otherwise, they may miss the detection.

(Zhou *et al.*, 2020) in their recent publication reviewed present techniques into diagnostics of IFs for dynamic systems while compared and discussed their strengths and limitations. In their work they mentioned that the model-based IF detection is appropriate for systems with clear functional dependencies. However, when systems are very complex, then a deep and fundamental understanding of the system becomes very difficult. Hence, the model-based methods can not make an accurate IF detection which limit the application of model-based methods in real physical systems. The other considered approaches in the filed of the IF detection in this research are the data-driven approaches which have provided efficient solutions for IF detection in the complex industrial systems. This paper also claims that the current research in the field of IF detection needs more development because there are many unsolved problems which needs to be answered.

furthermore, (Lo *et al.*, 2019) and (Cai *et al.*, 2018) review applications of ML methods and different BNs in fault diagnosis respectively. These papers explain that although that the ML approaches are quite useful in the field of fault detection and diagnosis but a single ML method may not solves all the available issues in the fault diagnosis of a

	HDBN	DBN	SBN	Model-based	Knowledge-based	Other data-driven based
Graphical	Yes	Yes	Yes	No	Partly	No
Solution Confidence	Yes	Yes	No	Partly	Yes	No
Nonlinearity	Yes	Yes	Yes	Yes	Partly	Partly
Complexity	Yes	Yes	Yes	Partly	No	Partly
Time-dependence	Yes	Yes	No	Yes	Partly	Partly
No discretization	Yes	Yes	No	Yes	Partly	Partly
Discrete variables	Yes	No	Yes	Yes	Yes	Yes
Data from different sources	Yes	Yes	Yes	No	Partly	No
Uncertainty	Yes	Yes	Yes	Partly	No	Partly
Feedback loop	Yes	Yes	Yes	Yes	Partly	Partly
Hypothesis updating	Yes	Yes	Yes	No	Partly	No

Table 2.1: Comparison of the hybrid dynamic Bayesian network method with others to detect and isolate intermittent faults. In this table, HDBN indicates the hybrid dynamic Bayesian network, DBN indicates the dynamic Bayesian network and SBN indicates the static Bayesian network, (Zhou *et al.*, 2020; Lo *et al.*, 2019; Cai *et al.*, 2018; Salmerón *et al.*, 2018; Cai *et al.*, 2017)

physical system. However, some methods can complement other methods on an specific system. A compression of different fault detection techniques has been demonstrated in Table (2.1).

Moreover, the problem of intermittent fault detection for complex systems with a combination of discrete and continuous variables is a crucial one, especially when the system dynamics are not deterministic, all conditions of the system is not directly observed, and the sensors are subject to noise and disturbance.

One of the possible solution to this task is proposed based on the framework of HDBNs. This model contains both continuous variables representing the state of the system and discrete variables representing discrete changes such as failures. It can also model a variety of faults, including intermittent faults, burst faults, measurement errors, and gradual drifts.

The characteristics of the HDBN method is presented in Table (2.1) compared with other methods. The table is a compilation of commonly accepted characterization collected from the literature.

In (Salmerón *et al.*, 2018) the most important approaches for HDBN inference has been demonstrated. The authors explain that how the presence of the time-dependent variables along with the increasing number of variables can add to the inference complexity in complex dynamic systems which boost the need for approximate inference instead of exact inference.

This method along with some of these characteristics are presented in details in Chapter 7.

The final objective in this research is related to the intermittent fault prediction. Different fault prediction methods have been discussed in this chapter, however, not many of them have been developed to predict the intermittent fault in a complex dynamic system.

In this thesis, the Gaussian Process Regression (GPR) method has been selected for the intermittent fault prediction because GPR model is an example of a flexible, probabilistic, nonparametric Bayesian model with uncertainty predictions. It offers a range of advantages for modelling from data and has been therefore used for dynamic systems identification and time-series modelling.

Moreover, a completely different approach is given by GPs, by neglecting the parametric model viewpoint and instead define a prior probability distribution over all possible functions directly.

The reasons for choosing GPR for intermittent fault prediction over other methods in this thesis are summarised as:

- As Bayesian methods, GPR models allow one to quantify uncertainty in predictions resulting not just from intrinsic noise but also the errors in the parameter estimation procedure. Furthermore, many methods for model selection and hyperparameter selection in Bayesian methods are immediately applicable to GPs.
- GPR is non-parametric and hence can model essentially arbitrary functions of the input points.
- GPR models provide a natural way to introduce kernels into a regression modelling framework. By careful choice of kernels, GPR models can sometimes take advantage of structure in the data.

Furthermore, in addition to their use in regression, GPs apply to integration, global optimization, mixture-of-experts models, unsupervised learning models, and more. GPs do also allow the data to speak very clearly about themselves. This method is explained in depth in Chapter 8.

In the next chapter, the emphasis will be on the selected methodologies and solutions presented in this research.

Chapter 3

Methodology

3.1 Methodology and solutions

The work addressed in this thesis incorporates the development of the path to achieving the Intermittent Fault Detection and Prediction (IFDP) system capable of reliably detecting and predicting intermittent faults as they occur and to classify the location of the faults to aid in the maintenance of said systems. Hence, in order to achieve this, it is necessary to develop various stages of modelling to eventually achieve the goal of creating a robust IFDP filter (Figure 1.2). The sequence of development leading to this goal is detailed in the list below:

(a) Develop a mathematical presentation of the intermittent fault.

(b) Design two novel observers, NUI and feed-forward observers, for fault detection in general nonlinear systems (not only Lipschitz nonlinear systems, see Appendix A) with unmatched uncertainties where the uncertainties (unknown inputs) are not in the same channel as outputs necessarily. The designed observers are more sensitive to the uncertainties (unknown inputs) and consequently more robust to the intermittent faults. The designed observers have less limitations compare to other nonlinear observers and will give more degrees of freedom to the designer. As a result, more systems are capable of using these observers for robust fault detection.

(c) Create the nonlinear mathematical model of the systems under consideration in Matlab and validate them against the real system (nonlinear system modelling).

(d) Develop the state-space equations to represent the main dynamics of key components of the system models under investigation.

(e) Use the nonlinear state-space equations to design the observer-based FD filters for intermittent fault detection in the said systems.

(f) Develop improvements to the designed FD filters by using the residual evaluation and novel adaptive threshold design. The designed adaptive threshold presented

in this research considers the nonlinearity and unmatched uncertainty (unknown input) of the system in its design process to be more robust in comparison with the static thresholds or adaptive thresholds which do not consider those factors (residual and adaptive threshold design).

(g) Construct and compare probabilistic data-driven fault detection filters to detect and isolate intermittent fault in the experimental fuel rig system.

(h) Apply an appropriate non-parametric and probabilistic data-driven method, GPR, for the fuel rig system to predict the forthcoming intermittent faults for the steps ahead (intermittent fault prediction).

(i) Combine the intermittent fault detection and prediction techniques to create the IFDP system to detect and predict intermittent faults for the experimental fuel rig system (novel IFDP technique(s), Figure1.2).

3.2 Case studies

The efficiency of the proposed methodologies are mainly examined by an experimental fuel rig based in the Integrated Vehicle Health Management (IVHM) Centre Laboratory, at Cranfield University. The test rig nominally represents a Unmanned Aircraft Vehicle (UAV) fuel system consisting of several pipes, valves, pump, sensors and tanks.

This test rig is used to provide experimental data to develop and validate the proposed techniques. Moreover, the rig provides an opportunity to introduce intermittent fault behaviour for validation of the FD and Fault Prediction (FP) methods.

Furthermore, there are other case studies such as car suspension presented in this research which are used to present other model-based fault detection methods. Finally, the presented techniques enable the designer to choose the most appropriate fault detection filter for the system under investigation.

3.3 Software and toolboxes

In the first stage of this research, the model has been implemented and simulated in MATLAB® environment.

In the second stage of this research, the probabilistic data-driven methods are developed in BNT version 5.2 or newer in MATLAB® (versions 5.0 and 5.1 have a memory leak which seems to sometimes crash BNT). Some of the advantages of BNT over other software are:

- Very good implementation of the inference algorithms.
- The possibility of implementing dynamic BNs.
- The possibility to work with discrete nodes that have continuous parents (thanks to softmax CPDs).

- Good computational power.
- Easy to manipulate the data, for training and testing, generated with the Matlab m-files and Simulink models.

However, the BNT is slower than its equivalent written in C.

Furthermore, GPML toolbox version 4.2.2. in MATLAB® environment is used for further analysis and decision making.

3.4 Conclusions

The presented methods in this research are used to detect, isolate and predict the intermittent fault when the system subjected to the uncertainty. In this research

- First the model-based fault detection filters are developed to detect the intermittent faults in the complex systems subjected to the uncertainty (unknown inputs). Although these methods are very promising in detecting the faults but were performed poorly when the complexity of the system increases.
- Then the probabilistic data-driven methods were deployed to detect, isolate and predict the intermittent faults. Nevertheless, all the information captured from the model-based techniques were adopted in the data-driven methods as available reliable knowledge about the system under investigation.
- Finally, the effectiveness of the presented methods are validated using appropriate application systems.

The main application system, the experimental IVHM fuel rig, to validate the proposed techniques is presented in the next chapter.

Chapter 4

Introduction to the Application System

4.1 Introduction

This chapter provides a full description of the existing fuel rig and it's functioning in full as this is necessary, as background, for the data generating and mathematical modelling of the rig which are presented in the next chapters.

Fuel systems are some of the most critical subsystems of an Unmanned Aerial Vehicle (UAV), being responsible of delivering the required amount of flow to the engine(s) in order to ensure the accomplishment of the UAV mission(s). They need therefore to be continuously monitored during the mission in order to be able to capture fuel system failure modes and their symptoms during different operating conditions.

The rig on which this research project has focused is the IVHM centre fuel system test bed which can accommodate component and sensor specific faults and also it is capable of emulating different types of missions with the typical format of ramp up to taxi1 - taxi1 - take off - cruise - landing- taxi2 - ramp to shut down. As part of the rig design, sensors are placed in the system. The logic associated with each sensor set is capable of identifying every failure mode in the fuel system. Each of the sensors, interrogated by the diagnostic rules, requires a clear definition regarding the threshold between normal and abnormal conditions.

The capability of the system and the model to show effects on the sensor output of different faults with varying degrees of severity have been explored. The results of these trials in both normal and faulty operation show that the test system has excellent repeatability of measurements within the defined operational range (Niculita *et al.*, 2013).

This rig is then used in to collect hybrid data (discrete and continuous) in different phases from ramp up to shut down when the system is stable and controlled by a feedback controller. These data are then used for the validation and evaluation of the developed methods and algorithms for the intermittent fault detection, isolation and prediction.

This chapter is organized as follows: in section 4.2 the full description of the fuel rig is presented. Then Section 4.3 is given the fuel rig functional description. Finally, the model of the rig's components, the experimental conditions and the discussions are presented in Sections 4.4, 4.5 and 4.6 respectively.

4.2 IVHM experimental fuel rig description

The IVHM experimental fuel rig is responsible for delivering the appropriate quantity of fuel to the engine(sump tank).

The fuel rig in this research is sufficiently simple to be readily represented in software and to be easily realised in the laboratory. When the fuel rig is in its engine feeding mode, then the fuel will be transfer from the main tank to the sump tank via pipes between these two tanks. The external gear pump will help to move the fuel when the pressure before the pump is lower than the pressure after the pump. When the fuel rig is in its recirculation mode, the valve 2 and valve 4 will be open and some of the fuel will be returned to the main tank on it's way to the sump tank which is considered as a leak to the fuel system.

A connection between the drain valve of the sump tank and the refill valve of the main tank has been considered in order to realize a general reset of the test rig.

The considered example system in this research contains all the important components of an UAV fuel system; including valves, pumps, pipes, storage tanks; arranged in a configuration representative of that used in a large aircraft. It will not use real fuel, but initially water. A fluid such as hydraulic oil which has similar viscosity and density characteristics to real fuel is also acceptable.

The fuel rig also contains sufficient instrumentation (e.g. pressure sensors, flow meters, leak detectors and vibration sensors) so as to be able to develop tests and damage parameter measurements to diagnosis and prognosis developing system faults.

Moreover, it is capable of producing data for five different faults (Niculita *et al.*, 2013):

1) Clogged filter

2) Faulty gear pump

3) Faulty solenoid (shut off) valve

4) Leaking pipe

5) Clogged pipe .

Generally, the rig consists of the following representative components (Figure 4.1):

- Main and sump tanks,
- External gear pump,
- Filter,
- Polyurethane tubing,
- Solenoid shut-off valve, direct proportional valves,
- Non-return valve,
- Control modules for the pump, direct proportional valves and shut-off valve Instrumentation.

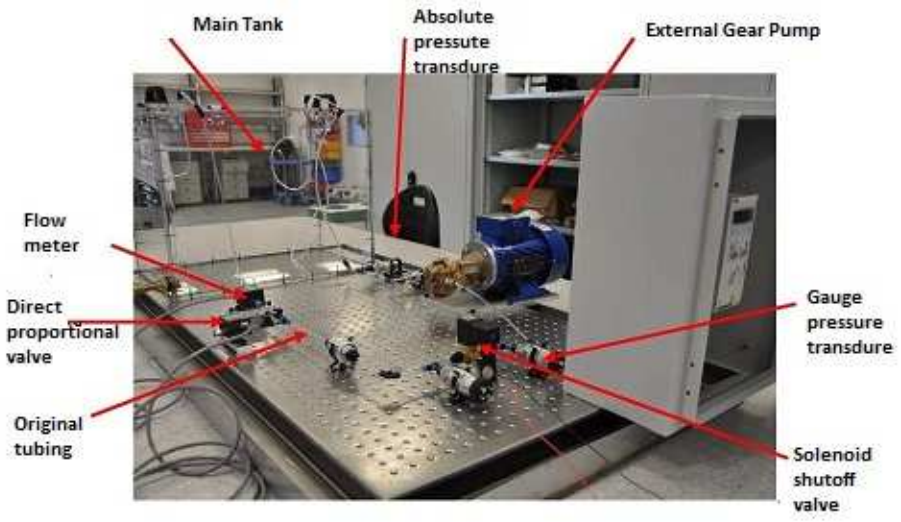

Figure 4.1: Fuel system test bed.

A connection between the drain valve of the sump tank and the refill valve of the main tank has been considered in order to realize a general reset of the test rig.

To control and acquire data from the fuel system, a system using National Instruments LabView virtual instrumentation has been developed using a CDAQ 9172 device with five compact modules: NI 9485, NI 9205, NI 9472, NI 9401 and NI 9263 (Figure 4.2).

Also, National Instruments LabView software version 8.6 is used to customize the control for the entire system. There are three main controls in the system:

- Pump control unit: manual or mission profile selection; speed; feedback loop,
- Valve control unit: shut-off valve position control,
- Direct proportional valve control unit: operated via the filter and gear pump fault sliders.

Moreover, the rig is operated in two modes:

- Engine feeding mode (Figure 4.3)
- Recirculation mode (Figure 4.4),

Moreover, the presented data to the user are from three parts:

- Pump speed in the pump control unit,
- Pressures in different points of the system (e.g. before filter, after filter, after pump, after shut-off valve),
- Time traces of pump speed, pressure, flow for both operational modes.

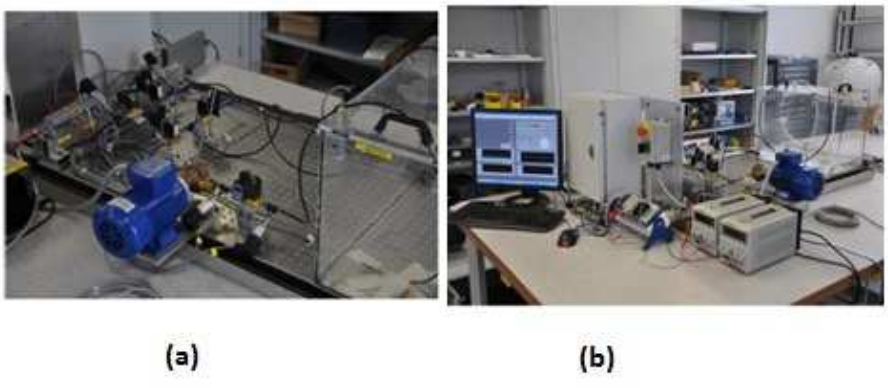

Figure 4.2: Fuel rig controls and data capture.

And these values are displayed according to running mode selection.

For pump control the user has two options: to run the pump manually by adjusting the pump speed or to set a couple of parameters (time, speed) in order to run the rig through a mission profile. This profile consists of UAV flight sectors: accelerate to taxi1 - taxi1 - take-off - cruise - landing - taxi2 - decelerate to shut-down.

From the hardware and software point of view, the control system is ready to accommodate all initially planned five failure modes in a plug and play manner although in this thesis only faulty solenoid shut-off valve, which has been installed and ready to provide data is considered.

4.3 IVHM fuel rig functional description

It is important to understand the functioning of the individual elements of the system, along with the sensors and the control system before considering the operation of the whole.

It is worth mentioning that these are low-cost components and sensors so all the operational data and equipment characterisation is not as readily available from manufacturers as it would be for flight worthy equipment (Niculita *et al.*, 2012).

4.3.1 Sensors

The instrumentation on the rig has a suite of nine sensors:

- Four gauge pressure sensors,
- Two absolute pressure sensors,
- Two flow meters,

and each of these elements can introduce an error in measurement.

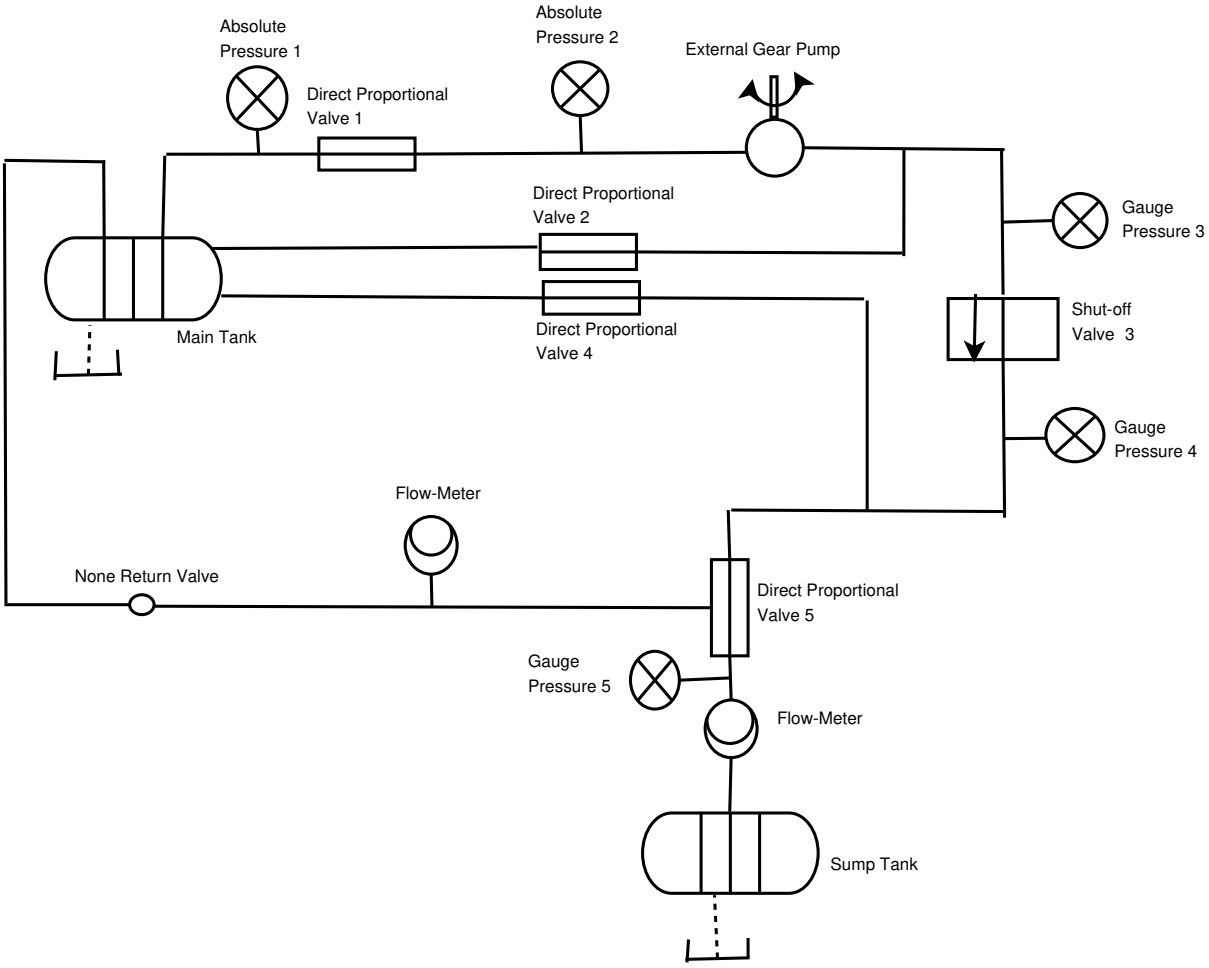

Figure 4.3: Fuel system demonstrator, engine feed mode.

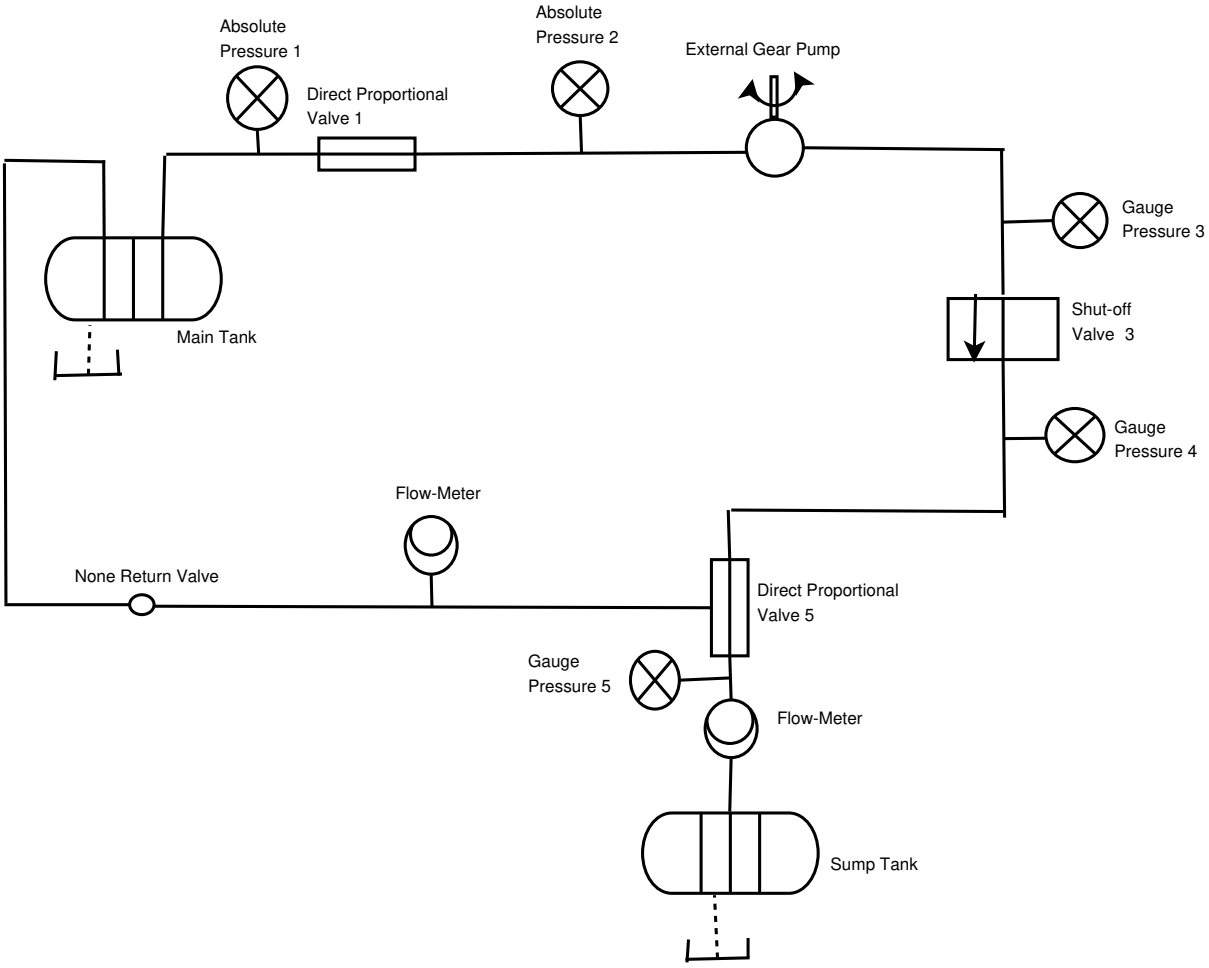

Figure 4.4: Fuel system demonstrator, recirculation mode.

GE Druck PMP 1400 industrial gauge pressure sensors

These sensors were initially specified and ordered for the rig. During experimental trials, it was realized that the pressure before the pump could drop below atmospheric. For this reason, a set of two absolute pressure sensors were obtained. Typical accuracy is rated at $\pm 0.15\%$ of span covering a range of $(0-4)$ *bar* with a 0.5 V output. More information about these sensors can be found in (Niculita *et al.*, 2012).

IMP absolute pressure sensors

The IMP industrial pressure transducer has a piezo-resistive ceramic pressure sensor giving it resistance to different fluids. The housing is made from stainless steel to ensure the product is suitable for a wide range of applications. Accuracy is rated as smaller than $\pm 0.25\%$ FS, Best Fit Straight Line (BFSL), covering the range $(0-5)$ *bar* with $(0-5)$ V output. More information about these sensors can be found in (Niculita *et al.*, 2012).

Omega engineering flow meters

The FLR 1011 series of flow meters from Omega Engineering is capable of measuring flows of $(0-2)$ $\frac{l}{min}$. Omega FLR 1011 flow meters used in the recirculation and engine modes has the characteristics in error terms introduced in (Niculita *et al.*, 2012).

4.3.2 External gear pump

For this rig, an Oberdorfen external gear pump with internal relief valve, model number $N999R$, was used. The data-sheet can be found in (Niculita *et al.*, 2012). The pump housing and gears are made of top quality bronze, shafts are 3030 stainless steel. Bearings are designed of high-performance carbon graphite selected for wear resistance and long service life. The gear pump is a positive displacement pump. Each shaft revolution displaces a definite amount of liquid relatively unaffected by the back pressure in the discharge line. According to the manufacturer, shaft rotational speed and volumetric flow rate are directly proportional. This has been measured on the rig without any other elements (valves, etc.) being present and the results are presented in Figure (4.5).

Figure (4.5) presents the variation of volumetric flow rate with rotational speed. The straight line between the measured points corresponding to 100rpm, 200rpm, 300rpm, 400rpm, and 500rpm shows the direct proportionality between volumetric flow rate and rotational speed by the instrumentation response.

Moreover, as shown in (Niculita *et al.*, 2012), the characteristics of volumetric flow and absolute pressure before and after the pump at $300rpm$ are not constant. As this test was conducted in (Niculita *et al.*, 2012) at $300rpm$ the frequency for one gear pump rotation is 5 Hz and the time between peaks should be 0.2 seconds. The sampling frequency for all measurements is set to 1 kHz, samples to read set to 500, and acquisition mode being set to continuous. Volumetric flow rates are varying between 0.605 and 0.61 $\frac{l}{min}$, pressure rates before and after pump are following a repetitive cycle synchronized with pump speed.

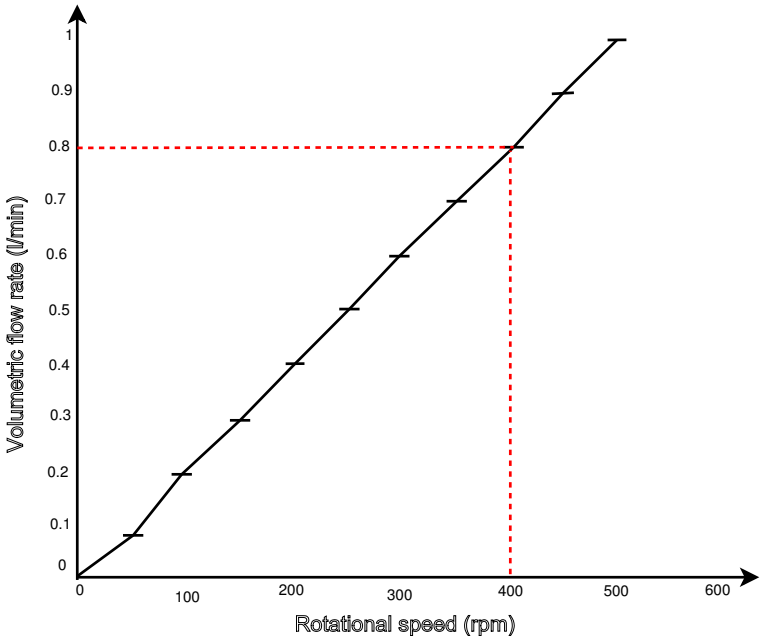

Figure 4.5: Pump characteristic.

4.3.3 Piping

Initially SMC clear pipe with $6mm/4mm$ OD/ID was used on the rig. Tests demonstrated that this tubing dilated under pressure and once a certain pressure was exceeded the tubing did not rapidly recover to its original dimensions, i.e. starting at $200rpm$ with an excursion to $500rpm$ upon return to $200rpm$ significantly different results were obtained. In order to deal with this issue, SMC polyurethane black tubing with $6mm/4mm$ OD/ID and a maximum pressure of $8bar$ was chosen. This has an operating temperature range of $-20°C$ to $60°C$ and a minimum bending radius of $15mm$. A simple way of checking functionality is to measure pressure drop against rotational speed for $1m$ and $0.5m$ lengths of this type of pipe, results shown in (Niculita *et al.*, 2012). It is clear that pressure losses increase directly with pipe length, if the length doubles then the pressure loss also doubles, as expected across the $(0-500)rpm$ range. Further testing has shown equally good, repeatable, results with this tubing.

4.3.4 Direct proportional valves

In order to simulate component degradation, some of the elements of the fuel system test rig (e.g. the filter) were replaced with Burkert 2833 Direct-acting Proportional Valves (DPV)s, $4mm$ internal diameter when fully open. These direct-acting proportional valves can be used as flow control valves and are characterized by low loss, low hysteresis, high repeatability and high sensitivity. These valves were characterized by measuring volumetric flow rate and pressure loss through the valve against valve opening positions, the opening positions being $(0-100)\%$ for $(0-10)$ Volts. More information about these sensors can be found in (Niculita *et al.*, 2012).

Reservoir: 3 arm	
Property	Value
Branch 1 Inflow Loss Coefficient	$1e-08$
Branch 1 Pipe Diameter	$0.004m$
Branch 1 Height above Base	$0.4m$
Branch 2 Outflow Loss Coefficient	$1e-08$
Branch 2 Inflow Loss Coefficient	$1e-08$
Branch 2 Pipe Diameter	$0.004m$
Branch 2 Height above Base	$0.4m$
Branch 3 Outflow Loss Coefficient	$1e-08$
Branch 3 Inflow Loss Coefficient	$1e-08$
Branch 3 Pipe Diameter	$0.004m$
Branch 3 Height above Base	0
Height of Top above Base	$0.4m$
Base Level above Base	$0m$
Horizontal Cross-Section Area	$0.15m^2$
Surface Pressure	1.01325 bar
Initial Liquid Level	$0.05m$
Initial Temperature	Not Set
Initial Pressure	Not Set

Table 4.1: Main tank parameters.

4.3.5 Shut-off valve

The shut-off valve introduces a pressure loss in the system. At zero rpm, the pressure on the inlet side of the shut-off valve is smaller than the outlet side because the pump inserts a bigger pressure loss in the system compared with the flow-meter. The pipes to the main tank have almost similar lengths so the pressure loss introduced by them is balanced on both sides of the shut-off valve.

4.4 Models of the fuel rig components

4.4.1 Main tank

The tank is represented in rig by the reservoir: 3-arm component. The parameters that need to be set for this component are: branch out flow/in flow loss coefficient, branch pipe diameter, branch height above base, and height of the top above base, base level above reference system, horizontal cross-sectional area, and initial liquid level (see Table 4.1).

Valve: Ball	
Property	Value
Diameter	$0.003m$
Valve Opening	1 ratio
Loss Coefficient vs Opening	Ball Valve Loss Coefficient
Correction for Re<= 10000	Laminar/Turbulant Loss Coefficient
Results On/Off	1.On

Table 4.2: Shut-off valve parameters.

Valve: Ball	
Property	Value
Diameter	$0.004m$
Valve Opening	1 ratio
Loss Coefficient vs Opening	DP Valve Loss Coefficient
Correction for Re<= 10000	Laminar/Turbulant Loss Coefficient
Results On/Off	1.On
Volumetric Flow Rate	Not Set
Mass Flow Rate	Not Set
Gas Flow Factor vs Valve Opening	Not Set
Flow Coefficient vs Valve Opening	Not Set
Joule Thomson Surface	Not Set

Table 4.3: Direct proportional valve parameters.

4.4.2 Valves

Shut-off valve

The shut-off valve is presented by a valve: Ball component. For this component, the internal diameter and the opening ratio has to be set along with the loss coefficient versus valve opening (see Table 4.2).

Direct proportional valve

The direct proportional valves are modelled as control Ball valves generic components. For these type of components input parameters needing setting are (see Table 4.3):

- The internal diameter,
- The opening ratio,
- The loss coefficient versus valve opening.

Pipe: Cylindrical Rigid	
Property	Value
Length	$0.04m$
Diameter	$0.004m$
Pipe Profile	Not Set
External Pressure vs Time	Not Set
Results On/Off	1.On
Friction Data (Cylindrical Pipe)	Sub Form

Sub-Form	
Property	Value
Friction Option	1. < 1 >::Colebrook-White Equation Approximate
Unsteady Friction Option	Not Set
Absolute Roughness	$0.1mm$
Hazen-Williams Friction Coefficient	110
Friction Factor	0.02

Table 4.4: Pipe parameters.

4.4.3 External gear pump

The Oberdorfen *N*999*R* external gear pump model was built through experimental trials to determine the pump characteristic.

Tests have been done across the whole established pump operational range and model is returning volumetric flow rates values with an error smaller than 0.1% compared with the data acquired on the rig (e.g. for the pressure rise value across the pump of 0.527595 *bar* and rotational speed of $500.0083rpm$ on the test rig it has been measured a volumetric flow rate value of $0.913354\frac{l}{min}$. While the model using the pressure rise value 0.5276 *bar* and rotational speed of $500.008rpm$ returns a volumetric flow value of $0.914785\frac{l}{min}$).

4.4.4 Pipes

A pipe is characterized by the pipe cylindrical rigid component and length, absolute roughness and internal diameter parameters have to be set. Additionally, friction data parameters were introduced (see Table 4.4).

The pump model was used to generate the volumetric flow rate used for pipe testing. The test observed the pressure drop variation with the change of the pipe length for the pipe model.

4.5 experimental conditions

As mentioned earlier the fuel rig can be configured to run in two different modes, engine feed mode and recirculation mode. The analysis presented in this thesis will only

Component	Failure mode
Shut-off valve	Stuck mid range
Gear pump	Degraded – less flow for the same pump speed
Pipe 3 (pipe before the shut-off valve)	Leaking
pipe 4 (pipe after the shut-off valve)	Leaking
Nozzle	Clogged
Filter	Clogged

Table 4.5: Failure modes considered for investigation

consider the engine feed mode. There are five pressure sensor sets in this fuel rig and the logic associated with each sensor set is capable of identifying every failure mode considered into the analysis.

The overall function of the IVHM fuel system, under engine feed operation mode, will be to provide fuel where:

- Input flows are considered to be voltage and representing electrical energy as a discrete signal.
- Output flow is volume and representing liquid material.

For the overall function case, the voltage is selected for the electric energy along with the signal type of flow, and the volume option is selected as output flow for material. The justification for these selections is that they describe the fuel system's main task which is to deliver a specific amount of fuel by setting control values (pump speed, valve positions) and powering several components.

In general, five components are considered to be affected by a particular failure mode :filter, gear pump, shut-off valve, one of the pipes and nozzle (see Table (4.5)).

For each of the failure modes, various degrees of severities can be simulated in real conditions on the fuel system by making use of direct-acting proportional valves.

4.5.1 parameter and data collection

The data for the experiments presented in this thesis was captured using NI Labview Interface which is a systems engineering software. The goal was to observer the pressure changes in all pressure sensors (outputs) when IF (shut-off valve clogging) was injected to the system undre the following conditions:

- Steady state conditions were used during the experiments carried out to obtain the data (pump running at 400 rpm)
- The fuel rig system is a Single Input-Mmulti Output (SIMO) system
- The input parameter is the pump speed

- The output being represented by the volumetric flow rate, and the pressure rates at five different locations (pressure sensors $1-5$) as continuous values
- The characteristic of the valves (open, closed, partly closed) were captured as discrete values
- The sampling time for collecting the discrete data was 0.001
- To inject the IF into the fuel rig the status of the shut-off valve (controller) was changed manually
- Input pressure to the pipe at the end of the run is the atmospheric pressure
- The fuel in this rifg is water at room temperature
- The transition rate for valves (time between opening and closing:$(0-90\%)$) is around 20 mls

The corresponding simulation results of this experiment is presented in Chapter 6 for further analysis.

4.6 Discussions

This chapter has introduced the IVHM fuel rig and its operational capabilities.

This fuel rig is build and used in IVHM centre, Cranfiel university for assessment of various fault diagnostic software tools and techniques. The IVHM Centre fuel rig has been fully commissioned and a verification and validation process was carried out. Tests considered both component and system level approach. Control and data acquisition software was verified as well and the initial requirements have been met. Filter blockage and faulty pump type of faults have been simulated and their effects on pressure and flow rates at defined locations at a range of fault severities have been monitored and the results can readily be used for diagnostic test purposes (Niculita *et al.*, 2013).

Since the IVHM Centre fuel rig meets the initial requirements of a demonstrator for the diagnostic and prognostic research area, then the intermittent fault diagnosis of this rig where no changes were made on the hardware of the system was performed for this research (Sedighi *et al.*, 2015). The process and the results of the IF detection, isolation and prediction using this rig is presented in the following chapters.

Chapter 5

Model (Observer)-Based Intermittent Fault Detection

5.1 Introduction

In this chapter, two novel model (observer)-based FD filters are introduced as follows:

1) NUIO-based FD: the existing theory on the NUIO design to a wider class of nonlinear systems with general nonlinearity is extended, which are subject to bounded unmatched unknown inputs (disturbances) and experiencing intermittent faults.
In the proposed method, first, the nonlinearities, unknown inputs (disturbances) and sensor noises are estimated from the available known parameters, and their corresponding error equations are obtained. Subsequently, the system error is written in terms of other parameter errors. If each error is stable, then the stability of the system error will be achieved. This approach provides a straight-forward technique in designing NUIO, which can reduce the state estimation errors against the unmatched unknown inputs (disturbances).
Then the robust FD problem of the considered system is studied when a nonlinear unknown input observer is provided and sufficient conditions are given to make the observers asymptotically stable .
Note that in model-based FD the terms uncertainty, unknown input and disturbance are the same.

2) Feed-forward observer-based FD: provided an approach to design feed-forward observer for nonlinear systems with Lipchitz nonlinearity and bounded unknown inputs (disturbances/uncertainties) to ensure the sensitivity against intermittent faults. The proposed observer design guarantees the system error stability. Some variables and scalars are also introduced to design observer's parameters, which bring more degrees of flexibility available to the designer. The designed observer is used to propose a precision fault detection scheme including an adaptive threshold design to detect intermittent faults. The efficiency of the considered approach is examined by the intermittent failure case in the suspension system of a vehicle.

This chapter is organized as follows: Section 5.2 presents the mathematical description of the nonlinear system of interest to this thesis. The design of the NUIO and feedforward observer along with theorems and error dynamic stability analysis are addressed in Sections 5.3-5.6 respectively. The residual and appropriate adaptive threshold is designed in Section 5.7 while the numerical examples and simulation results are presented in Sections 5.8 and 5.9. Finally, the conclusions are given in Section 5.10.

5.2 System description

State-space equations are a set of differential equations make use of the notions of states, inputs, outputs and dynamics to describe the behaviour of a system (Adhyaru, 2012).
The state is a collection of variables that summarize the past of a system for prediction of the future. For an engineering system the state is composed of the variables required to account for the conservation of mass, momentum and energy (Samadi & Saif, 2017). For system modelling the state variables are gathered in a vector, $x \in \mathbb{R}^n$ called the state vector, the control variables are represented by another vector $u \in \mathbb{R}^p$ and the measured signals by the vector $y \in \mathbb{R}^q$. Then a system can then be represented by a set of differential equation (Mattei, 2001):

$$\begin{aligned} \dot{x} &= h_x(x,u) \\ y &= h_y(x,u_y) \end{aligned} \tag{5.1}$$

where $\dot{x} = \frac{dx}{dt}$.

A model of this form is called a state space model and the dimension of the state vector is called the order of the system (Choi & Chung, 1996).
The model consists of two functions. The function h_x gives the velocity of the state vector as a function of state x and control u, and the function h_y gives the measured values as functions of state x and control u, (Mutoh, 2009).
The system is called time-invariant because the functions h_x and h_y do not depend explicitly on time t. It is possible to have more general time-varying systems where these functions do depend on time. A linear state-space system can then be represented by

$$\begin{aligned} \dot{x} &= Ax(t) + Bu(t) \\ y &= Cx(t) + Du_y(t) \end{aligned} \tag{5.2}$$

where A, B, C and D are constant matrices.
A more general system is obtained by letting the output be a linear combination of the states of the following system,

$$y = c_n x_1 + c_{n-1} x_2 + \cdots + c_1 x_n + d_1 u_y, \tag{5.3}$$

where this system can be modeled in the state-space form as (Koenig, 2005):

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 1 \\ -a_n & -a_{n-1} & \cdots & \cdots & -a_1 \end{pmatrix} x + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} u$$

$$y = \begin{pmatrix} c_n & c_{n-1} & \cdots & c_1 \end{pmatrix} x + d_1 u_y. \tag{5.4}$$

Now, consider a class of nonlinear systems defined by the following state-space form (Peñarrocha *et al.*, 2009)as follows:

$$\dot{x}(t) = h_x(x,u,\mu)$$

$$y(t) = h_y(x,\mu_y). \tag{5.5}$$

If the nonlinear function $h_x(x,u,\mu)$ is differentiable with respect to x, then this class of the system may be expressed in terms of a linear unforced part, and nonlinear state dependent controlled parts, (Rajmani, 1998) as follows:

$$\dot{x}(t) = Ax(t) + Bu(t) + D\mu(t) + Sg(x,u,t)$$

$$y(t) = Cx(t) + D_y\mu_y(t) \tag{5.6}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$ present state, input and output vectors, respectively; $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$ and $S \in \mathbb{R}^{n\times s}$ are known matrices correspondingly; $D \in \mathbb{R}^{n\times q}$ and $D_y \in \mathbb{R}^{p\times r}$ are referred to the known distribution matrices of the unknown input (disturbance) and sensor noise, respectively; $\mu(t) \in \mathbb{R}^q$ and $\mu_y(t) \in \mathbb{R}^r$ are the unknown bounded vectors which describe the unknown input and/or any kind of modeling uncertainty such as noise, time-varying term, and parameter variation in both component/ actuators and sensors, respectively.
This chapter treats general nonlinearities that depend on unmeasured states, but for illustration, a nonlinearity of the form $g(x,u,t) \in \mathbb{R}^s$ has included in the design procedure. Note that to analyze the error stability the error estimation equation should be perform in fault-free case, where $f_i(t) = f_{i_s}(t) = 0$ while $f_i \in \mathbb{R}^{r_i}$ and $f_{i_s} \in \mathbb{R}^{r_s}$ are the actuator/component and sensor faults correspondingly.

5.3 Nonlinear unknown input observer design

Prior to presenting the NUIO design, the following assumptions are made:

Assumption 1

- *No measurement depends directly on one of the states, which is affected by the unknown input (disturbance): $CD = 0$.*
- *Matrices D, S and D_y are full column rank matrices.*
- *The output $y(t)$ and its derivative $\dot{y}(t)$ are available, where $\dot{y}(t)$ could be estimated using a robust differentiator from the output y.*

Assumption 2

- *The unknown input (disturbance) is bounded to some positive constant d,*

$$\|\mu(t)\| \leq d. \tag{5.7}$$

- *The sensor noise, $\mu_y(t)$, is not constant and*

$$\|\dot{\mu}_y(t)\| \leq \kappa \|e_x(t)\| \tag{5.8}$$

where κ is a positive constant and $e_x(t)$ is the state estimation error vector which will define later.

The NUIO is usually designed such that its state estimation error vector $e_x(t)$, approaches to zero asymptotically, regardless of the presence of the unknown input (disturbance) in the system. To design such an observer normally the rank condition, rank(CD)=rank(D), must be satisfied, (Mondal *et al.*, 2009; Chen & Saig, 2006a,b). However, for the system under investigation this rank condition is not satisfied since $CD = 0$ (see Assumption 1).

Hence to avoid the effect of the unknown input/disturbance term, $\mu(t)$, on the error estimation of the system, the novel observer of the following form is given

$$\begin{aligned} \dot{z}(t) &= Nz(t) + Ly(t) + Gu(t) + HSg(\hat{x},u,t) + HD\hat{\mu}(t) + H_\mu D_y \hat{\mu}_y(t) \\ \hat{x}(t) &= z(t) - Ey(t) \end{aligned} \tag{5.9}$$

where $z \in \mathbb{R}^n$ is the state observer and $\hat{x}(t)\mathbb{R}^n$, $g(\hat{x},u,t)\mathbb{R}^s$, $\hat{\mu}(t)\mathbb{R}^q$ and $\hat{\mu}_y(t)\mathbb{R}^r$ are the estimations of $x(t)$, $g(x,u,t)$, $\mu(t)$ and $\mu_y(t)$ respectively. Matrices $N \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{n\times p}$, $G \in \mathbb{R}^{n\times m}$, $H \in \mathbb{R}^{n\times n}$, $H_\mu \in \mathbb{R}^{n\times p}$ and $E \in \mathbb{R}^{n\times p}$ should be determined as explained below.

It is desired to design an observer such that the estimated states $\hat{x}(t)$ tend to actual states $x(t)$ eventually. Hence the error equation of the form

$$e_x(t) = x(t) - \hat{x}(t) = x(t) - z(t) + Ey(t) \tag{5.10}$$

is defined for system (5.6) and observer (5.9) in terms of x.
Then, estimation error (5.10) is rewritten as follows:

$$\begin{aligned} \dot{e_x}(t) &= \dot{x}(t) - \dot{z}(t) + EC\dot{x}(t) + ED_y\dot{\mu}_y(t) \\ &= Ne_x(t) + (HA - NH - LC)x(t) + (HB - G) \\ &\quad u(t) + HS(g(x,u,t) - g(\hat{x},u,t)) + HD(\mu(t) - \\ &\quad \hat{\mu}(t)) + H_\mu D_y(\mu_y(t) - \hat{\mu}_y(t)) + ED_y\dot{\mu}_y(t) \end{aligned} \tag{5.11}$$

where

$$H = I_n + EC \tag{5.12}$$

and

$$H_\mu = -(NE + L). \tag{5.13}$$

By constructive design, the following matrix equations give a NUIO:

$$HD \neq 0 \tag{5.14}$$

$$HB - G = 0 \tag{5.15}$$

$$HA - NH - LC = 0. \tag{5.16}$$

The error equation given in (5.11) is rewritten as:

$$\dot{e_x}(t) = Ne_x(t) + A_gHSe_g(t) + A_dHDe_\mu(t) + A_iH_\mu D_y e_i(t) + ED_y\dot{\mu}_y(t) \tag{5.17}$$

where $e_g(x,u,t) = g(x,u,t) - g(\hat{x},u,t)$, $e_\mu(t) = \mu(t) - \hat{\mu}(t)$ and $e_i(t) = \mu_y(t) - \hat{\mu}_y(t)$ denote the nonlinearity, unknown input and sensor noise errors, respectively. Matrices A_g, A_d, and A_i are design matrices which will be defined later in Section (5.4).
In order to write (5.9) in well-known NUIO form, the gain matrix K is introduced in such a way that

$$N = HA - KC. \tag{5.18}$$

For stability, K must be chosen such that N is Hurwitz, where poles of all elements of N have negative real part, which is always possible under the detectability assumption of the pair (HA, C). If strengthen the detectability assumption to observability of (HA, C), then the eigenvalues of N can be placed arbitrary, see (Darouach *et al.*, 1994; Chen *et al.*, 1996).

Then the procedure to design the observer (5.9), can be summarized as follows:

- Select a full rank matrix $D_1 \in \mathbb{R}^{n\times d_1}$ which for each $d_1 \leq m$ satisfies

$$CD_1 \neq 0,$$

$$HD_1 = 0, \tag{5.19}$$

and rank(CD_1)= rank(D_1). Matrix D_1 should also make the pair (HA, C) detectable.

- Substitute (5.12) into (5.19), to obtain matrix E of the form

$$E = -D_1(CD_1)^+ + \Gamma(I_{d_1} - (CD_1)(CD_1)^+). \tag{5.20}$$

Since D_1 is a full column rank one necessary condition for $ECD_1 = -D_1$ to have solution is that CD_1 is also of full column rank. If CD_1 is full column rank, then all possible solutions of $ECD_1 = -D_1$ must have the form (5.20) where $(CD_1)^+$ is the pseudo-inverse of (CD_1) and Γ is an arbitrary matrix.

- Using (5.20), H can be found out from equation (5.12).

- Substitute matrix H into (5.15) to obtain matrix G.

- Assume that the pair (HA, C) is an observable pair and P_0 is a Symmetric Positive Definite (s.p.d.) solution of the following Algebraic Riccati Equation (ARE),

$$(HA)^T P_0 + P_0(HA) - P_0 C^T R^{-1} C P_0 = -Q_0 \tag{5.21}$$

where $Q_0 \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{p\times p}$ are arbitrary s.p.d. matrices. Then by selecting the gain matrix $K = P_0 C^T R^{-1}$, the matrix $N = HA - KC$ is an stable matrix.

- Once K is obtained, substitute N, K and (5.12) into (5.16) to achieve the observer gain L as:

$$L = -HAE + K(I_p + CE). \tag{5.22}$$

5.4 Sufficient conditions for existence of the NUIO

To show that the observer given in (5.9) is indeed a NUIO, it is desired to achieve the stability of the errors $e_\mu(t)$, $e_g(x,u,t)$ and $e_i(t)$ to make the error dynamics of the system (5.17) stable.
In the next subsections, the stability of these errors will be investigated.

5.4.1 Stability analysis of the nonlinearity error $e_g(x,u,t)$

The nonlinearity $g(x,u,t)$ could be estimated from the available known signals,

$$g(\hat{x},u,t) = G_{11}y(t) + G_{12}\dot{y}(t) + G_{13}\hat{x}(t) + G_{14}\dot{\hat{x}}(t) + G_{15}u(t) \tag{5.23}$$

where matrices $G_{11} \in \mathbb{R}^{s\times p}$, $G_{12} \in \mathbb{R}^{s\times p}$, $G_{13} \in \mathbb{R}^{s\times n}$, $G_{14} \in \mathbb{R}^{s\times n}$ and $G_{15} \in \mathbb{R}^{s\times m}$ should be designed in order to obtain $g(\hat{x},u,t)$, (Imsland & Fossen, 2007; Liu & Peng, 2002).

Moreover, from Equation (5.6), $g(x,u,t)$ is given by

$$g(x,u,t) = S^{+}\dot{x}(t) - S^{+}Ax(t) - S^{+}Bu(t) - S^{+}D\mu(t). \tag{5.24}$$

By substituting equations (5.23) and (5.24) into $e_g(x,u,t)$, the nonlinearity error is presented as

$$\begin{aligned} e_g(x,u,t) &= g(x,u,t) - g(\hat{x},u,t) \\ &= -(S^{+}A + G_{11}C)x(t) + (S^{+} - G_{12}C)\dot{x}(t) - \\ &\quad (S^{+}B + G_{15})u(t) - S^{+}D\mu(t) - G_{12}D_y\dot{\mu}_y(t) \\ &\quad -G_{11}D_y\mu_y(t) - G_{13}\hat{x}(t) - G_{14}\dot{\hat{x}}(t). \end{aligned} \tag{5.25}$$

which could be modified into the following form

$$\begin{aligned} e_g(x,u,t) &= -(S^{+}A + G_{11}C + G_{13})x(t) + (S^{+} - G_{12}C \\ &\quad -G_{14})\dot{x}(t) - (S^{+}B + G_{15})u(t) - S^{+}D\mu(t) \\ &\quad -G_{12}D_y\dot{\mu}_y(t) - G_{11}D_y\mu_y(t) + G_{13}e_x(t) + \\ &\quad G_{14}\dot{e}_x(t). \end{aligned} \tag{5.26}$$

To verify the stability of $e_g(x,u,t)$, the following conditions must hold

$$\begin{gathered} G_{13} = -S^{+}A - G_{11}C \\ G_{14} = S^{+} - G_{12}C \\ G_{15} = -S^{+}B \\ S^{+}D = 0 \\ -G_{11}D_y = 0 \\ -G_{12}D_y = 0 \end{gathered} \tag{5.27}$$

Therefore, the error $e_g(x,u,t)$, will have the following form

$$\begin{aligned} e_g(x,u,t) &= (-S^+A - G_{11}C)e_x(t) + (S^+ - G_{12}C)\dot{e}_x(t) \\ &= G_{13}e_x(t) + G_{14}\dot{e}_x(t). \end{aligned} \tag{5.28}$$

which is a function of $e_x(t)$. Since $e_x(t) \to 0$, then $e_g(x,u,t) \to 0$ is asymptotically stable.

5.4.2 Stability analysis of the unknown input (disturbance) error $e_\mu(t)$

The unknown input (disturbance) $\mu(t)$, can also be estimated from the available known signals as

$$\hat{\mu}(t) = M_{11}y(t) + M_{12}\dot{y}(t) + M_{13}\hat{x}(t) + M_{14}\dot{\hat{x}}(t) + M_{15}Sg(\hat{x},u,t) + M_{16}u(t) \tag{5.29}$$

where matrices $M_{11} \in \mathbb{R}^{q\times p}$, $M_{12} \in \mathbb{R}^{q\times p}$, $M_{13} \in \mathbb{R}^{q\times n}$, $M_{14} \in \mathbb{R}^{q\times n}$, $M_{15} \in \mathbb{R}^{q\times n}$ and $M_{16} \in \mathbb{R}^{q\times m}$ should be designed.
The unknown input (disturbance) $\mu(t)$, clearly is given from the states equation (5.6),

$$\mu(t) = D^+\dot{x}(t) - D^+Ax(t) - D^+Bu(t) - D^+Sg(x,u,t). \tag{5.30}$$

By using (5.29) and (5.30), the unknown input (uncertainty) error equation, becomes

$$\begin{aligned} e_\mu(t) &= \mu(t) - \hat{\mu}(t) \\ &= -(D^+A + M_{11}C)x(t) + (D^+ - M_{12}C)\dot{x}(t) - (D^+ \\ &\quad B + M_{16})u(t) - D^+Sg(x,u,t) - M_{15}Sg(\hat{x},u,t) \\ &\quad -M_{13}\hat{x}(t) - M_{14}\dot{\hat{x}}(t) - M_{12}D_y\dot{\mu}_y - M_{11}D_y\mu_y. \end{aligned} \tag{5.31}$$

Since $e_x(t) = x(t) - \hat{x}(t)$, then (5.31) could be rearranged as

$$\begin{aligned} e_\mu(t) &= -(D^+A + M_{11}C + M_{13})x(t) + (D^+ - M_{12}C - \\ &\quad M_{14})\dot{x}(t) - (D^+B + M_{16})u(t) - (D^+ + M_{15})S \\ &\quad g(x,u,t) + M_{13}e_x(t) + M_{14}\dot{e}_x(t) + M_{15}S(G_{13} \\ &\quad e_x(t) + G_{14}\dot{e}_x(t). \end{aligned} \tag{5.32}$$

Now, if the following conditions hold

$$M_{14} = D^+ - M_{12}C$$

$$M_{13} = -D^{+}A - M_{11}C$$

$$M_{16} = -D^{+}B$$

$$M_{15} = -D^{+}$$

$$M_{12}D_y = 0$$

$$M_{11}D_y = 0 \tag{5.33}$$

then Equation (5.32) can be rewritten as

$$e_{\mu}(t) = (M_{15}SG_{13} + M_{13})e_x(t) + (M_{15}SG_{14} + M_{14})\dot{e_x}(t). \tag{5.34}$$

Equation (5.34) shows that the unknown input (disturbance) error, $e_{\mu}(t)$ is a function of $e_x(t)$. Obviously if $e_x(t) \to 0$, then $e_{\mu}(t) \to 0$, is asymptotically stable.

5.4.3 Stability analysis of the sensor noise error $e_i(t)$

Similarly, the sensor noise $\mu_y(t)$, could be estimated from the available known signals as follows

$$\hat{\mu_y}(t) = F_{11}y(t) + F_{13}\hat{x}(t) \tag{5.35}$$

where matrices $F_{11} \in \mathbb{R}^{r \times p}$ and $F_{13} \in \mathbb{R}^{r \times n}$ need to be designed.

Rearranging Equation (5.6), will also give $\mu_y(t)$,

$$\mu_y(t) = D_y^{+}y(t) - D_y^{+}Cx(t). \tag{5.36}$$

By substituting equations (5.35) and (5.36) into the sensor noise error, $e_i(t)$ is expressed as

$$e_i(t) = \mu_y(t) - \hat{\mu_y}(t) =$$

$$D_y^{+}y(t) - D_y^{+}Cx(t) - F_{11}y(t) - F_{13}\hat{x}(t). \tag{5.37}$$

Equation (5.37) could be modified by substituting $\hat{x}(t) = x(t) - e_x(t)$ of the form

$$e_i(t) = (-F_{11}C - F_{13})x(t) + D_y^{+}D_y\mu_y(t) - F_{11}D_y\mu_y(t) + F_{13}e_x(t). \tag{5.38}$$

Thus, if the following conditions hold

$$(I_r - F_{11}D_y) = 0$$

$$F_{13} + F_{11}C = 0 \tag{5.39}$$

where I_r is an identity matrix of size r, the sensor noise error equation (5.38) could be represented as

$$e_i(t) = -F_{11}Ce_x(t)$$

$$= F_{13}e_x(t). \tag{5.40}$$

This equation shows that the sensor noise error, $e_i(t)$ is a function of $e_x(t)$, and will converge to zero if $e_x(t)$ tends to zero asymptotically.

5.4.4 Sufficient conditions for NUIO existence

By substituting errors (5.28), (5.34) and (5.40) into the error equation (5.17), the error dynamic of the system could be reformulated as a linear equality of the form

$$\bar{M}\dot{e}_x(t) = \bar{G}e_x(t) + ED_y\dot{\mu}_y(t) \tag{5.41}$$

with

$$\bar{M} = I_n - A_dHD(M_{15}SG_{14} + M_{14}) - A_gHSG_{14}, \tag{5.42}$$

and

$$\bar{G} = N + A_dHD(M_{15}SG_{13} + M_{13}) + A_gHSG_{13} + A_iH_\mu D_yF_{13}, \tag{5.43}$$

where I_n is an identity matrix of size n.
Since $\bar{M}$ is a nonsingular matrix, $A_d \in \mathbb{R}^{n\times n}$, $A_g \in \mathbb{R}^{n\times n}$ and $A_i \in \mathbb{R}^{n\times n}$ should be selected in order to make $\bar{M}^{-1}\bar{G}$ Hurwitz. The error equation (5.41) can then be rewritten as

$$\dot{e}_x(t) = \bar{M}^{-1}\bar{G}e_x(t) + \bar{M}^{-1}ED_y\dot{\mu}_y(t). \tag{5.44}$$

To show the asymptotic stability of the error equation (5.44), the following theorem needs to be made.
Note that to simplify the notations, the index x will be omitted henceforth.

Theorem 1: The error system (5.44) will be asymptotically stable if the following Matrix Inequality (MI)

$$\bar{N}_1P + P\bar{N}_1 + \varepsilon_1P^2 + \frac{\kappa^2}{\varepsilon_1}\bar{N}_2\bar{N}_2^T < 0 \tag{5.45}$$

for some $\varepsilon_1 > 0$, has the s.p.d. solution P where $\bar{N}_1 = \bar{M}^{-1}\bar{G}$ and $\bar{N}_2 = \bar{M}^{-1}ED_y$.

Proof: By substituting the error equation (5.44) into the time derivative of Lyapunov equation, $V(e) = e^TPe$, the following will be obtained

$$\dot{V} = e^T\left(\bar{N_1}^TP + P\bar{N}_1\right)e + e^TP\bar{N}_2\dot{\mu}_y(t) + \dot{\mu}_y^T(t)\bar{N}_2^TPe. \tag{5.46}$$

Since for any matrices X, Y and any positive number $\varepsilon > 0$, if

$$\varepsilon \left(\frac{1}{\varepsilon}X - Y\right)^T \left(\frac{1}{\varepsilon}X - Y\right) \geq 0$$

then

$$X^T Y + Y^T X \leq \frac{1}{\varepsilon}X^T X + \varepsilon Y^T Y, \qquad \forall \varepsilon > 0$$

hence for any $\varepsilon_1 > 0$,

$$e^T P \bar{N}_2 \dot{\mu}_y(t) + \dot{\mu}_y(t)^T \bar{N}_2^T P e \leq \varepsilon_1 e^T P P e + \frac{1}{\varepsilon_1}\bar{N}_2^T \|\dot{\mu}_y(t)\|^2$$

$$\leq (\varepsilon_1 P^2 + \frac{\kappa^2}{\varepsilon_1}\bar{N}_2\bar{N}_2^T)\|e(t)\|^2. \tag{5.47}$$

see Assumption 2.
Therefore the following inequality will be satisfied

$$\dot{V} \leq e(t)^T[(\bar{N}_1^T P + P\bar{N}_1) + \varepsilon_1 P^2 + \frac{\kappa^2}{\varepsilon_1}\bar{N}_2\bar{N}_2^T]e(t) < 0. \tag{5.48}$$

Corollary 1: Assume that the conditions of *Theorem 1:* are satisfied and P is the s.p.d. solution of

$$\bar{N}_1 P + P\bar{N}_1 = -Q \tag{5.49}$$

with an appropriate s.p.d. matrix Q.
Then the error equation (5.44) is asymptotically stable if there exists $\varepsilon_1 > 0$ such that

$$\lambda_{min}(Q) > \varepsilon_1 \lambda_{max}^2(P) + \omega_1 \tag{5.50}$$

where $\omega_1 = \frac{\kappa^2}{\varepsilon_1}\sigma_M(\bar{N}_2)$, and σ_M represents the maximum singular value of the corresponding matrix.
It is easy then to show that (5.48) gives

$$\dot{V} \leq e(t)^T[-Q + \varepsilon_1 P^2 + \frac{\kappa^2}{\varepsilon_1}\bar{N}_2\bar{N}_2^T]e(t)$$

$$\leq [-\lambda_{min}(Q) + \varepsilon_1 \lambda_{max}^2(P) + \frac{\kappa^2}{\varepsilon_1}\sigma_M(\bar{N}_2)]\|e(t)\|^2 < 0 \tag{5.51}$$

which also satisfies (5.50).

Hence the asymptotic stability of the system error estimation (5.17) and the existence of the NUIO given by (5.9) are guaranteed.

5.5 Nonlinear feed-forward observer design

Consider the nonlinear system described in (5.6). Prior to feed-forward observer design, the following assumptions are made:

Assumption 3

- *The pair (A,C) is observable.*
- *Nonlinearity $g(x,u,t)$ is assumed to be globally Lipschitz in x (see Appendix A) with Lipschitz constant κ, i.e.*

$$\|g(x,u,t)-g(\hat{x},u,t)\| \leq \kappa\|x-\hat{x}\|. \tag{5.52}$$

- *Measurement depends directly on one of the states, which is affected by the unknown input (uncertainty), $CD \neq 0$.*
- *The disturbance is bounded to some positive constant α, $\| \mu(t) \| \leq \alpha$.*

The nonlinear feed-forward observer is designed such that its state estimation error vector $e_x(t)$ approaches to zero asymptotically, regardless of the presence of the unknown input term in the system. Hence the observer of the following form is introduced:

$$\begin{aligned} \dot{z}(t) &= Nz(t)+Ly(t)+Gu(t)+HSg(\hat{x},u,t)+\Omega\nu \\ \hat{x}(t) &= z(t)-Ey(t) \end{aligned} \tag{5.53}$$

where $z \in \mathbb{R}^n$ is the state observer, with matrices $N \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{n\times p}$, $G \in \mathbb{R}^{n\times m}$, $H \in \mathbb{R}^{n\times n}$, and $E \in \mathbb{R}^{n\times p}$ which should be obtain. The $\hat{x}$ is an estimate of x. $\Omega \in \mathbb{R}^{n\times m}$ is the feed-forward injection map and $\nu \in \mathbb{R}^m$ is an external feed-forward compensation signal. Note that Ω is selected such as $C\Omega$ to be a nonsingular matrix. It is desired to design the observer such that $\hat{x}$ tends to x eventually.
Hence the state estimation error equation of the form

$$e_x(t) = x(t)-\hat{x}(t) = x(t)-z(t)+Ey(t) \tag{5.54}$$

is defined for the system (5.6) and observer (5.53) in terms of $x(t)$.
It is straightforward to rewritten equation (5.54) as follows:

$$\begin{aligned} \dot{e_x}(t) &= Ne_x(t)+(HA-NH-LC)x(t)+(HB-G)u(t)+ \\ &\quad HD\mu(t)+HS(g(t,u,x)-g(\hat{x},u,t))-\Omega\nu \end{aligned} \tag{5.55}$$

where

$$H = I_n - EC. \tag{5.56}$$

If the following conditions hold:

$$HD \neq 0 \tag{5.57}$$

$$HB - G = 0 \tag{5.58}$$

$$HA - NH - LC = 0 \tag{5.59}$$

then the error equation (5.55) will find the following form:

$$\dot{e_x}(t) \quad = \quad Ne_x(t) + HS(g(x,u,t) - g(\hat{x},u,t)) + HD\mu(t) - \Omega\nu \tag{5.60}$$

Equation (5.60) shows that the error, $e_x(t)$, is sensitive to both unknown input and the nonlinearity.

To design the feed-forward observer the following definitions and assumption are given:

1) Define the output error $e_y(t)$ as:

$$e_y(t) \quad = \quad y(t) - \hat{y}(t) = Ce_x(t). \tag{5.61}$$

2) Define ν as:

$$\nu = \Theta \frac{y(t) - \hat{y}(t)}{\|y(t) - \hat{y}(t)\|} = \Theta \frac{Ce_x(t)}{\|Ce_x(t)\|} \tag{5.62}$$

where $\Theta \in \mathbb{R}^{m \times m}$ is a diagonal matrix which satisfies

$$\lambda_{min}(\Theta) \geq \alpha \parallel F \parallel, \tag{5.63}$$

and $F \in \mathbb{R}^{m \times m}$ will be defined later.

3) Let P be the s.p.d. solution of the Lyapunov equation

$$NP + PN^T = -Q_f \tag{5.64}$$

Where Q_f is an arbitrary s.p.d. matrix. Hence the feed-forward injection map can

be present as

$$\Omega = P^{-1}C^T\Theta^{-1}. \tag{5.65}$$

Assumption 4

Assume that there exists a matrix $F \in \mathbb{R}^{m\times m}$, such as:

$$HD = \Omega F = P^{-1}C^T\Theta^{-1}F. \tag{5.66}$$

Then in order to design the observer (5.53) the following steps are made:

- From the Assumption 3 and condition (5.56), Matrix E will be obtain as:

  $$E = (\Omega F - D)(CD)^+. \tag{5.67}$$

  where $(CD)^+$ is a pseudo inverse of CD.

- Next by substituting matrix E into (5.56) matrix H is obtained.

- Then by substituting H into (5.58) matrix G is given.

- Assume that the pair (HA,C) is an observable pair and P_0 is the s.p.d. solution of the following ARE, then the gain matrix K is selected to make the matrix $N = HA - KC$ stable,

  $$(HA)^T P_0 + P_0(HA) - P_0C^TR^{-1}CP_0 = -Q_0 \tag{5.68}$$

  where $Q_0 \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{p\times p}$ are arbitrary s.p.d. matrices. Hence by selecting $K = P_0C^TR^{-1}$ the matrix N will be an stable matrix.

- Once K is calculated, by substituting N, K and H into (5.59) the observer gain L could be achieved of the form

  $$L = HAE + K(I_p - CE) \tag{5.69}$$

  where $I_p \in \mathbb{R}^{p\times p}$ indicates an Identity matrix of size p.

Eventually all the design matrices for constructing the observer (5.53) are obtained.

5.6 Stability analysis of the error system

In this section, the behaviour of the error system (5.55) in the fault-free case is studied. To analyze the error stability the following theorem is defined:

Theorem 2: Assume that the conditions (5.56)-(5.59) are satisfied and there is no fault in the system, then the error system (5.60) is asymptotically stable if the following matrix inequality:

$$P_1N + N^TP_1 + \varepsilon_1P_1^2 + \frac{\kappa^2}{\varepsilon_1}S^TH^THS < 0 \tag{5.70}$$

for some positive number ε_1, has a s.p.d. solution P_1.

Corollary 2: If P_1 is a solution of the following ARE

$$N^TP_1 + P_1N + \varepsilon_1P_1^2 + \frac{\kappa^2}{\varepsilon_1}\sigma_M(HS) = -Q_\Theta \tag{5.71}$$

where Q_Θ is the appropriate p.s.d matrix and $\sigma_M(HS)$ is the maximum singular values of the corresponding matrix respectively, then the error system (5.60) is asymptotically stable.
Note that for simplicity the index x will be omitted from e_x .

Proof: Consider the following Lyapunov equation

$$V = e(t)^TP_1e(t) \tag{5.72}$$

where P_1 is an s.p.d. matrix.

The time-derivative of Lyapunov equation (5.72) is

$$\begin{aligned}\dot{V} &= e(t)^TP_1\dot{e}(t) + \dot{e}(t)^TP_1e(t)\\ &= e(t)^T\left(N^TP_1 + P_1N\right)e(t) + \left(g(x,u,t) - g(\hat{x},u,t)\right)^TS^TH^TP_1e(t) + e(t)^T\\ &\quad P_1HS(g(x,u,t) - g(\hat{x},u,t)) + 2e(t)^THD\mu(t) - 2e(t)^TP_1\Omega\nu. \end{aligned}\tag{5.73}$$

Since for any matrices X, Y and any positive number $\varepsilon > 0$ the following inequality satisfies

$$\varepsilon\left(\frac{1}{\varepsilon}X - Y\right)^T\left(\frac{1}{\varepsilon}X - Y\right) \geq 0$$

then

$$X^TY + Y^TX \leq \frac{1}{\varepsilon}X^TX + \varepsilon Y^TY, \qquad \forall\varepsilon > 0$$

hence for any $\varepsilon_1 > 0$,

$$\begin{aligned} & e(t)^T P_1 HS(g(x,u,t) - g(\hat{x},u,t)) + (g(x,u,t) - g(\hat{x},u,t))^T S^T H^T P_1 e(t) \\ \leq\ & \varepsilon_1 e(t)^T P_1 P_1 e(t) + \frac{1}{\varepsilon} HSS^T H^T \|(g(x,u,t) - g(\hat{x},u,t)\|^2 \\ \leq\ & \varepsilon_1 e(t)^T P_1 P_1 e(t) + \frac{\kappa^2}{\varepsilon_1} HSS^T H^T \|x(t) - \hat{x}(t)\|^2 \\ =\ & (\varepsilon_1 P_1^2 + \frac{\kappa^2}{\varepsilon_1} HSS^T H^T)\|e(t)\|^2. \end{aligned} \tag{5.74}$$

Then by substituting (5.62), (5.65), (5.66) and (5.74) into (5.73),

$$\begin{aligned} \dot{V(t)} \leq\ & e(t)^T \left[\left(N^T P_1 + P_1 N\right) + \varepsilon_1 P_1^2 + \frac{\kappa^2}{\varepsilon_1} \sigma_M(HS) \right] e(t) + 2e(t)^T C^T \Theta^{-1} F \mu(t) \\ -\ & 2e(t)^T C^T \frac{Ce(t)}{\| Ce(t) \|} \\ \leq\ & -e(t)^T Q_\Theta e(t) + 2 \| e(t)^T C^T \| \left(\frac{\alpha_1 \| F \|}{\lambda_{min}(\Theta)} - 1\right) \\ \leq\ & -\lambda_{min}(Q_\Theta) \| e(t) \|^2 \end{aligned} \tag{5.75}$$

which implies that (5.75) is negative definite. Therefore the error system (5.60) in fault free case is asymptotically stable if there exists an s.p.d. P_1 satisfying (5.70).

5.7 Model (observer)-based intermittent fault detection

The FD consists essentially of two steps, residual generation and residual evaluation including threshold, see (Chen *et al.*, 2015; Qning *et al.*, 2014)
The purpose of the first step is to generate a signal, the residual, which is supposed to be nonzero in the presence of a fault and zero otherwise. However, the residual is almost always nonzero due to disturbances and model perturbations, even if there is no fault.

The purpose of the second step of the FD algorithm is thus to evaluate the residual and draw conclusions on the presence of a fault. This is done by comparing some function of the residual to a threshold, see (Frank, 1995; Emami-Naeini & Rock, 1988).

5.7.1 Residual generation

Consider the nonlinear system (5.6) with the intermittent fault $f_i(t)$,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + D\mu(t) + Sg(x,u,t) \\ y(t) &= Cx(t) + D_y \mu_y(t) + K_s f_{i_s}(t) \end{aligned} \tag{5.76}$$

where $K_s \in \mathbb{R}^{p \times r_s}$ is the known distribution matrix of the intermittent fault $f_{i_s} \in \mathbb{R}^{r_s}$.

For effective fault detection, the effect on sensitivity due to the disturbance in the residual vector must be small while the sensitivity due to the faults should be large. Hence, the generated residual $r(t)$ should be as sensitive as possible to the fault $f_{i_s}(t)$ and as robust as possible to unknown input (disturbance) $\mu(t)$ and sensor noise $\mu_y(t)$.
Then the fault detection filter for system (5.76) may have the following form:

$$\begin{aligned}
\dot{z}(t) &= Nz(t) + Ly(t) + Gu(t) + HSg(\hat{x}, u, t) + HD\hat{\mu}(t) + \\
&\quad HD_y\hat{\mu}_y(t) \\
\hat{x}(t) &= z(t) - Ey(t) \\
\hat{y}(t) &= C\hat{x}(t) + D_y\hat{\mu}_y \\
r(t) &= \xi(y(t) - \hat{y}(t)) = \xi(Ce_x(t) - D_yF_{11}Ce_x(t) + K_sf_{i_s}(t)) \\
&= \xi C\bar{\omega}e_x(t) + \xi K_sf_{i_s}(t)
\end{aligned} \tag{5.77}$$

where $\xi \in \mathbb{R}^{n_\xi \times p}$ is a weighting matrix which should be designed and

$$\bar{\omega} = I_n - D_yF_{11}. \tag{5.78}$$

The problem can be stated as finding ξ, such that the following aims are achieved, (Ahmadizadeh *et al.*, 2014):

- The effect of the unknown input (disturbance) signals on the residual signal are as small as possible while the effect of the fault signal is as large as possible.
- The effect of parametric uncertainties on the residual signal is as small as possible.
- The fault detection system is robust and stable in the presence of exogenous signals and uncertainties.

In other words, the main objection is to show that the residuals are away from zero when faults have occurred; however, the residual tends to zero in "no-fault" situation.

5.7.2 Residual evaluation

A common choice of evaluation signal is the following 2-norm,

$$r_{eval} = \|r\|_2 \triangleq \sqrt{\int_0^\infty |r(\tau)|^2 d\tau}. \tag{5.79}$$

Since evaluation of (5.79) cannot be easily realized, because the value of $\|r\|_2$ is not known until $t = \infty$, and it is reasonable to assume that the faults could be detected, if occurs over the finite time interval, therefore (5.79) could be modified to

$$r_{eval} = \|r(t)\|_2 \triangleq \sqrt{\int_0^t |r(\tau)|^2 d\tau} \tag{5.80}$$

where τ is the time window and it is finite, (Puig *et al.*, 2012). Benefit of using the 2-norm is that it is then straight-forward to optimize the residual generator to minimize the influence of unknown input (disturbance) $\mu(t)$. For simplicity the index 2 will be ignored from this point.
For the evaluation signal (5.80), the occupancies of faults can be alarmed if

$$\|r(t)\| > T_r \Longrightarrow A\ fault\ is\ detected \tag{5.81}$$

and

$$\|r(t)\| \leq T_r \Longrightarrow No\ fault\ is\ detected. \tag{5.82}$$

where, T_r, indicates the threshold. Hence the value of threshold gives an explicit bound for the fault-free case and thus provides a valuable guideline for robust threshold selection, (Qning *et al.*, 2014; Puig *et al.*, 2012; Wang & Wang, 2007).

Adaptive threshold:

Since with a fixed threshold it is most likely that the false alarms occur (Wu *et al.*, 2017), for instance, due to the dynamics of the system, and this can be seen as a breach in the fixed threshold where no fault exists, hence an adaptive threshold, T_r, is obtained based on the residual dynamics in fault-free case to minimize the false negative and false positive alarms.
To design the adaptive threshold for nonlinear system (5.76) and evaluation signal (5.80), the residual $r(t) = \xi C\bar{\omega}e_x(t) + \xi K_s f_{i_s}(t)$ should be redefined as follows

$$r(t) = r_e(t) + r_{f_{i_s}}(t), \tag{5.83}$$

where $r_e(t) = r(t)\mid_{\mu(t)=0,\mu_y(t)=0,f_{i_s}(t)=0}$ and $r_{f_{i_s}}(t) = r(t)\mid_{\mu(t)=0,\mu_y(t)=0}$ are the residuals due to the state errors and intermittent fault respectively.

To design T_r, it is also needed to define the residual due to the unknown input (disturbance), $r_\mu(t) = r(\acute{t})\mid_{f_{i_s}(t)=0}$ where $t = [0, t_1, \cdots, t_i, \cdots, t_n]$ and $\acute{t} = [t_i, \cdots, t_n]$, where $\acute{t}$ is the time when the system is stable.
Therefore T_r could be express as follows:

$$T_r = \sup\|r_\mu(\acute{t})\| + \|r_e(t)\| \geq 0. \tag{5.84}$$

Since the unknown input (disturbance) is bounded to a positive scalar d (see Assumption 2), then

$$\sup\|r_\mu(\acute{t})\| = \delta_d \geq 0 \tag{5.85}$$

where δ_d is a positive constant number. Hence from (5.84), $\|r_e(t)\| = T_r - \delta_d$, which results that

$$\|r_e(t)\| \leq T_r. \tag{5.86}$$

In other hand to show that T_r is the upper bound of residual $\|r(t)\|$, consider (5.83) as follows:

$$\|r(t)\| = \|r_e(t) + r_{f_i}(t)\|. \tag{5.87}$$

In faulty case $\|r_{f_i}(t)\| > \beta > 0$, where β is a positive constant. Consequently $\|r(t)\| = T_r + \beta > 0$ which conclude $\|r(t)\| > T_r$.
If there was no fault in the system, then $\|r_{f_i}(t)\| = 0$, therefore from (5.87), $\|r(t)\| = \|r_e(t)\|$ which results in $\|r(t)\| \leq T_r$.
Hence according to the obtained results the designed residuals and adaptive thresholds are able to detect intermittent faults while occurred.

5.8 NUIO-based intermittent fault detection example

To illustrate the effectiveness of the NUIO designed in this chapter, the model of a three-phase current motor is used, see (Brik & Zeitz, 1988; Nikoukhah, 1995). The state equations of this system are given by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= A_1 x_1 + A_2 x_2 + A_3 x_3 + D_1 \mu(t) \\ \dot{x}_3 &= u + A_4 x_3 + S_1 g(x,u,t) \end{aligned} \tag{5.88}$$

where $x = (x_1, x_2, x_3)^T$ is the state vector, u is the control input and A_1, A_2, A_3, A_4, D_1 and S_1 are constants (see table (5.1)).

The intermittent fault $f_{i_s}(t)$ is generated as a combination of impulses at different amplitudes which will be occurred in discrete intervals. Thus, the fault can be modelled as:

$$f_{i_s}(t) = \begin{cases} 0 \times d_y & for \quad 0 \leq t < 3s \\ 1 \times 10^{-1} \times d_y & for \quad 3s \leq t < 5s \\ 0 \times d_y & for \quad 5s \leq t < 8s \\ 2.5 \times 10^{-1} \times d_y & for \quad 8s \leq t < 12s \\ 0 \times d_y & for \quad 12s \leq t < 15s \\ 5 \times 10^{-1} \times d_y & for \quad 15s \leq t < 21s \\ 0 \times d_y & for \quad 21s \leq t < 26s \\ 7 \times 10^{-1} \times d_y & for \quad 26s \leq t < 35s \\ 0 \times d_y & for \quad 35s \leq t < 40s \\ 9 \times 10^{-1} \times d_y & for \quad 40s \leq t < 50s \end{cases} \tag{5.89}$$

where $d_y > 0$ is constant and t indicates the time. The outputs of the system are given by

$$y(t) = Cx(t) + K_s f_{i_s}(t). \tag{5.90}$$

Moreover, the unknown input (disturbance), $\mu(t)$ is an unknown function which is bounded to some positive constant, $d = 0.001$ (see Assumption 2) and the nonlinearity function is , $g(x,u,t) = \frac{1}{2}\sin(x_1)$. The numerical values of the different parameters are listed in Table (5.1).

Using the parameter values given in Table (5.1), the following state-space matrices

Parameter	Numerical value
A_1	-0.2149
A_2	-0.2703
A_3	-1.1990
A_4	-0.3222
D_1	1.00
S_1	1.9

Table 5.1: Numerical values of the system parameters

are established for the example system (5.88),

$$A = \begin{pmatrix} 0 & 1.0 & 0 \\ -0.2149 & -0.2703 & -1.1990 \\ 0 & 0 & -0.3222 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, S = \begin{pmatrix} 0 \\ 0 \\ 1.90 \end{pmatrix}, K_i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

It is clear that for this system assumption 1 is satisfied and $CD = 0$. As this assumption is satisfied to design the NUIO for the illustrated system the following steps should be carried out.

- First, select $D_1 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T$ which satisfies the rank condition

  $$rank(CD_1) = rank(D_1) = 1$$

  and makes $HD_1 = 0$ and $CD_1 \neq 0$.

- Second, substitute D_1 into equation (5.20), matrix E will be then driven as:

  $$E = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

- Then substitute E into (5.12) to establish matrix H

  $$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

- Next from $N = HA - KC$ and ARE (5.21), matrix N and control gain K are ob-

tained as follows:

$$N = \begin{pmatrix} -0.3466 & 1.0 & 0.0662 \\ -0.3905 & -0.2703 & -0.6922 \\ 0.0662 & 0 & -1.1878 \end{pmatrix},$$

$$K = \begin{pmatrix} 0.3466 & -0.0662 \\ 0.1756 & -0.5068 \\ -0.0662 & 1.1878 \end{pmatrix}.$$

- Finally by substituting N and K into (5.22), the observer gain L is given

$$L = \begin{pmatrix} 0.3466 & 0 \\ 0.1756 & -1.1990 \\ -0.0662 & 0 \end{pmatrix}.$$

The obtained parameters satisfy the conditions (5.27), (5.33) and (5.39), and guarantee the existence of the designed NUIO for the system under investigation.
Matrices $\bar{M}$ and $\bar{G}$ are also obtained from the equations (5.42) and (5.43)

$$\bar{M} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1.010 & -0.0200 \\ 0 & 0 & 1 \end{pmatrix},$$

$$(\bar{M})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.9901 & 0.0198 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\bar{G} = \begin{pmatrix} -0.3466 & 1.000 & 0.0762 \\ -0.3927 & -0.2730 & -0.7142 \\ 0.0662 & 0 & -1.1878 \end{pmatrix}.$$

Since $\det\bar{M} \neq 0$, therefore $\bar{M}$ is invertible. Hence, in order to make $\bar{M}^{-1}\bar{G}$ Hurwitz the following design matrices are selected

$$A_d = \begin{pmatrix} -0.01 & 0 & 0 \\ 0 & -0.01 & 0 \\ 0 & 0 & -0.01 \end{pmatrix},$$

$$A_i = \begin{pmatrix} -0.01 & 0 & 0 \\ 0 & -0.01 & 0 \\ 0 & 0 & -0.01 \end{pmatrix},$$

$$A_g = \begin{pmatrix} -0.02 & 0 & 0 \\ 0 & -0.02 & 0 \\ 0 & 0 & -0.02 \end{pmatrix}.$$

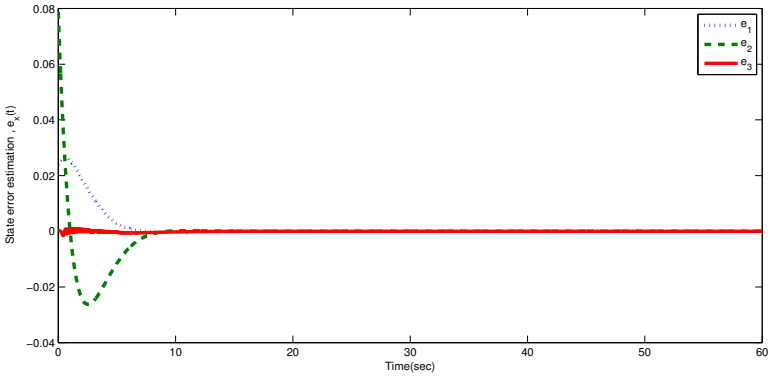

Figure 5.1: States error responses ($e = x - \hat{x}$).

Hurwitz matrix: An square matrix $\bar{M}^{-1}\bar{G}$ is called Hurwitz or stable matrix when all of its eigenvalues have strictly negative real parts $\|Re(\bar{M}^{-1}\bar{G})\| < 0$. Hence, By selecting these design matrices, $\bar{M}^{-1}\bar{G}$ becomes

$$\bar{M}^{-1}\bar{G} = \begin{pmatrix} -0.3466 & 1 & 0.0762 \\ -0.3875 & -0.2703 & -0.7306 \\ 0.0662 & 0 & -1.1878 \end{pmatrix}.$$

where every eigenvalue of $\bar{M}^{-1}\bar{G}$ has strictly negative real part,

$$\begin{aligned} \lambda_{e1} &= -0.2869 + 0.6482i \\ \lambda_{e2} &= -0.2869 - 0.6482i \\ \lambda_{e3} &= -1.2308 \end{aligned} \tag{5.91}$$

λ_e presents the eigenvalues of $\bar{M}^{-1}\bar{G}$ and the negative real part of these eigenvalues demonstrate that $\bar{M}^{-1}\bar{G}$ is Hurwitz.

To see the capability of the proposed NUIO in detecting intermittent fault, simulation results are presented in Figures (5.1)-(5.5). From Figure (5.1), which shows the states error's responses, it can be seen that the observers perform as expected and the state estimation errors do tend to zero asymptotically. It also demonstrates that the proposed design approach minimizes the effects of the unknown inputs (disturbances) to the state estimation errors and gives a straightforward way to design a robust observer for intermittent fault detection where the bounded unknown inputs (disturbances) exist.

Moreover, figures (5.2)-(5.5) show that the intermittent fault has been detected using the designed adaptive threshold for both choice of outputs.

When an adaptive threshold is designed, as shown in Figures (5.2)-(5.5), the system dynamics do not breach the threshold, (Sedighi *et al.*, 2013). Moreover, the proposed adaptive threshold is insensitive to faults of specific amplitude. The adaptive threshold approach provides the capability to ignore small intermittent disturbances that manifest

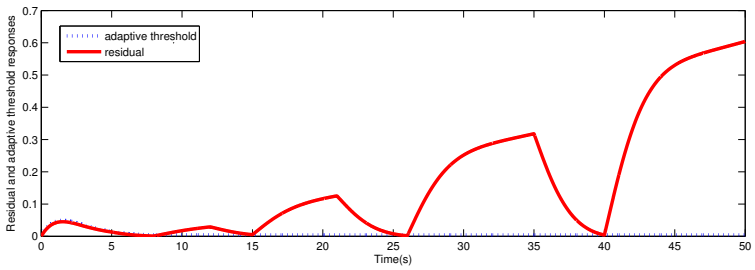

Figure 5.2: Residual and adaptive threshold responses to detect the intermittent faults (first choice of output, y_1).

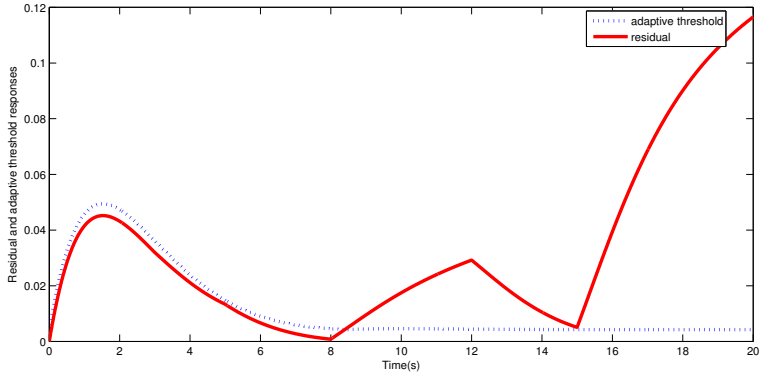

Figure 5.3: Residual and adaptive threshold responses to detect the intermittent faults (first choice of output, y_1), zoom on the first few seconds of Figure 5.2.

as system noise and do not have a serious impact on the system performance.

5.9 Feed-forward observer-based intermittent fault detection example

To demonstrate the efficiency of the outcomes achieved in sections $5.5 - 5.7$, consider the dynamic characteristics of a car suspension system. The system is treated as a Mass-Spring-Damper (MSD) system where n masses, springs, and dampers are linked in series (Figliola & Beasley, 2014; Wang *et al.*, 2007; Patton *et al.*, 1989).
Since the study of a full-suspension model is truly complex including all four suspensions (tire) systems working individually, then the quarter-car suspension system is expressed in the three levels of complexity.

The one-degree of freedom model shown in Figure (5.6a) considers displacement, r_1, of the sprung mass, m_1, of the vehicle and the primary suspension stiffness, k_1, and damping, c_1, only. Here the unsprang mass (mass of the wheels and other components such as lower control arms) and the mass of the tires are not considered.

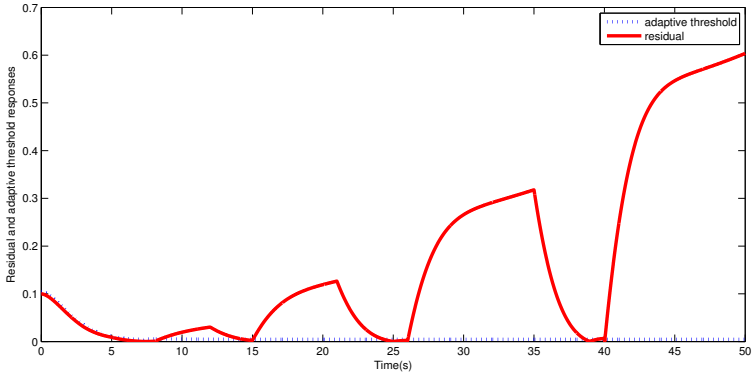

Figure 5.4: Residual and adaptive threshold responses to detect the intermittent faults (second choice of output, y_2).

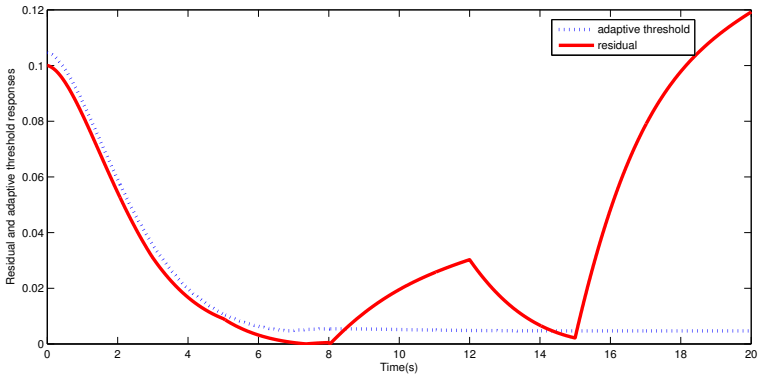

Figure 5.5: Residual and adaptive threshold responses to detect the intermittent faults (second choice of output, y_2), zoom on the first few seconds of Figure 5.4.

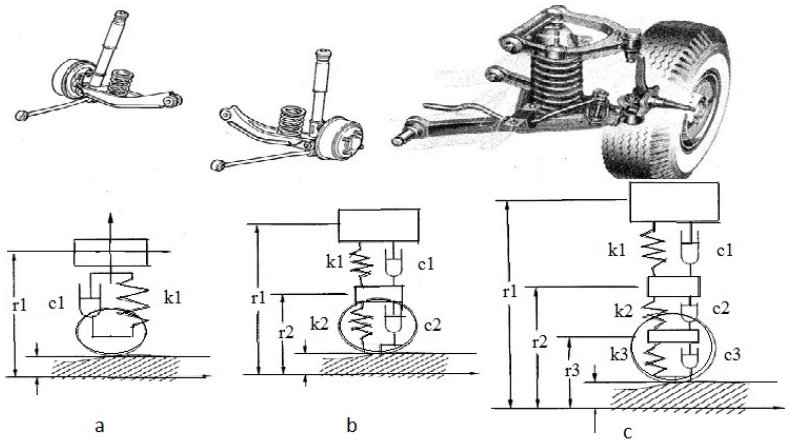

Figure 5.6: The model vehicle suspension system.

The two degrees of freedom model shown in Figure (5.6b) accounts for the dynamics of the unsprang mass and introduces the second equation of motion and degrees of freedom for the displacement, r_2, of the unsprang mass, m_2, springs and dampers with k_2 and c_2. In this model, the tires are massless.

A three-degree of freedom, the model is shown in Figure (5.6c) where the dynamics of the tires are added to the analysis by treating them as a mass-spring-damper (Genta, 1997; Gillespie, 1992).

A MSD model is normally designed by a series of differential equations. The model consists of a finite number of masses, springs, and dampers on a line. It is assumed that n masses, springs and dampers are connected serially.

Note that the system which will be designed in this chapter may be extended. Hence, the designer can choose any number of masses, springs and dampers to build the desired system with more complexity.

In Figure (5.7), two springs, dampers and masses were linked together in series (Wang *et al.*, 2007) where x_1 and x_2 indicate the position and velocity of the first mass and x_3 and x_4 indicate the position and velocity of the second mass, respectively. A_{nl} is a nonsingular damping device whose damping force is $F_{A_{nl}} = C_{nl}\text{sign}(x_2)\ln(1+\mid x_2 \mid)$, with $C_{nl} \geq 0$.

An arbitrary and unknown force, w, is enforced on the second mass. The known input forces, u_1 and u_2 are applied to both masses 1 and 2, subsequently. The state variables x_1 and x_4 are measurable, hence to estimate the state variables x_2 and x_3 the observers are designed.
The state equations of the model are given such as,

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1+k_2}{m_1} & -\frac{b_1+b_2}{m_1} & \frac{k_2}{m_1} & \frac{b_2}{m_1} \\ 0 & 0 & 0 & 1 \\ \frac{k_2}{m_2} & \frac{b_2}{m_2} & -\frac{k_2}{m_2} & -\frac{b_2}{m_2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

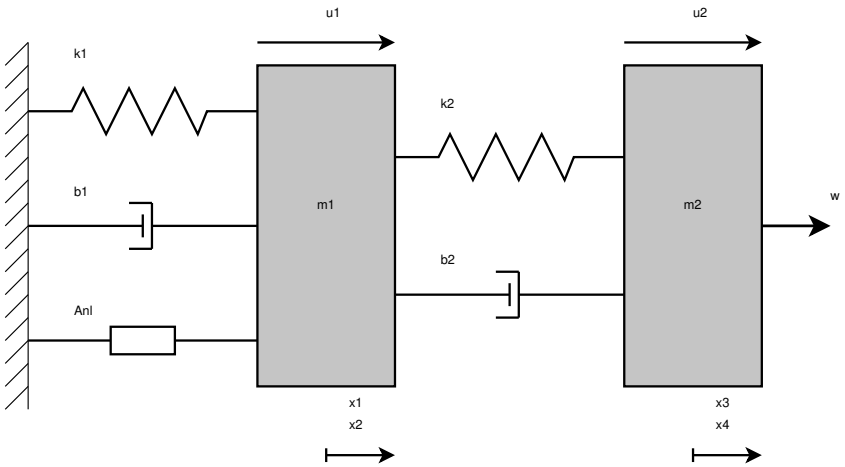

Figure 5.7: The mass-spring-damper system.

,

$$+\begin{pmatrix} 0 \\ -\frac{1}{m_1} \\ 0 \\ 0 \end{pmatrix} \rho(x_2) + \begin{pmatrix} 0 & 0 \\ \frac{1}{m_1} & 0 \\ 0 & 0 \\ 0 & \frac{1}{m_2} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{m_2} \end{pmatrix} w, \tag{5.92}$$

$$\rho(x_2) = C_{nl}\,\mathrm{sign}(x_2)\ln(1+|x_2|). \tag{5.93}$$

The output of the model is presented as, $y = Cx + \eta_y$, while $C \in \Re^{n\times n}$. And η_y indicates an additive offset (intermittent sensor fault/output error) on each output.

5.9.1 Intermittent fault detection

Collapsing suspension because of coil spring failure sounds to be a growing issue, created by a combination of latter harsh winter conditions and weight-saving designs. A plastic layer is coating the coil springs while they are built to decrease the risk of corrosion. Over time, contact between coils when the spring is compressed frequently in service may cause damage to this coating. Regularly the failure of the coil spring appears to be created by corrosion, is accelerated by salt enforced to the roads in winter. Consequently, the electrolytic action between the salt solution, created by road salting, and the iron in the spring produces free hydrogen atoms which enter the steel and can cause microscopic cracking. Cracks propagate and combine, eventually leading to the spring failure, `http://www.theaa.com`.
Cracks and corrosion both can be classified as intermittent faults. Assume that at each failure, the length of spring will change suddenly. Hence, a fault in the position i, is defined as a change in the length of the i-th spring, $L_f = (L_0 + f_{i_s} L_0)$, while in all other parts of the model the length of the springs will remain as $L_f = L_0$, where L_0 indicates the initial length of the spring.

The fault, $f_{i_s}(t)$ is a time varying of the form $f_{i_s}(t) = dd_i y_{n_c}(t)$, where the constants, dd_i, for $i = [1, \cdots, 4]$, indicate the maximum fault amplitudes. The selected output is shown by y_{n_c} where $n_c = 1, \cdots, n$ and the time is defined by t.
Consequently, for the model under the investigation the intermittent fault, $f_{i_s}(t)$, is presented as,

$$f_{i_s}(t) = \begin{cases} 0 & for \quad 0 \leq t < 5s \\ dd_1 & for \quad 5s \leq t < 7s \\ 0 & for \quad 7s \leq t < 11s \\ dd_2 & for \quad 11s \leq t < 14s \\ 0 & for \quad 14s \leq t < 18s \\ dd_3 & for \quad 18s \leq t < 25s \\ 0 & for \quad 25s \leq t < 28s \\ dd_4 & for \quad 28s \leq t < 40s \end{cases} \tag{5.94}$$

with constants, $dd_1 = 0.0025$, $dd_2 = 0.01$, $dd_3 = 0.15$ and $dd_4 = 0.25$.

5.9.2 Simulation results

Assume that the model parameters have got the following values, $m_1 = 5kg$, $m_2 = 1kg$, $k_1 = 30\frac{N}{m}$, $k_2 = 10\frac{N}{m}$, $b_1 = 4\frac{Ns}{m}$, $b_2 = 2\frac{Ns}{m}$, $C_{nl} = 5N$ and $w(t) = 0.04\sin(t) + 2N$, then to design the appropriate observer (5.53), the following matrices are obtained,

$$E = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -1 \end{pmatrix}, \; G = \begin{pmatrix} 0 & 0 \\ 0.2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \; H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$L = \begin{pmatrix} -0.3501 & 0 \\ -4.5838 & 0.4000 \\ 0.2936 & 1.000 \\ -0.8632 & 0 \end{pmatrix}, \; K = \begin{pmatrix} -0.3501 & 0.7234 \\ -4.5838 & -0.2351 \\ 0.2936 & 1.4022 \\ -0.8632 & 0.500 \end{pmatrix},$$

$$N = \begin{pmatrix} 0 & 1 & 0.3501 & -0.7234 \\ -8.000 & -1.200 & 6.5836 & 0.6351 \\ 0 & 0 & -0.2936 & -0.4022 \\ 0 & 0 & 0.8632 & -0.5000 \end{pmatrix}.$$

From (5.71) the s.p.d. matrices Q_Θ and P_1 are obtained as follows:

$$Q_\Theta = \begin{pmatrix} 4.7840 & -16.1760 & 1.6550 & 2.3160 \\ -16.1760 & 69.7660 & -6.2630 & -9.0150 \\ 1.6550 & -6.2630 & 0.6920 & 0.9280 \\ 2.3160 & -9.0150 & 0.9280 & 1.3031 \end{pmatrix},$$

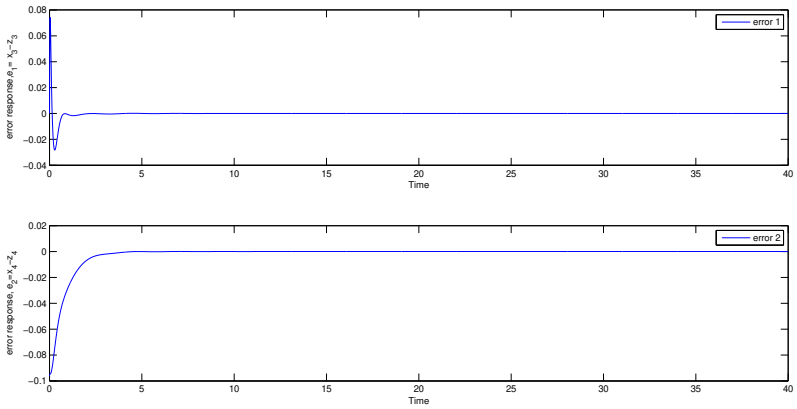

Figure 5.8: The error estimation responses in presence of unknown inputs.

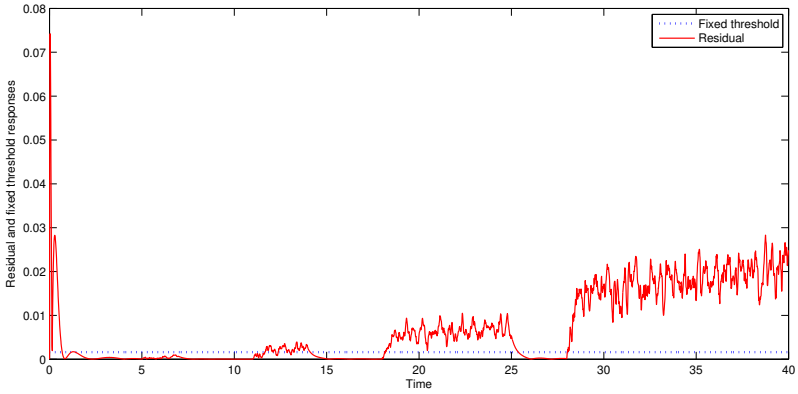

Figure 5.9: The residual and fixed threshold responses in presence of the intermittent fault.

$$P_1 = \begin{pmatrix} 3.0828 & -2.7134 & -0.0139 & -0.4510 \\ -2.7134 & 37.7228 & -1.6128 & -1.1082 \\ -0.0139 & -1.6128 & 0.4494 & 0.5322 \\ -0.4510 & -1.1085 & 0.5322 & 2.2219 \end{pmatrix}.$$

Select $\varepsilon_1 = 1 \times 10^{-5}$, then $\lambda \min(Q_\Theta) = 0.0125$, which satisfies equation (5.75) to be definite negative. Hence, the error stability of the MSD model is also guaranteed.

Figure (5.8) shows the behaviour of the state errors and demonstrates that the errors between the actual and estimated states are stable and converge to zero asymptotically even though uncertainties within the system exist. Figure (5.8) also shows that the designed observer satisfies the stability of the error regardless of any bounded uncertainties in the absence of faults.

Figures (5.9) and (5.10) show the residual and fixed threshold responses when the system is influenced by unknown input with the known bound. Figure (5.10) shows that with a fixed threshold false alarm can occur due to the dynamics of the system. This can

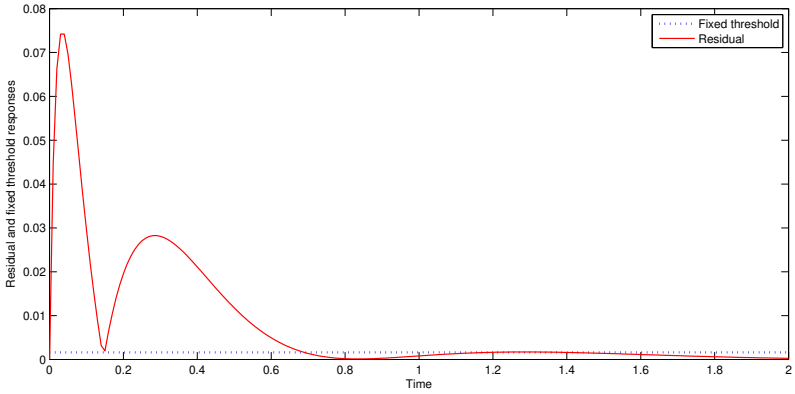

Figure 5.10: The residual and fixed threshold responses in presence of the intermittent fault (first few seconds).

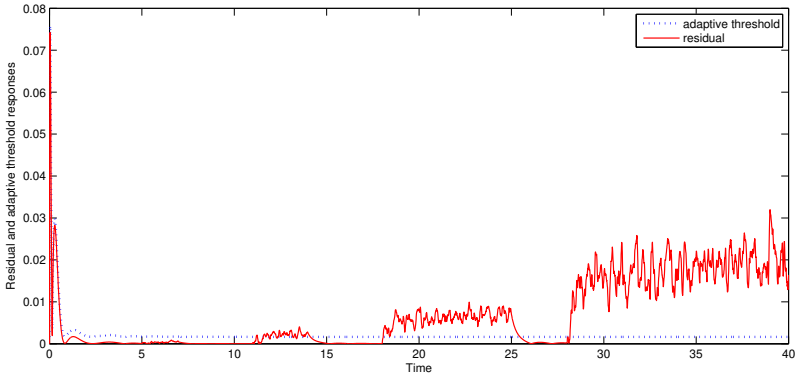

Figure 5.11: The residual and adaptive threshold responses in presence of the intermittent fault.

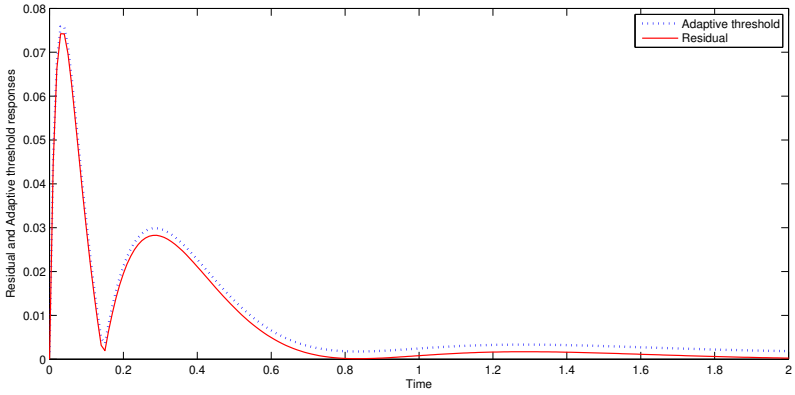

Figure 5.12: The residual and adaptive threshold responses in presence of the intermittent fault (first few seconds).

be seen as a breach in the fixed threshold at the beginning of the operation of the system where no fault exists.
In this case, faults can be detected more precisely if an adaptive threshold is designed. When an adaptive threshold is designed, as shown in Figures (5.11) and (5.12) the same system dynamics do not breach the threshold. The adaptive threshold approach, therefore, provides the capability to ignore small intermittent disturbances that manifest as system noise and do not have a serious impact on the system operation.
The simulation results also show that the proposed design approach was minimizing the effects of any uncertainties and has presented a sufficient method to design a robust observer for intermittent fault detection where the bounded disturbances exist.

5.10 Conclusions

In the first part of this chapter, a novel robust NUIO for a class of nonlinear systems with bounded unmatched unknown inputs (disturbances) was proposed. Furthermore, the fault detection filter for this class of nonlinear systems with the adaptive threshold design was also provided. The results show that the designed error dynamics are stable. In this method, sufficient design parameters (D_1, G_i, M_i, and F_i, $i = 1, \cdots, n$), where n is the number of states, and the nonunique design matrices A_g, A_d and A_i provide extra degrees of freedom to design NUIO.
The advantages of the presented NUIO are summarized as follows:

- First, The restriction of the NUIO rank condition has been ignored without losing the design convenience by introducing auxiliary disturbance's distribution matrix D_1.
- Second, there is no need to consider only the Lipschitz nonlinearity. It could be applied to the wider class of nonlinear systems with any form of general nonlinearity.
- Finally, the sensor fault detection method by designing the adaptive threshold is developed.

However, the main advantage of the proposed method is the possibility to convert the nonlinear system to the linear system by applying the linear equality mentioned in the design procedure, and thus makes the difficult NUIO design problem an easy task for the considered class of nonlinear systems.
The effectiveness of the techniques is illustrated with the help of a numerical example. The simulation results show that the designed NUIO can indeed make the state estimation errors to asymptotically converge to zero regardless of the bounded unknown inputs (disturbances).
Through this work, if the existing conditions of NUIO were satisfied, the detection of the intermittent faults using adaptive threshold can be achieved for descriptor systems.

Moreover, in the second part of this chapter, a robust nonlinear feed-forward observer has been designed for a class of nonlinear systems whose nonlinear function satisfies Lipschitz condition, and the unknown input term is bounded. In this approach, a design

matrix has been proposed to provide extra degrees of freedom to the designer to develop the residual.
The main advantage of the proposed approach is the intermittent fault detection by designing a residual and an appropriate adaptive threshold while the designed threshold should be highly sensitive to the intermittent faults only.
Eventually, the effectiveness of the technique is shown by the help of a numerical example. The simulation results also demonstrate that the generated residual and adaptive threshold can indeed detect the intermittent faults regardless of the bounded disturbances/unknown inputs.

5.10.1 Limitations

Although, there are some limitations for this work such as,

- the simulation results demonstrate that the performance of the model (observer)-based fault detection techniques decrease significantly as system complexity increases (see Appendix E).
- When the intermittent faults are very small, as small as disturbance, hence they are almost impossible to detect and if the threshold has been designed to be that sensitive, then there is a possibility to detect the noises and disturbances as well,
- If the intermittent fault appears for a very short period and then goes into the rest period (deactivate period) then there is a possibility that the designed adaptive threshold was not able to detect the fault or may detect it with some delays,

But, the proposed method could make the difficult intermittent fault detection an easier task for the considered class of nonlinear systems.

In the next chapter, one of these fault detection methods will be applied to detect intermittent fault in the experimental fuel rig example system under consideration.

Chapter 6

Model (Observer)-Based Intermittent Fault Detection Application

6.1 Introduction

In this chapter, the main objective is to discuss the application of observer-based intermittent fault detection technique for the experimental fuel rig system.
Complete design sequences for the intermittent fault detection of actuators, components, input and output sensors of the fuel rig described in Chapter 4, are developed in this chapter.
The approach to be used here is the nonlinear unknown input (NUI) observer-based FDI methodology (described in deep in the previous chapter) to detect any system, actuator or sensor intermittent fault.

To validate the results of this approach, the collected experimental data from the fuel rig, described in Chapter 4 will be used.

This chapter is organized as follows: Section 6.2 the mathematical modelling of the fuel rig system is presented. In Section 6.3 the intermittent fault which has been injected to the fuel rig system artificially. and the design procedure to detect it, is proposed. Finally, the simulations results along with the discussions and conclusions are presented in Sections 6.4 and 6.5 respectively.

6.2 System modelling

As explained in Chapter 4, pipes' length and diameter, pump characteristics, loss coefficient versus valve opening characteristics, shut-off valve pressure drop when fully opened and tank's capacity has been identified within the design phase by carrying out various scenarios in a controlled simulation environment. Volumetric flow rates in the mainline and pressure rates at five different locations were also calculated using the physical model.

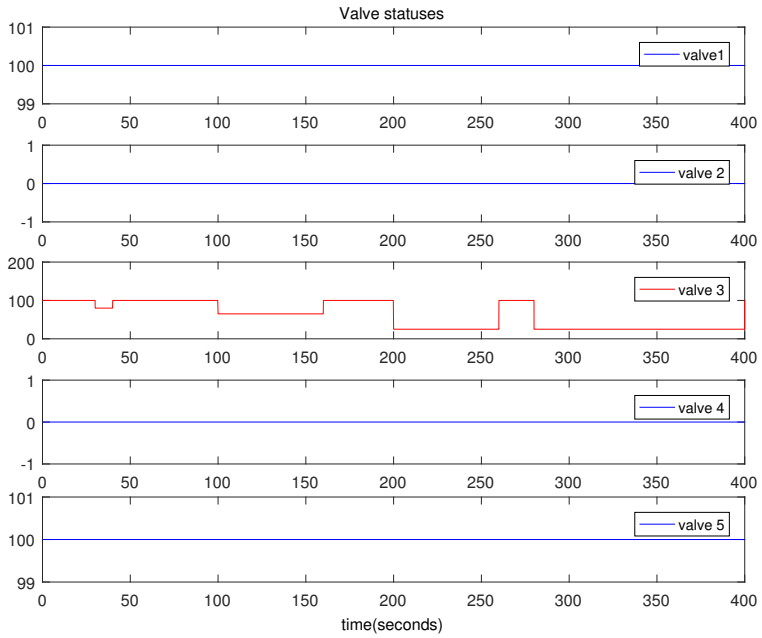

Figure 6.1: The statuses of all valves in the fuel rig.

For the healthy state of the fuel system (Figure 4.3), the DPVs were set as follows (Figure 6.1):

- DPV1 - fully open,
- DPV2 - fully closed,
- DPV3 - fully open (the faulty valve),
- DPV4 - fully closed
- DPV5 - fully open.

Pressure sensors data were recorded for 360 seconds to have a good estimation. The pump rotational speed was set at 400rpm (Figure 6.2) and the feedback loop of the pump control unit was active, so the pump speed was constant for the entire testing session.

Moreover, there was no flow into the system, $Q_{in} = 0$ and it is assumed that the fuel temperature is constant during the operations. Note that the fluid dynamic phenomena connected to flows through pipes is neglected. On the rig, the main tank supply fluid through a pump with an array of the valves connected to the pump.

The modelling will focus on the fluidic side of the rig and the approach used here, is mechanistic/physical modelling based-on the hardware of the fuel rig and the fluid properties.

Each element in this system is modelled as a subsystem and the overall model consists of all such models to represent the overall system. It should be noted that while some of the pipes depicted are very short they are included in the model of the system to make the overall model equations solvable.

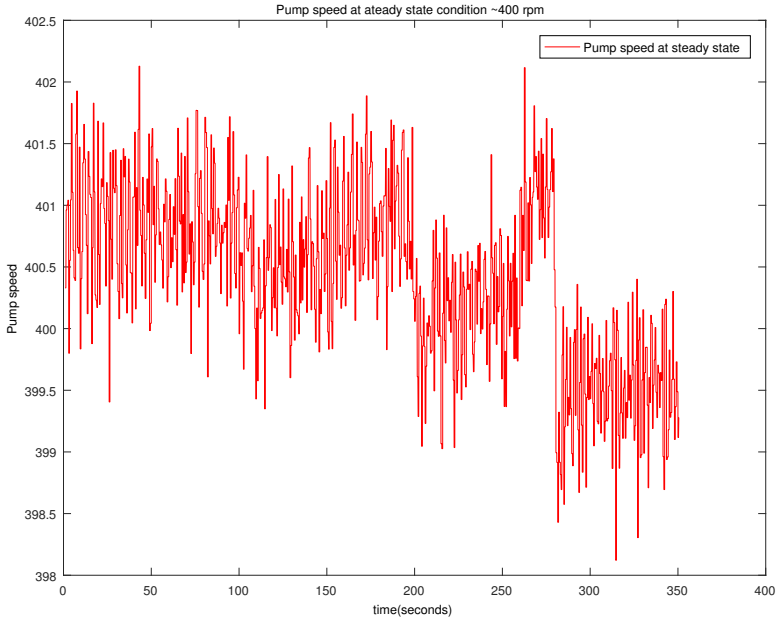

Figure 6.2: Gear pump speed.

6.2.1 Mathematical modelling of the physical functions

Tank model

The overall system process is a non-self-regulator because in this process a positive displacement gear pump is used to drain the water from the tank and the tank along with the discharge valves cannot maintain the process.

The mathematical equations provide a pressure output for a given height which in turn is dependent on the flow in and flow out of the tank (Young, 2018; Meenatchisundaram, 2015; Marshall, 1978). A fluid capacitor (tank) has increasing fluid storage with increasing pressure, and is defined by the volume of the water in the tank which is defined by:

$$V = A_t h \tag{6.1}$$

Where V indicates the volume of the liquid, A_t and h indicate the cross-sectional and hight of the liquid in the tank respectively. Moreover, $\frac{dV}{dt}$ shows the rate of change of liquid volume in the tank. Hence,

$$\frac{dV}{dt} = Q_{in} - Q_{out} \tag{6.2}$$

Then from (6.1) and (6.2) the following equation is obtained:

$$h(t) = \frac{\int (Q_{in} - Q_{out})dt}{A_t} \tag{6.3}$$

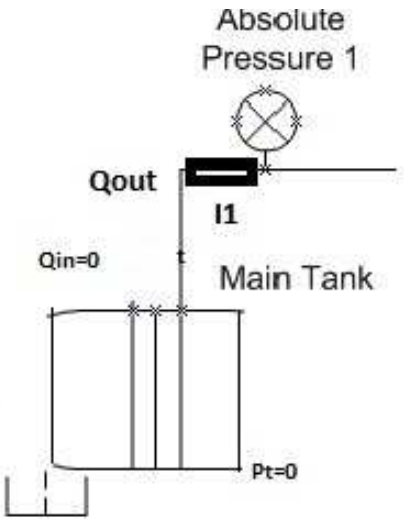

Figure 6.3: The Tank model.

The height and pressure of the fluid in the tank under investigation are then defined by equations (6.3) and (6.4).

$$P_t = \rho g h(t) \tag{6.4}$$

Substitute equation (6.3) into (6.4) where $Q_{in} = 0$, will give:

$$\dot{P}_t = C_t Q_{out} \tag{6.5}$$

With $C_t = -\frac{\rho g}{A_t}$. Q, A_t, h, ρ, P and g present volumetric flow $(\frac{m^3}{s})$, tank cross-section (m^2), height (m), density $(\frac{Kg}{m^3})$, pressure (Pa) and gravity $(\frac{m}{s^2})$ respectively (see figure 6.3).

Valve model

The valve equation below gives the volumetric flow rate from the valve for a given pressure differential across the valve (Marom *et al.*, 2012).

$$Q(t) = C_v A_v \sqrt{\frac{2}{\rho}} \Delta P \tag{6.6}$$

where Q, ρ, ΔP, C_v and A_v present volumetric flow $(\frac{m^3}{s})$, density $(\frac{Kg}{m^3})$, pressure difference (Pa), valve conductance (m^2) and proportional valve opening respectively. The equation (6.6) for fast opening valves is nonlinear and could be rewritten as

$$\Delta P = RQ^2 \tag{6.7}$$

where $R = \frac{\rho}{2C_v^2 A_v^2}$.

Pump model

The pump, motor and gearbox are represented in the model based on results found by practical experiment (Civelek, 2006). The ramp of pump time is equal to valve transition time. It means that the pump will reach the maximum voltage when the valve reach it's state. The pump is also required to validate the pipe and valve model as it is the only physically measurable input signal available to drive the pipe and valve subsystem models.
For a pump with positive displacement,

$$\begin{aligned} Power_{in} &= \tau\omega \\ Power_{loss} &= f(friction,\ viscous,\ effects,\ \cdots) \\ Power_{out} &= \Delta P \times Q. \end{aligned} \tag{6.8}$$

where τ and ω represent nominal displacement and rotational (shaft) speed respectively. $\tau\omega$ is considered as input speed of the pump in (rpm).
Moreover, equations (6.8), could be represented as follows

$$Power_{out} = \eta_m Power_{in} \tag{6.9}$$

where η_m is the pump volumetric efficient and typically gear pumps have efficiencies around 85% . Hence the following equation is obtained for pressure difference around the gear pump:

$$\Delta P = \frac{\eta_m \tau\omega}{Q} \tag{6.10}$$

Pipe model

The defining equation for the pipe subsystem model is based on the compressibility of fluid in the system due to the pressure acting upon it (Figure 6.4) and is shown in following equations, (Analysis, 2004).

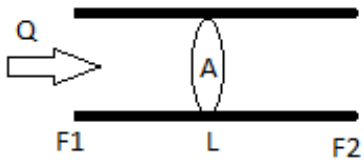

Figure 6.4: The pipe model.

Consider an incompressible fluid element flowing down a pipe with constant cross

sectional area A based on Newton's motion law:

$$\begin{aligned} F &= F_1 - F_2 = (P_1 - P_2)A \\ &= m\frac{dV}{dt} = AL\rho\frac{dV}{dt} \end{aligned} \tag{6.11}$$

$$P_1 - P_2 = I\frac{dQ}{dt} \tag{6.12}$$

This equation shows that the net force, (F),causes the mass, (m), to accelerate, (a), where

$$F = ma = PA \tag{6.13}$$

is the Newton's motion law and $I = \frac{\rho L}{A}$. L and A are the length and cross-section of the corresponding pipe respectively, $Q = AV$, $m = \rho LA$ P, and $V_0 = A \times \Sigma L$ present volumetric flow $(\frac{m^3}{s})$, mass, pressure (Pa) and original volume of pipe (m^3) respectively.

6.2.2 Overall system model

When constructing the overall model presented in Figure (6.5) the subsystems are simply parameterized and connected to represent the complete system. For parameterizations of the component models, information was taken from measurement, data-sheets and based on an experiment.

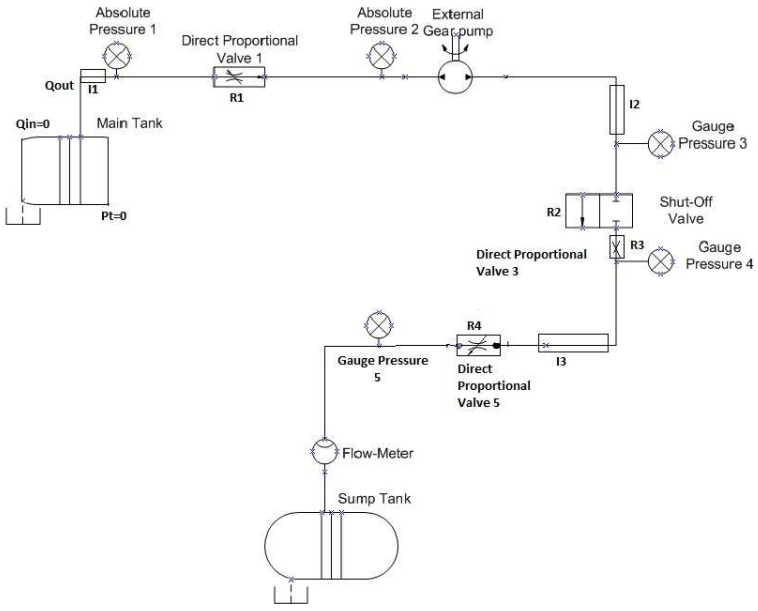

Figure 6.5: The fuel rig system demonstrator. This is a transparent engine-mode version of Figure (4.3) in Chapter 4 when there is no return of fuel/fluid to the main tank.

Figure (6.5) shows that there is only one capacitor in the system which makes the fuel rig a Single-Input Multi-Output (SIMO) system. Hence the system is presented with the second-order differential equation. Similarly, the state space equations of this

system replace the second-order differential equation with a single first-order matrix differential equation (equation 6.23).
To obtain the state-space equations of the fuel rig presented in Figure (6.5), the system is divided into subsystems and the equations for each subsystem is defined.

The total pressure change in different parts of the fuel rig could be presented as,

$$\Delta P = P_{start} - P_{end} \tag{6.14}$$

where P_{start} presents the starting pressure of the system and P_{end} presents the final pressure. So, to estimate the pressure of the fifth sensor, the starting point could be the main tank and the final point is the pressure sensor 5 at the end of the fuel rig, hence, equation (6.14) is rewritten as:

$$\Delta P = P_1 - P_5. \tag{6.15}$$

Similarly, to estimate the pressure sensor 1, the starting point can be started from valve 1 and the final point is the pressure sensor 1. The same method is repeated to estimate all of the five pressure sensors.

To obtain (6.15), the system is divided into subsystems where the pressure difference for each subsystem is presented as follows:

- The pressure difference for the pipe 1 is:

$$\Delta P_1 = I_1 \frac{dQ_t}{dt} \tag{6.16}$$

- The pressure difference for the valve 1 is:

$$\Delta P_2 = R_1 Q_t \tag{6.17}$$

- The pressure difference for the pipe 2 and the pump is:

$$\Delta P_3 = I_2 \frac{dQ_t}{dt} + K_u U t \tag{6.18}$$

  where $K_u = \frac{\eta_m}{Q_p}$. Q_p represents the flow through the cross section of the inlet port of a pump and $U = \tau\omega$ is the input to the system. Note that the gear pump is a positive displacement. They will produce the same flow at a given speed (rpm) no matter the discharge pressure because they are constant machines and Q_p is a constant value.

- The pressure difference for the valve 2 and the valve 3 is:

$$\Delta P_4 = (R_2 + R_3) Q_t \tag{6.19}$$

- The pressure difference for the valve 4 and the pipe 3 is:

$$\Delta P_5 = R_4 Q_t + I_3 \frac{dQ_t}{dt}, \tag{6.20}$$

- Consequently, equation (6.14) could be represented as

$$\begin{aligned} \Delta P &= \Delta P_1 + \Delta P_2 + \Delta P_3 + \Delta P_4 + \Delta P_5 \\ &\quad (I_1 + I_2 + I_3) C_t \frac{d^2 P}{dt^2} + (R_1 + R_2 + R_3 + R_4) C_t (\frac{dP}{dt}) + P \\ &= K_u U(t). \end{aligned} \tag{6.21}$$

Note that from the equation (6.21) and afterwards ΔP is replaced by P for simplicity. Finally, the related state-space equations could have taken the following form,

$$\begin{aligned} x_1 &= P \\ x_2 &= \dot{x_1} = \frac{dP}{dt} \\ \dot{x_2} &= \frac{d^2 P}{dt^2} \\ &= -\frac{(R_1 + R_2 + R_3 + R_4)}{(I_1 + I_2 + I_3)} x_2 - \frac{1}{(I_1 + I_2 + I_3) C_t} x_1 + \frac{K_u}{(I_1 + I_2 + I_3) C_t} \\ &\quad U(t). \end{aligned} \tag{6.22}$$

Consequently equation (6.22) could be rewritten in matrix form as follows,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + D\mu(t) + Sg(x,u,t) + K_s f_{i_s}(t) \\ y(t) &= Cx(t) \end{aligned} \tag{6.23}$$

where $D \in \mathbb{R}^{n \times q}$ is referred to the known distribution matrix of the unknown input (uncertainty), and $\mu(t) \in \mathbb{R}^q$ is an unknown vector but bounded to a positive constant d,

$$\|\mu(t)\| \leq d, \tag{6.24}$$

which describes the unknown input and/or any kind of modelling uncertainty such as random noise, time-varying term, modelling errors, and parameter variation, etc.,.

The value of the upper bounds of the unknown inputs for each pressure sensor are presented in Table (6.1) (Niculita *et al.*, 2013).

The state-space matrices are then presented as:

$$\dot{x} = \begin{pmatrix} \dot{x_1} \\ \dot{x_2} \end{pmatrix},$$

Pressure sensor (PS)	Upper bound of unknown input (d)
PS_1	0.0105
PS_2	0.0099
PS_3	0.4463
PS_4	0.0949
PS_5	0.0499

Table 6.1: The upper bound of unknown input (disturbance) for each pressure sensor (Niculita *et al.*, 2013)

$$A = \begin{pmatrix} 0 & 1 \\ -\frac{1}{(I_1+I_2+I_3)C_t} & -\frac{(R_1+R_2+R_3+R_4)}{(I_1+I_2+I_3)} \end{pmatrix} \tag{6.25}$$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ \frac{K_u}{(I_1+I_2+I_3)C_t} \end{pmatrix}$$

The Reynold number, R_e, of the fuel in the system under investigation is less than 2000, and the fuel is laminar.

Reynold number: Reynold number is a dimensional number expressing the ratio between the inertia and the viscous forces (Connor, 2020):

$$Re = \frac{\rho V d}{\mu} \tag{6.26}$$

where ρ, V,d, and μ indicate the density, mean velocity of the fluid, a characteristic linear dimension, (travelled length of the fluid), and dynamic viscosity of the fluid respectively.

If $Re < 2000$ then the flow is laminar, if $Re > 4000$ then the flow is turbulent and if the $2000 < Re < 4000$ then the flow is transition flow.

In the system under investigation, the pipe diameters is, $d = 0.4m$, the friction factor for water is 0.2, $\rho = 997\frac{Kg}{m^3}$, the the dynamic viscosity of the water is $\mu = 1.002\frac{}{NS}m^2$ (Fowler, 2007) and the tank surface pressure is $1.01325bar$ (Niculita *et al.*, 2013). Moreover, $V = \frac{Q}{A}$, where $Q = 0.784\frac{l}{min} = 0.000013\frac{m^3}{s}$ is the flow meter through the pipe and A is the pipe cross sectional area (m^2).

by substituting these values in the above equation, the Re number for the water in the system under investigation is obtained: $Re = 4.14 < 2000$ which shows the flow is laminar.

This indicates that the system acts linearly. However the fast opening of the DPVs causing nonlinear responses in the system.

The nonlinear characteristic of a DPV is modelled and explained in (Zhu & Jin, 2016). The nonlinear factors in DPV are the orifice area and the solenoid forces. Although the solenoids used in the DPVs are termed 'proportional', but the solenoid forces are not completely linear with coil current and the spool displacement. Moreover, the nonlinear-

ity model of the orifice area $A(x)$, must be approximated by the method of linearization based on the Taylor series expansion which makes the nonlinearity of the DPV obvious. The other source of nonlinearity in pipes with the fast-opening DPV is the water hammering effect (Zhu & Jin, 2016). Quick opening DPV, can lead to the nonlinear water hammer effect when water stops or changes directions suddenly.

Hence a general form of the Lipschitz nonlinearity, $g(x,u,t) = x_i^2 \in \mathbb{R}^s$, due to fast opening valves behaviour (Niculita *et al.*, 2013), is added to the state-space equations to cover the nonlinear responses of the valves.

$$S = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; g = \begin{pmatrix} 1 \\ x_2^2 \end{pmatrix}, \; D = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

and the outputs of the fuel rig are ,

$$y = \begin{pmatrix} P_1 & P_2 & P_3 & P_4 & P_5, \end{pmatrix} \tag{6.27}$$

where each output is analyzed individually. For example the fifth pressure sensor output is presented as:

$$y = P_5 = x_2, \quad C = \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

The controllability and observability matrices for this system is of full rank (rank=2) which demonstrate that the system is both controllable and observable :

$$Obs = \begin{pmatrix} C \\ CA \end{pmatrix} = 1 \times 10^4 \begin{pmatrix} 0.0001 & 0 \\ 0 & 0.0001 \\ 0 & 0.0001 \\ -7.4595 & -0.0001 \end{pmatrix}$$

$$Cont = \begin{pmatrix} B & BA \end{pmatrix} = 1 \times 10^{-6} \begin{pmatrix} 0.0 & -0.3337 \\ -0.3337 & 0.0 \end{pmatrix}$$

If all five pressure outputs were evaluated at the same time, then the output of the system is presented as

$$y = Cx$$

where y is presented in equation (6.22) and C is:

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.28}$$

Which makes the system unobservable, because the matrices A in (6.21) and C in (6.23) have not the same number of columns.

It is clear that for this system (if only component disturbances were considered) then the Assumption 1 in the previous chapter is satisfied and $CD = 0$. Hence the following steps should be carried out.

- First, select $D_1 = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$ which satisfies the rank condition

$$rank(CD_1) = rank(D_1) = 1$$

and makes $HD_1 = 0$ and $CD_1 \neq 0$.

- Second substitute D_1 into the equation (5.20), matrix E will be then driven as:

$$E = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

- Then substitute E into (5.12) to establish matrix H

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

- Next from $N = HA - KC$ and ARE (5.21), matrix N and control gain K are obtained as follows:

$$N = \begin{pmatrix} -0.3782 & 0.7614 \\ -0.2386 & -0.7876 \end{pmatrix},$$

$$K = \begin{pmatrix} 0.3782 & 0.2386 \\ 0.2386 & 0.7876 \end{pmatrix}.$$

- Finally by substituting N and K into (5.22), the observer gain L is given

$$L = \begin{pmatrix} 0.3782 & 1 \\ 0.2386 & 0 \end{pmatrix}.$$

The obtained parameters satisfy the conditions (5.27), (5.33) and (5.39), and guarantee the existence of the designed NUIO for the system under investigation.

Matrices $\bar{M}$ and $\bar{G}$ are also obtained from the equations (5.42) and (5.43)

$$\bar{M} = \begin{pmatrix} 1.03 & 0 \\ 0 & 1.21 \end{pmatrix},$$

$$(\bar{M})^{-1} = \begin{pmatrix} 0.9709 & 0 \\ 0 & 0.8264 \end{pmatrix},$$

$$\bar{G} = \begin{pmatrix} -0.4212 & 0.7614 \\ -0.2386 & -1.0476 \end{pmatrix}.$$

Since $\det \bar{M} \neq 0$, therefore $\bar{M}$ is invertible. Hence, in order to make $\bar{M}^{-1}\bar{G}$ Hurwitz the following design matrices are selected

$$A_d = \begin{pmatrix} -0.01 & 0 \\ 0 & -0.2 \end{pmatrix},$$

$$A_i = \begin{pmatrix} -0.013 & 0 \\ 0 & -0.05 \end{pmatrix},$$

$$A_g = \begin{pmatrix} -0.02 & 0 \\ 0 & -0.01 \end{pmatrix}.$$

By selecting these design matrices, $\bar{M}^{-1}\bar{G}$ becomes

$$\bar{M}^{-1}\bar{G} = \begin{pmatrix} 0.4089 & 0.7392 \\ -0.1972 & -1.2198 \end{pmatrix}.$$

and every eigenvalue of $\bar{M}^{-1}\bar{G}$ has strictly negative real part which demonstrate that $\bar{M}^{-1}\bar{G}$ is Hurwitz.

Note that by considering the sensor disturbances in the system, the NUIO design becomes straight-forward as the condition $CD \neq 0$ is satisfied.

Next, the an intermittent fault has been injected to the system.

6.3 Intermittent fault detection

In our example system to show the better applicability of the proposed method, the intermittent fault artificially has been injected into the system. In the experimental fuel rig system, the intermittent fault, $f_{i_s}(t)$, is considered as a fault in $DPV3$.

$f_{i_s}(t)$, is defined as a time varying function of the form $f_{i_s}(t) = dd_i y_{n_c}(t)$, where dd_i, the maximum fault amplitudes, and y_{n_c} is the designer's choice of output ($n_c = 1, \cdots, 5$). Hence the intermittent fault, $f_{i_s}(t)$, could be modelled as follows,

$$f_{i_s}(t) = \begin{cases} f_{i_{s1}} & for \quad 0 \leq t < 30s \\ f_{i_{s2}} & for \quad 30s \leq t < 40s \\ f_{i_{s3}} & for \quad 40s \leq t < 100s \\ f_{i_{s4}} & for \quad 100s \leq t < 160s \\ f_{i_{s5}} & for \quad 160s \leq t < 200s \\ f_{i_{s6}} & for \quad 200s \leq t < 260s \\ f_{i_{s7}} & for \quad 260s \leq t < 280s \\ f_{i_{s8}} & for \quad 280s \leq t < 360s \end{cases} \tag{6.29}$$

where $dd_i \geq 0$, for $(i = 1, \cdots, n)$ are constants and take the values in Table (6.2), and t indicates the time in seconds (Figure 6.6).

dd	Value
dd_1	0
dd_2	0.0020
dd_3	0
dd_4	0.0035
dd_5	0
dd_6	0.0050
dd_7	0
dd_8	0.080

Table 6.2: Numerical values for dd_i

Residual design

As mentioned earlier, the intermittent fault detection system in general consists of two parts: residual generation and residual evaluation including thresholds. The residual for this system is generated as explained in (5.77) as follows:

$$r(t) = \xi(y(t) - \hat{y}(t)) = \xi C\bar{\omega}e_x(t) + \xi K_s f_{i_s}(t) \tag{6.30}$$

The design matrix, ξ, should be define to achieve the following aims the way that:

- $\xi HD = 0$.
- $\xi HS = 0$.
- $\xi K_s \neq 0$.

Hence by choosing

$$\xi = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

all of these conditions are satisfied. In the next section the effectiveness of the designed threshold to evaluate these residuals are demonstrated.

6.4 Simulation Results and Discussions

The model has been implemented and simulated in MATLAB/Simulink® environment. To assess the model performance, simulation results have been compared with experimental data obtained from a fuel rig system while the intermittent fault has been injected manually. The intermittent fault was considered as shutoff valve is getting clogged gradually with some rest periods in-between. Eventually, the shutoff valve clogging is reached to 80% approximately.

The numerical values of the different parameters for the simulations are listed in Table (6.3) and the simulation results, which highlight the modelling capability, are illustrated in the following figures.

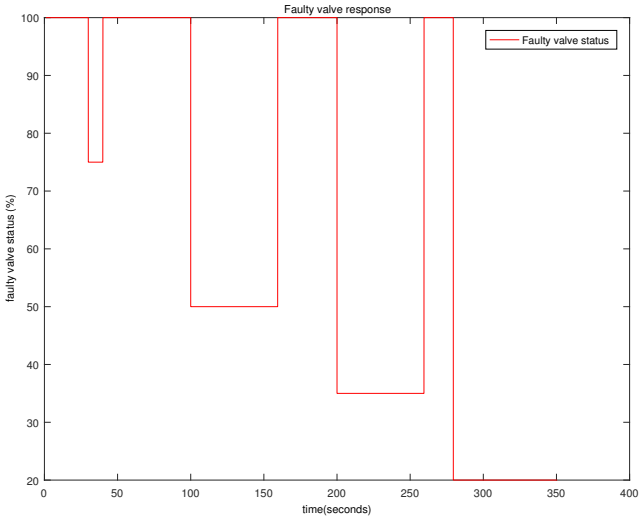

Figure 6.6: Intermittent fault dynamic of the faulty shut-off valve.

Parameter	Value	Unit
ρ	999.0479	$\frac{kg}{m^3}$
g	9.8	$\frac{m}{s^2}$
A_t	0.15	m^2
h_t	0.4	m
C_{d_1}	0.77	$-$
C_{d_2}	0.05	$-$
C_{d_3}	0.77	$-$
C_{d_4}	0.77	$-$
A_{v_1}	0.004	m
A_{v_2}	0.003	m
A_{v_3}	0.004	m
A_{v_4}	0.004	m
L_{p_1}	0.4	m
L_{p_2}	0.4	m
L_{p_3}	0.4	m
A_{p_1}	0.004	m
A_{p_2}	0.004	m
A_{p_3}	0.004	m
η_m	85%	$-$
Q_p	0.784	$\frac{l}{min}$
$\omega\tau$	400	rpm

Table 6.3: Numerical values of the system parameters

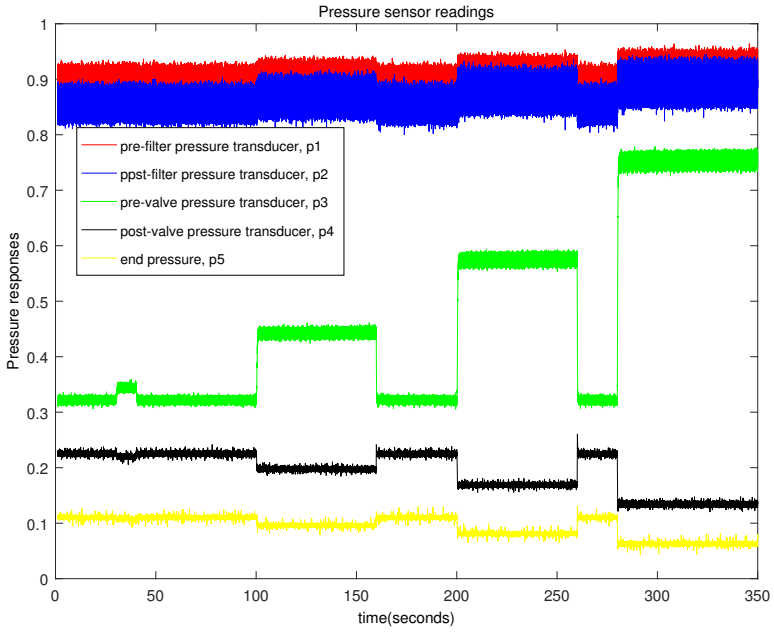

Figure 6.7: The real outputs of the pressure sensors with the initial conditions mentioned in Table (6.4).

Pressure	initial condition
Pre-filter pressure, P_{01}	0.901
Post-filter pressure,P_{02}	0.852
Pre-valve pressure,P_{03}	0.326
Post-valve pressure,P_{04}	0.229
End pressure,P_{05}	0.121

Table 6.4: Pressure sensors initial conditions where P_0 indicates the initial pressure for each sensor respectively.

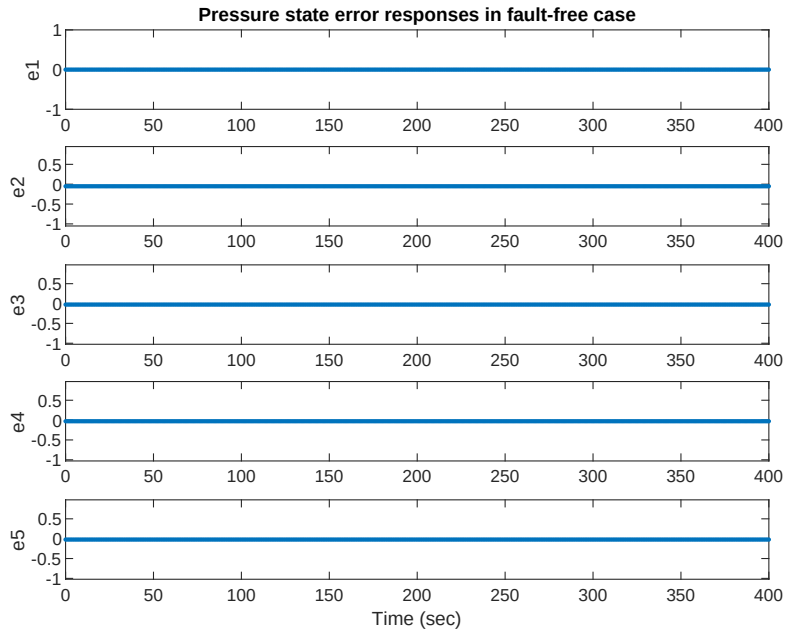

Figure 6.8: State errors in fault free case. In this figure, $e1 = y_1 - \hat{y}_1$ indicates the first pressure error, $e2 = y_2 - \hat{y}_2$ indicates the second pressure error, $e3 = y_3 - \hat{y}_3$ indicates the third pressure error, $e4 = y_4 - \hat{y}_4$ indicates the fourth pressure error, and $e5 = y_5 - \hat{y}_5$ indicates the fifth pressure error.

Figure (6.7) shows the real noisy and faulty outputs out of the five pressure sensors of the fuel rig system. These data are then used to design the NUIO fault detection filters for these sensors to detect the intermittent faults.

Figure (6.8) shows the accuracy of the mathematical modelling of the fuel rig system. This figure indicates the errors between the real outputs of the system and the estimated outputs from the mathematical modelling in the fault-free case. In Figure (6.8) the state errors tend to zero asymptotically which means that the real outputs of the system, (y), and their estimated outputs, $(\hat{y})$, merge towards each other quickly.

Figure (6.9) shows the errors of all five pressure sensors in the presence of an intermittent fault. The non-zero values in these errors indicate the presence of the intermittent fault in the system.

The errors in Figure (6.9) are then used to design the observer-based residuals and thresholds for each pressure sensors.

Figure (6.10) shows the effectiveness of the proposed observer-based fault detection method to detect intermittent fault using fixed thresholds. As this figure shows, in the third and fourth subfigures, the fixed threshold is not able to detect the intermittent fault and/or detect it very late.

The reason behind it could be because of the size of unknown inputs or disturbances in these sensors (see Table 6.1) .

The fixed threshold is the upper bound of available unknown inputs and if these boundaries become significantly huge at some point, then the effect of the faults will be masked by the disturbances and a fixed threshold is not able to detect the fault and

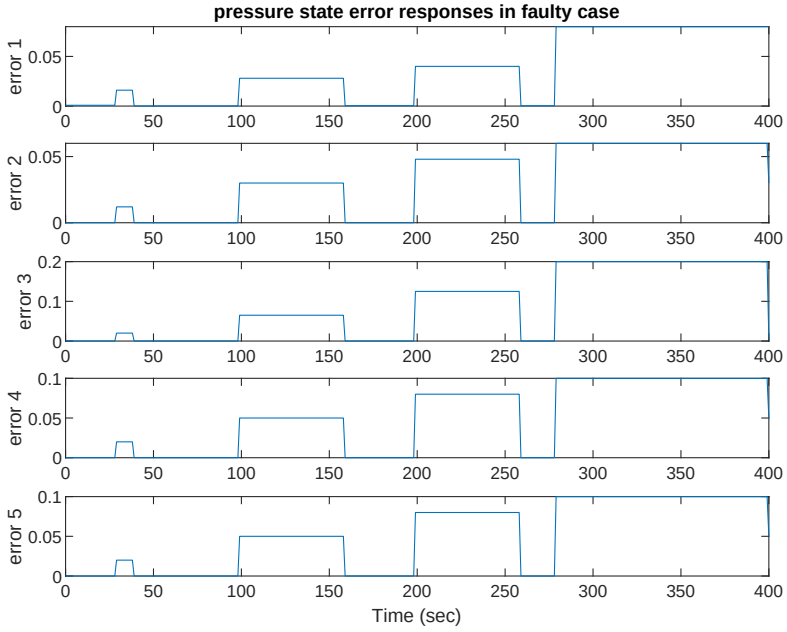

Figure 6.9: State errors in the presence of intermittent fault. In this figure, error1 $= y_1 - \hat{y}_1$ indicates the first pressure error, error2$= y_2 - \hat{y}_2$ indicates the second pressure error, error3$= y_3 - \hat{y}_3$ indicates the third pressure error, error4$= y_4 - \hat{y}_4$ indicates the fourth pressure error, and error5 $= y_5 - \hat{y}_5$ indicates the fifth pressure error.

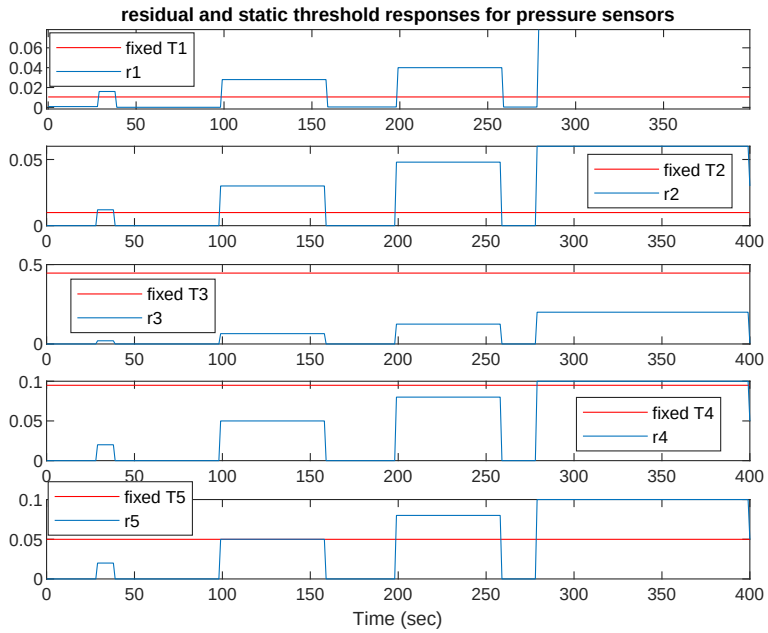

Figure 6.10: The system residuals along with the fixed thresholds. In this figure T indicates the threshold and r is the corresponding residual.

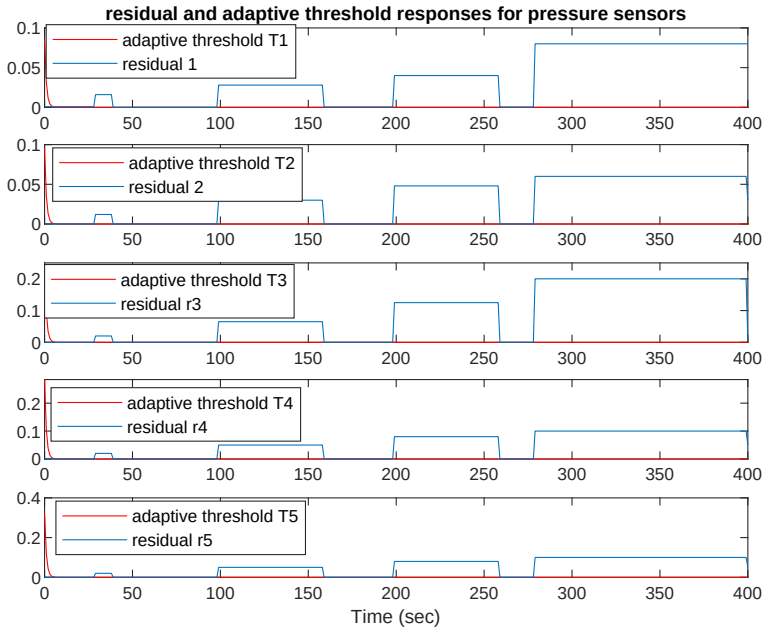

Figure 6.11: The system residuals along with the adaptive thresholds. In this figure T indicates the threshold and r is the corresponding residual.

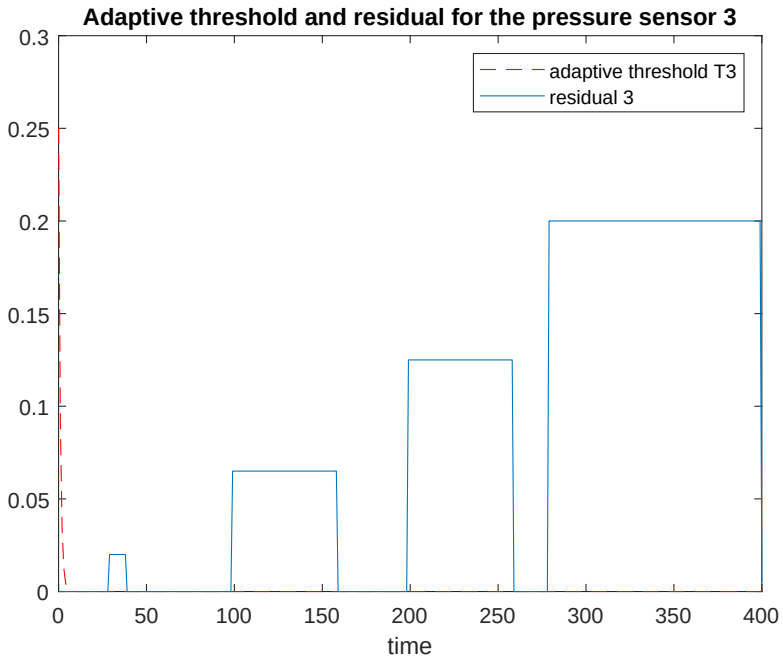

Figure 6.12: The system residual along with the adaptive thresholds for the pressure sensor 3 (pre-clogged valve sensor).

will increase the rate of false-positive (and NFF) significantly.

One solution to this problem could be designing an adaptive threshold. Figure (6.11) shows that in all five subfigures the intermittent fault was detected before it makes a major breakdown to the system responses. This figure also indicates that pressure sensors which are closer to the fault location show severe changes in its output in compare with sensors which are far from the fault location.

The third and fourth subfigures which show the residual of the pre and post clogged filter, demonstrate the most changes in their responses in comparison to others because they are closer to the fault position, although, the fixed threshold missed the fault in these parts of the system.

Moreover, the third sensor in this figure (pre clogged valve sensor), which is the closest sensor to the fault location had the best intermittent fault detection by the use of a designed adaptive threshold. The fault is detected as quickly as it started to make changes to the sensor output. The better view of the designed observer-based residual and adaptive threshold of this sensor is shown in Figure (6.12).

In this figure, the non-zero variable on the observer-based residual in the first few seconds is due to the unknown inputs (measurement errors). This Figure clearly shows that the designed adaptive threshold is sensitive to the fault only and did not detect the unknown inputs due to the disturbances, measuring errors, etc., instead of the intermittent fault.

The presented simulation results demonstrate that the designed adaptive threshold is reducing the false and miss alarm rates in comparison with the fixed thresholds significantly and Figures (6.11) and (6.12), show that the proposed FDI method along with the designed adaptive threshold is successfully detected the intermittent fault in the fuel rig system.

The simulation results also show that the proposed NUIO approach minimising the effect of the unknown inputs (uncertainties) to the state estimation errors and will give a straightforward way to design a robust observer-based residual and adaptive threshold for intermittent fault detection where bounded uncertainties exist.

6.5 Conclusions

Although the effectiveness of the proposed method to accurately detect the intermittent fault has been shown in this chapter but one main limitation to these methods is that in mathematical modelling the diagnostic system is sensitive to the modelling error. Usually, for the very complex systems, these errors are quite considerable and the effect of these modelling errors obscure the effect of fault (intermittent fault) and may cause false/miss alarms.

One solution to this problem could be to employ data-driven methods, particularly when the system under investigation is complex or the mathematical modelling of the system is difficult or costly to obtain.

In data-driven based fault diagnosis methods the precise mathematical model is not needed and is capable of dealing with incomplete information. Hence when the systems are more complex it is more appropriate to use data-driven based fault detection methods such as Bayesian networks.

In the next chapter the BN-based fault detection method is presented. The goal is to detect the intermittent fault and localize faulty components.

Chapter 7

Bayesian Network-Based Intermittent Fault Detection

7.1 Introduction

Systems, in real-world, are often dynamic and their variables and parameters change continuously with time. Hence, sensors in practice are capable of providing much more information, such as how the measured variable varies over time (i.e: the reading could be continuous over time rather than simply a static reading). However, considering dynamic factors in the system when performing fault detection adds difficulties to the task (Marrison, 1992).

Furthermore, in real-world, fault detection algorithms have to make decisions based on uncertain data or uncertain models (Oblak *et al.*, 2007) which are not usually measurable but will make the fault detection process more complicated.

BN, is considered as one of the methods which can handle the uncertainties as the inherent uncertainties in the system can be absorbed into the conditional probabilities (Figueroa & Sucar, 1999). BN is also capable of learning missing data entries (Jackman, 2000).

In this chapter, how the BNs (Static BN (SBN) and HDBN) can detect intermittent faults in a system are investigated.

This chapter aims to introduce a method that could apply to intermittent fault detection in systems with some dynamic aspects.

Therefore, the method is first described in a general way and then is demonstrated on the fuel rig system outlined in Chapter 4.

The diagnostic model in this chapter is built in two phases as follows:

- In the first phase, an SBN is developed and studied for the fuel rig system to detect and identify all possible root causes of intermittent fault.
- In the second phase, an HDBN is presented for the fuel rig system to create a network to represent the intermittent fault detection system which can handle the temporal dynamics in the system.

The developed networks results are then evaluated and compared.

In general, a Bayesian Network Fault Detection (BNFD) system may carry out the following steps after constructing the network in order to define the fault probability and faulty components:

- First, the prior probability distributions when no evidence is propagated to the network is identified.
- Second, the posterior probability distribution through the BN with given evidence is given to each component.
- Third, the ranking of the faulty components can be achieved (components with higher posterior probability with regards to their prior probability have higher rank).
- Finally, the faulty component will be localized and represent as hypothesis variables.

If the physical systems under investigation comprise both continuous and discrete quantities, then to introduce the time-dependent variables along with the discrete variables in the framework of probabilistic models, HDBN is introduced (Gasse *et al.*, 2012). In HDBN fault detection method if the given variable is continuous (has temporal dynamic), the corresponding node is also continuous and its probability distribution is supposed to be Gaussian (normal) (see Appendix C), (Iamsumang, 2015).

This chapter is organized as follows: in Section 7.2, BN is presented. Then BNFD is introduced in Section 7.3 and intermittent fault detection in the fuel rig using BNFD systems along with their simulation results, discussions and sensitivity analysis are presented in Sections 7.4 and 7.5 respectively. Finally, conclusions and the limitations of the proposed methods are presented in Section 7.6.

7.2 Bayesian network

In general, BNs approaches consist of the following steps:

- Finding DAG (structure of the network), denoted by G
- Finding CPT/CPD, one for each node ate G

The diagnostic inference is then possible by looking at the CPT/CPD when some variables values are known (priors). During the calculation, if some state of the nodes is not known, the propagation of priors finds the best hypothesis consistent with the actual data (Xiang *et al.*, 1990).

To build a DBN/HDBN, it is necessary to insert a certain number of regression nodes representing the values of given variables at previous time instants. The regression means a return to a former or less developed state and the regression nodes are all the nodes represented in the previous time slice (Qian & Dougherty, 2016). This number

of regressions for each variable depends on the particular dynamic system under consideration and is very much rely on the nature of the event which is investigated by the corresponding network.

Since the dynamic behaviour of some system components is abnormal for short time intervals and the occurrence and the length of these intervals are unpredictable which makes it difficult to react appropriately, hence the correct and appropriate number of regressions is very important. In cases such as intermittent fault detection, it is easy to miss the fault at the early stages. Hence, the number of regressions must be selected carefully. The expert knowledge or any available reliable information could help to choose this number.

7.2.1 Finding DAG

Finding the best DAG is the crucial step in BN design. Construction of a graph to describe a BN commonly could be achieved based on probabilistic methods using databases of records (Li & Chen, 2014) such as:

- Search and score approach: in this approach, a search through the space of possible DAGs is performed to find the best DAG. The number of DAGs, $f(n_n)$, as a function of the number of nodes, n_n, grows exponentially with n_n (Lee & van Beek, 2017),

$$f(n_n) = \sum_{i=1}^{n_n} (-1)^{i+1} \left(\frac{n_n}{i}\right) 2^{i(n_n - i)} f(n_n - i). \tag{7.1}$$

  Hence for the system with $n_n = 16$, using equation (7.1), there is 8.38×10^{46} possible DAGs and exhaustive search on the space of all DAGs is not practical. In practice, search and score is an initial approximation of the network structure from data. However, the accuracy of the learned BN is then largely affected by the richness of the data and the prior knowledge of the network ordering.

- K_2 algorithm: in this algorithm initially each node has no parents. Then the parents are added incrementally (the score of the resulting structure will increase). When the addition of a single parent does not increase the score, it stops adding parents to the node. Before the algorithm started, the possible parents of every variable must be defined (order) based on the experience, human knowledge and other available evidence. If the order is known, a search over this order is more efficient than searching over all DAGs (Lerner & Malka, 2011).

- MCMC algorithm: it is a stochastic process, where the current state depends only on the past state. By applying MCMC algorithm, the chain is the sequence of DAGs in which the search for the best DAG is performed. The MCMC process finds a very complex function, the DAG, that best agreed with the evidence contained in the database by applying a probabilistic approximation (Roberts & Rosenthal, 2009).

  The MCMC algorithm starts at a specific point in the space of DAGs. The search is performed through all the nearest neighbours, as it moves to the neighbour that

has the highest score. If no neighbour (the graph that can be generated from the current graph by adding, deleting or reserving a single arc) has the highest score than the current point, a local maximum has been found and the algorithm will stop (Andrieu & Atchade, 2007; Hoshino, 2008).

In this research, the model-based FD along with the mathematical modelling of the system, help the network designer to record all the variables of interest (the variables which are difficult, expensive and sometimes impossible to monitor by the sensors but important for the fault detection) to construct the network structure (Figure 7.1). The model-based FD method is also helpful to avoid unnecessary complexity and uncertainty due to the increasing number of variables and hidden layers.

7.2.2 Finding CPT/CPD

the conditional probabilities are the probability distributions of a node given its parents and are calculated based on the resulting structure. Usually, the nodes with no parents (root nodes) correspond to the prior probabilities (Probabilities before any evidence is given) and the nodes with parents correspond to the conditional probabilities (Xiang *et al.*, 1990).

The computation of conditional probability, $Pr(X_F|X_E)$, where X_F, is the variable of interest (eg., fault cause) and X_E, is the variable or set of variables that have been observed (eg., sensor observations) are called **inference**.

There are many different algorithms for calculating the inference in the network, which apply different trade-offs between speeds, complexity, generality and accuracy such as:

- Variable Elimination Algorithm (VEA), (Jung *et al.*, 2009) which allow the inference calculation with a generic structure
- Maximum Posterior Likelihood Algorithm (MPLA), (Dempster *et al.*, 1977)
- Monte carlo algorithm(Mau, 1996)
- Junction tree algorithm(Xia & Prasanna, 2008)
- Gibbs sampling algorithm(Lawrence, 2005)

In this research, the Junction Tree inference algorithm was applied.

In general, BNs in addition to their simple causal graphical structure, have some other appealing properties:

- The ability to update initial beliefs about the values of each variable (prior probabilities) in the face of new evidence via Bayes' theorem.
- A BN can perform different types of inference: deductive (top-down, forward, predictive), diagnostic (bottom-up, inverse, explanatory). This is useful since the same BNs can be used for both the detection (assessment) and the evaluation.
- Analysts can make probability judgments consistent with the direction of causality.

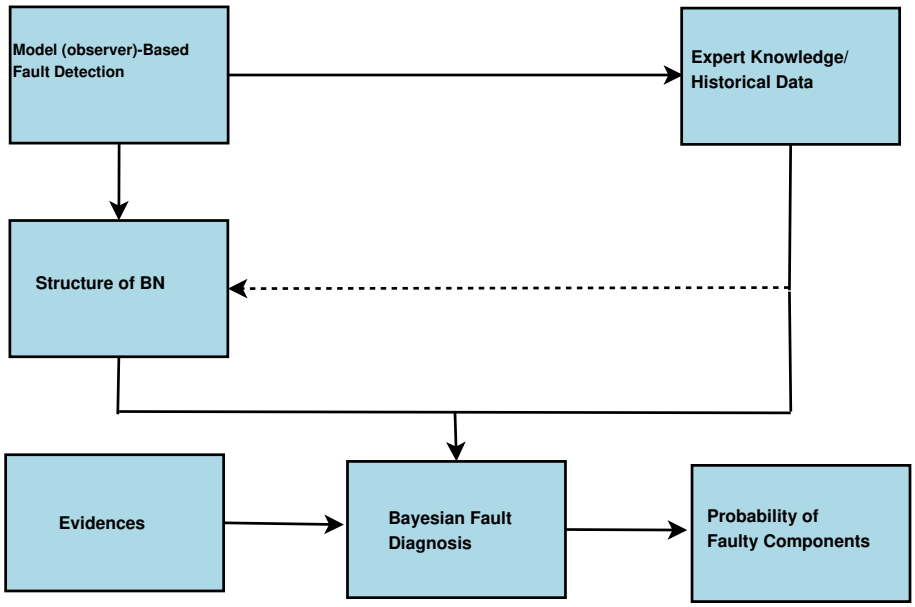

Figure 7.1: Flow chart of the BNFD network for fault diagnosis.

- Evidence can be entered into the model, and the effect on the other variables (nodes) can be observed (improving or worsening, and by what magnitude).
- BNs can work with data of different types and sources: they handle a mix of subjective and objective data, hence, supplement traditional experimental and statistical methods.

7.3 Intermittent fault detection using BNs

The complete sequence of intermittent fault detection approach using BN consists of the following steps:

- Data collection (historical data, laboratory data, environmental data, knowledge-based data).
- identifying the key variables, the failures and the operating modes.
- Describing casual relationships (static, dynamic).
- Construction of the network: a BN for the fault detection may be built,
  - Traditionally, by intuition (expert knowledge): the resulting network model may be incomplete and casual relationships may be incorrect;
  - By learning through data (structure learning): the richness of data, the prior knowledge about the process and other FDI methods results such as model-based FDI guarantee the accuracy;

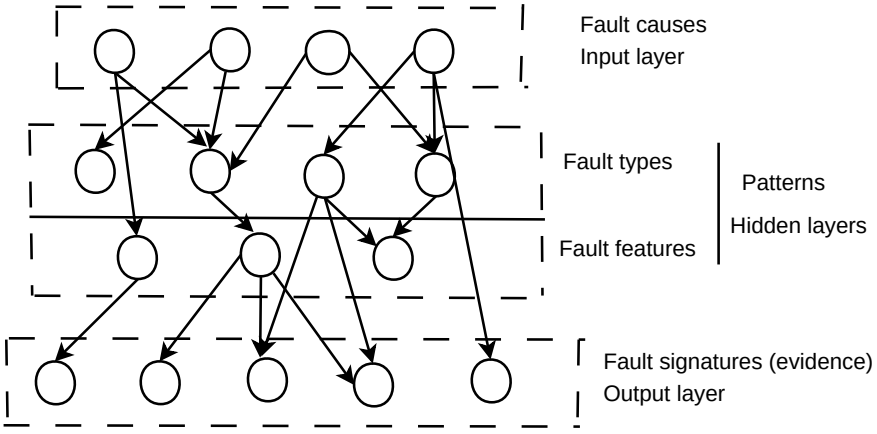

Figure 7.2: BN-based fault detection structure.

  - Combination of both: when using a combination of data and expert knowledge to construct the BN, usually the edges which are the same in all algorithms are kept and the remaining edges are removed based on the experience, expert knowledge and other evidence. For example, independent sensor edges among sensors can be removed because their values are always observed. Then the remaining edges are compared with the relationship among variables on the process.

- Arrange the groups of variables in layers. In the designed BN for FDI in Figure (7.2):
  - Variables which exhibit significant changes when the fault occurs present the output sets or evidence (fault signatures)
  - Variables that are independent of the faults are the input sets (hypothesises or fault causes)
  - Variables, the hidden nodes, which through them the relations between all the other nodes are learned (mediating or fault type/fault feature).

    The mediating variables are often introduced to facilitate the acquisition of CPDs/CPTs. If the introduction of mediating variables serves no purpose, then they should eliminated from the model; otherwise, they may add unnecessary complexity to the network and cause the designed network performs poorly (Kwoh & Gillies, 1996).

- Calculation of conditional probabilities: the BNFD is now ready to infer probabilities of faulty components. The system measurements are propagated to the BNFD system as evidence and the probability distribution for each hypothesis variable is inferred, where, **hypothesis variables** contains all the possible hypothesises on failure sources or conditions.

- Finally, the subsequent results through ranking the faulty components by their corresponding probability distribution is given.

In the complex systems subject to the noise or disturbances and the intermittent faults, it is very difficult to determine which hypothesis is the root cause of the intermittent

fault. There are many possible hypothesis variables at each time step of DBN/HDBN, and prune away the wrong hypothesis often is not possible until several time steps after the hypothesises are generated.

The presented approach may also be used to localize multiple faulty components that will be correlated to produce a single symptom.

7.4 Intermittent fault detection in the experimental fuel rig system using BNs

In this research, the IVHM fuel rig (and its equivalent mathematical model in Matlab) is used to generate data. The simulation time is about 360 seconds and the sampling time is 0.001 seconds.

In the considered experimental fuel rig all the sensor readings and their related pressure changes are continuous with the Gaussian distribution and there are also some discrete variables related to the valve statuses, tank water measurement, pipe statuses, etc. The values related to the presented intermittent faults are also Boolean (True/False). Hence with this variety of data types, detecting the intermittent fault in this system is challenging.

The control system and data collection can best be described with reference to the graphical user interface (GUI), National Instruments LabVIEW software version 8.6. Referring to Chapter 4 there are three main controllers:

- Pump control unit: manual or mission profile selection; speed feedback loop
- Valve control unit: shut off valve position control
- Direct proportional valve control unit: operated via the filter and gear pump fault sliders

And the data presented to the user in the GUI are as follows:

- Pump speed in the pump control unit,
- Pressures in different points of the system (e.g. before filter, after the filter, after the pump, after shut-off valve)
- Time traces of pump speed, pressure, flow for both operational modes. These are displayed according to running mode selection.

The designed BNs to address this FDI problem are developed in Matlab using Bayesian Network Toolbox (BNT) written by Kevin Murphy (`http://www.ai.mit.edu/~murphyk/Software/BNT/bnt.html`).

This toolbox was chosen due to its very good implementation of the inference algorithms, the possibility of implementing real DBNs, the possibility to work with discrete nodes that have continuous parents and its computational power. Also being written in Matlab makes it much easier to manipulate the data and model-based FDI results, generated in Matlab/Simulink earlier for the purpose of training and testing data.

Sensor	Mean=Mode	Standard Deviation
Pressure Sensor 1 (PS1)	0.9012	0.014
Pressure Sensor 2 (PS2)	0.8522	0.015
Pressure Sensor 3 (PS3)	0.3261	0.0036
Pressure Sensor 4 (PS4)	0.2292	0.02
Pressure Sensor 5 (PS5)	0.1220	0.03

Table 7.1: mode and mean values for each sensor data

7.4.1 Experimental fuel rig conditions and assumptions

In order to detect intermittent fault in the experimental fuel rig under investigation the following conditions are satisfied:

1) The potential components to be faulty in the fuel rig system based on expert knowledge (Niculita *et al.*, 2013), mathematical modelling, model-based FDI and other available evidence are:

   - Filter (DPV1): clogged.
   - Gear pump: Degraded (less flow for the same pump speed).
   - Shut-Off Valve (SOV): clogged and stuck mid-range.
   - pipe 4: leaking.
   - pipe 3: leaking.
   - Nozzle (DPV5): clogged.

   However, from the hardware and software point of view, the control system is ready to accommodate all initially planned failure modes in a plug and play manner although in the present research only one of them (clogged shut-off-valve (S-O-V)), is considered.

   The user has control over pump speed and shut-off valve position in manual or mission profile mode and direct proportional valve position in order to be able to inject the intermittent fault into the system.

2) The fuel rig data are also collected in both steady-state condition where the pressures and flows, etc., are constant over time (more or less) at pipelines, and the transient condition when the variables may change rapidly.

3) The temperature is not set but is considered as room temperature.

4) All the sensors are perfectly reliable (Niculita *et al.*, 2013). So, in this research, the fault associated with each sensor is not considered as one of the possible root causes of the intermittent fault.

   Sensor Faults: sometime the sensors which are collecting data/evidence are faulty. One of the fault to be considered is that they might be stuck over time. A stuck sensor node represents the stuck state of a sensor. In this situation, the sensor's

reading is the same over some time, regardless of what the underlying process state is. The sensor fault node does not share the same components (nodes) as the corresponding sensor node in the network but its value is updated using the current sensor value.

In the designed SBN and HDBN a direct casual relation (direct link) between hypothesis nodes and sensor nodes means that component has more/grater effect or influence on that sensor. The sensor faults usually have a direct link to its corresponding sensor.

5) The system has two operating modes, the engine-feeding mode which is **active**, and recirculation mode which is **inactive**. In this research, the engine-feeding mode is investigated.

6) The fuel rig system with the intermittent fault is also subjected to the uncertainty, and the following are the key issues, which can significantly affect the formulation of the hypotheses,

   - The time that the symptoms of the fault appear.
   - Their duration,
   - The time that the observation/measurements are made,
   - The time that the faults are induced.

   Hence, the model must be able to update the system given that observation. Also, the evidence must be made over time to capture the evolution of the system as it changes over time.

7) All the root nodes except ″No water″ have two states as follows:

   - State 1: being faulty.
   - State 2: not faulty.

   The root node ″No water″ has two states as follows:

   - State 1: true (there is no water, or fuel available).
   - State 2: false (there are enough water or any other fuel available in the system).

   The intermediate node ″Intermittent fault″ has two states as follows:

   - State 1: true (the system is faulty).
   - State 2: false (the system is healthy).

   The intermediate node ″Pressure change″ has two states as follows:

   - State 1: true (there is a pressure change in the system).
   - State 2: false (there is no pressure change in the system).

   The evidence node ″Sensors″ has five states as follows:

- State 1: PS1 is true (Pressure sensor 1 is evidence to the system).
- State 2: PS2 is true (Pressure sensor 2 is evidence to the system).
- State 3: PS3 is true (Pressure sensor 3 is evidence to the system).
- State 4: PS4 is true (Pressure sensor 4 is evidence to the system).
- State 5: PS5 is true (Pressure sensor 5 is evidence to the system).

8) The probability density function associated with the data provided by each sensor is characterized by a Gaussian (Normal) distribution (Table 7.1).

Assumptions:

In order to detect intermittent fault in the experimental fuel rig under investigation the following assumptions are made:

1) The fuel (water, oil) is always available.

2) Root nodes are orthogonal, meaning there are no interactions between them.

3) In this research to avoid exponentially increase the size of the network, decrease the compute time for the inference and manage the size of needed memory, a single class of faults which is the intermittent fault is considered. The fault variables are usually considered as Boolean variables where their nodes are represented by two states only (faulty and healthy).

7.4.2 Graph construction (finding DAG)

To design the BN for the IVHM fuel rig the nodes must always be numbered in topological order: parents (ancestors) before the child (descendants). Then to specify the graph structure, the size and type of the nodes (discrete or continuous) must be specified. The size of discrete nodes are the number of possible values each node can take on (binary) and the continuous nodes can be presented as a vector and their size are the length of that vector.

The first attempt of designed BN for the fuel rig system in a one-time slice is presented in Figure (7.3). However, the presented BN can be improved and simplified (Ockham's razor rule) by accommodating the expert knowledge, and other available reliable information.

Occam's razor is a principle from philosophy which is very useful in BN and is another reason why BNs are applicable for many complex systems. Occam's razor principle states that " unnecessarily complex models should not be preferred to simpler ones." Note that the simpler models mean fewer parameters to train, so faster computations and more generalisations and BNs automatically and quantitatively embody Occam's razor (MacKay, 1992). Thus the complex BN for the fuel rig system presented in Figure (7.3) is simplified and build based on the combination of the expert knowledge and data. The final SBN network is presented in Figure (7.4).

Moreover, The same DAG is used to design HDBN. The designed HDBN (Figure 7.5) is solved at time, $(t = 0)$, with prescribed initial conditions (prior probabilities).

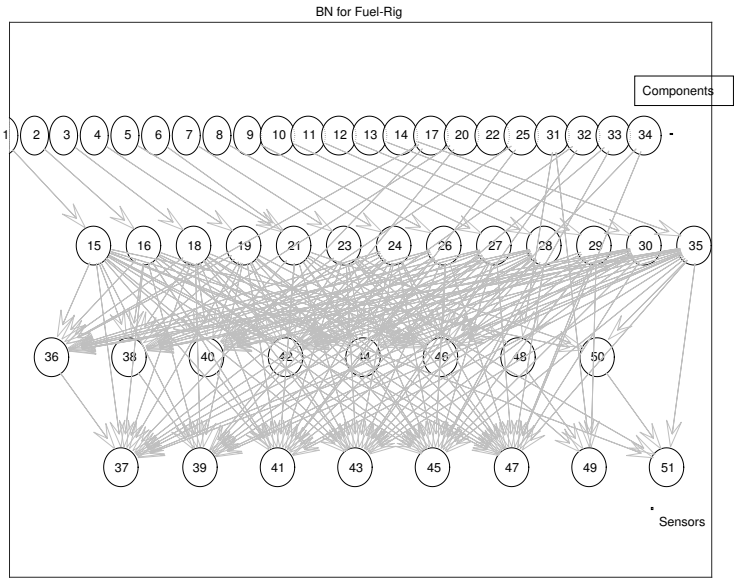

Figure 7.3: The static BN for the fuel rig (first attempt).

The results of this analysis will give the values which are used in computing of the network probabilities in the second time slice. The number of slices (regressions) in this system is 360 due to the sensitivity of the intermittent fault detection and this process is repeated over the whole time of the process.

In general, for the designed HDBN, the following probabilities need to be calculated:

- Prior probabilities: the prior probability of an uncertain quantity is the probability distribution that would express one's belief about this quantity before some evidence is taken into account, $Pr(X_0)$, which are defined based on the knowledge of experts, other FDI algorithm such as model-based fault detection, and other available evidence. This probability is the state prior probability at time $t = 0$.

- Temporal probabilities or interslice probabilities: dynamic networks are merely static networks straddling time periods or time slices. The future state of some nodes can be influenced by their prior state (feedback nodes). Some nodes can also affect other nodes in other time slices, $Pr(X_t|X_{t-1})$, which are the set of parents of X_t^i, that can be either in the current time slice (t) or previous time slice $(t-1)$. This is the probability between different slices.

- Posterior (Conditional) probabilities: for each time slice any new information is propagated to the model to update current priors, $Pr(Y_t|X_t)$, where Y_t indicates the evidence. When the posterior probability of a node are higher than its prior probability, then the abnormality shows itself (hypothesis variables).

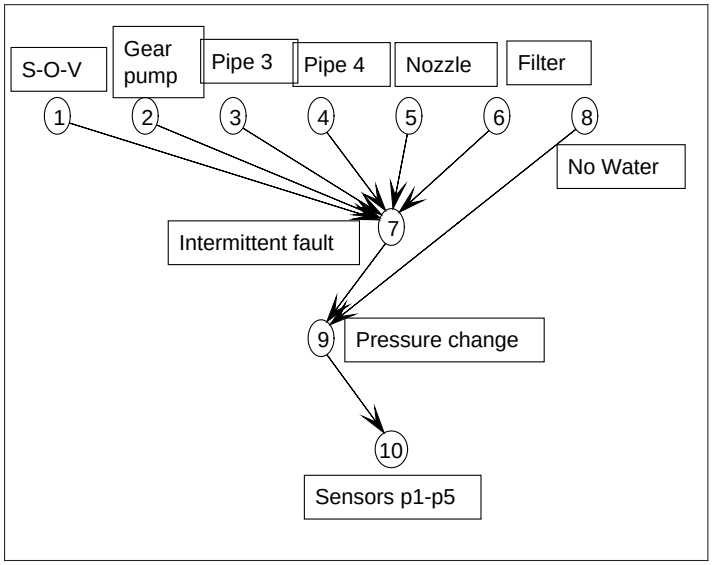

Figure 7.4: The SBN for the experimental fuel rig system.

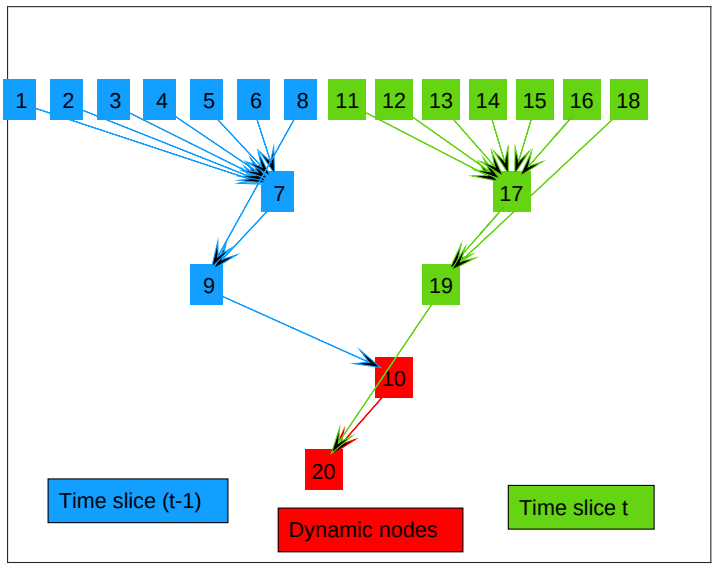

Figure 7.5: The HDBN for the experimental fuel rig system. In this figure, nodes 1- 10 represent the S-O-V, gear pump, pipe 3, pipe 4, nozzle, filter, intermittent fault, no water in the tank, pressure change and sensors at time slice (t-1) respectively, and nodes 11-20 represent the S-O-V, gear pump, pipe 3, pipe 4, nozzle, filter, intermittent fault, no water in the tank, pressure change and sensors at time slice (t) respectively. Nodes 10 and 20 are the pressure sensors in two different time slices which are the only dynamic nodes in this network.

7.4.3 Finding probability

To detect intermittent fault root causes in the experimental fuel rig system, prior probabilities are important to obtain the diagnostic inferences and the conditional probabilities. If some states of the nodes are not available, the propagation of the priors finds the best hypothesis consistent with the actual data.
As prior, it is easiest to give equal failure probability to all the components (which have the potential to be the fault cause) unless other information is available. However, in the real world, the prior probabilities of all the components in a network are not the same. In this research, the prior probabilities have been assigned to each root node using a combination of expert knowledge elicitation and data.
Then the components whose their posterior probabilities have increased with respect to their prior probabilities could be the potential root causes of the intermittent fault which are represented as hypothesis variables in the network.

Hence, in order to generate an initial fault candidate's set, a probability distribution, which is computed through the SBN and/or HDBN with given evidence, is attached to each component and then ranking of the faulty components is achieved.

When the observed information variables (sensors) is abnormal (faulty), then the probabilities of all the hypothesis variables will increase. The probabilities of the hypothesis variables which have direct casual relation with the given evidence/information (sensor) may be higher.

7.5 Simulation results and discussions

7.5.1 Prior probability

To obtain the prior probabilities for the root nodes (S-O-V, gear pump, pipe 3, pipe 4, nozzle, filter and no water), the expert knowledge and available literature on the fuel rig is used by setting a Beta priors as $x \sim Beta(a,b) = Uni(0,1)$, (see Appendix D). The results are summarized in Figures (7.7-7.13) and Table (7.2).

In Figure (7.6) if we take the average of all the Beta fits (see Appendix D) then the prior mean, M, will be $M = 0.15$.

However, one of the prior fits is very different from the others with high variance, V, $(M_5 = 653, V_5 = 0.0017)$.

Hence, it is needed to ignore the Beta fits which are very different from others especially if they have high variance. In this Figure by eliminating M_5, the average beta prior mean is $M = 0.0560$ which is closer to the expectation. The same method is used in this chapter to define the prior probabilities for all the root nodes of the experimental fuel system.

In Figures (7.7-7.13), M presents the prior mean and V presents the prior variance for each root node.

Components (root nodes)	Prior probabilities
Shut-off valve	0.05
Gear pump	0.02
Pipe 3 (pipe before the shut-off valve)	0.005
pipe 4 (pipe after the shut-off valve)	0.004
Nozzle	0.0002
Filter	0.012
No Water (no water in the tank)	0.0001

Table 7.2: Root nodes and their state 1 prior probabilities in SBN

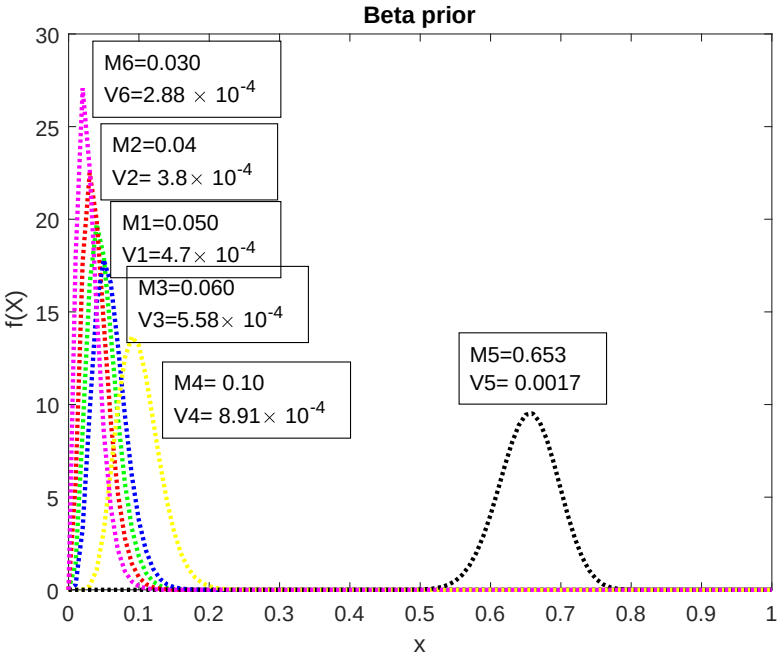

Figure 7.6: Beta prior for discrete nodes.

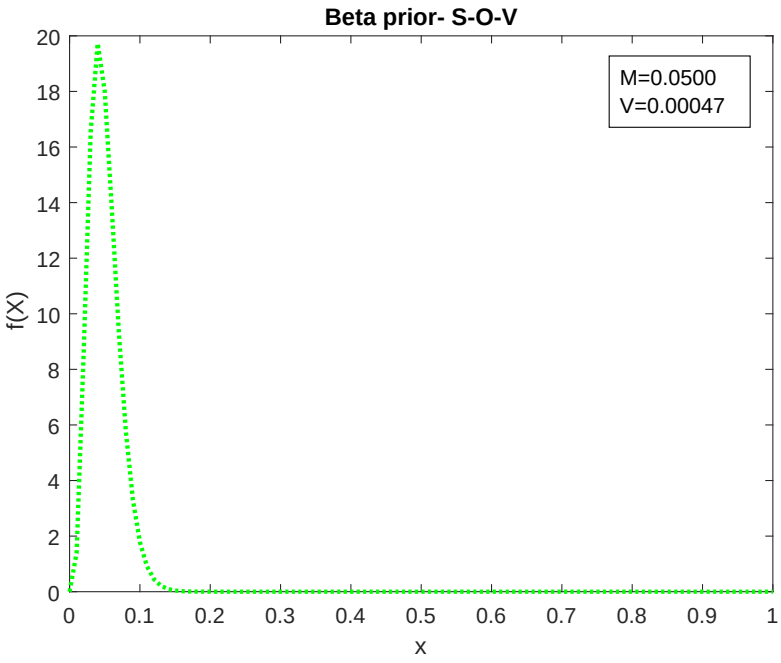

Figure 7.7: Beta prior for the S-O-V.

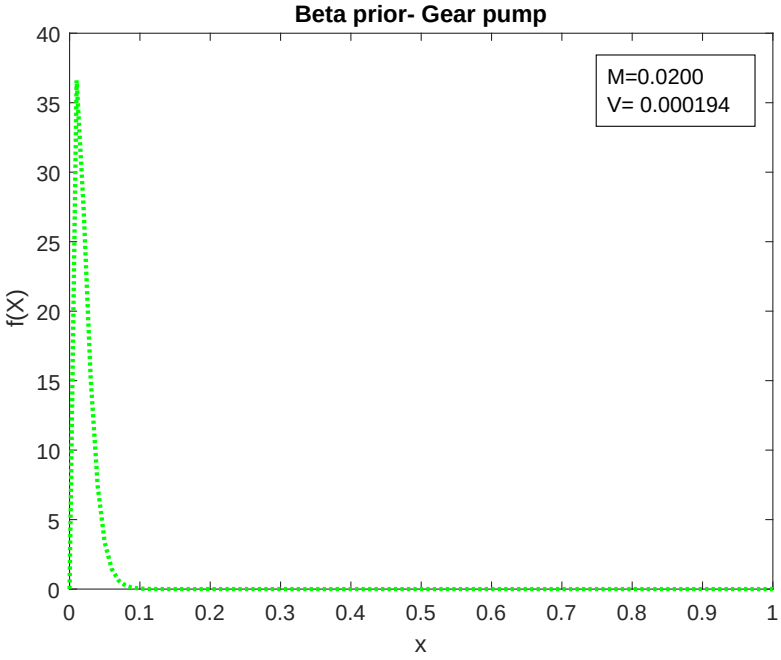

Figure 7.8: Beta prior for the gear pump.

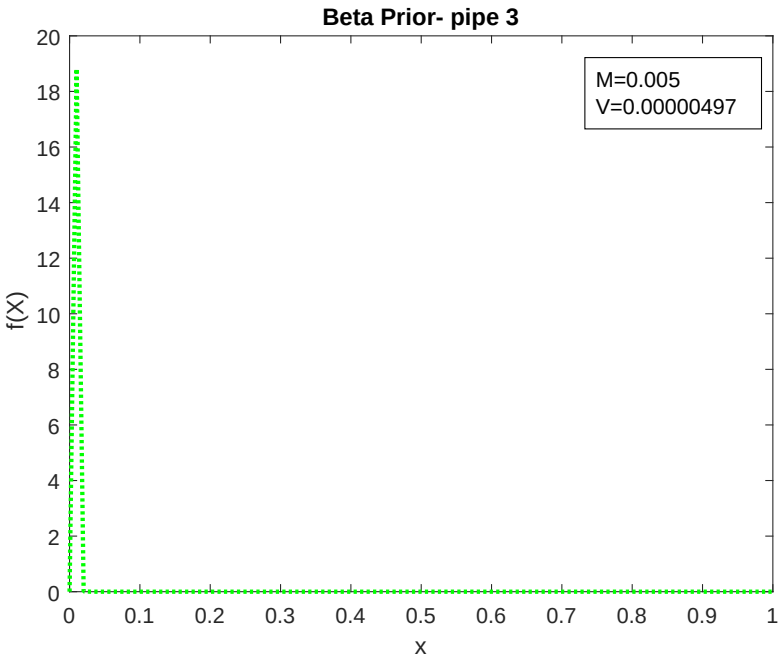

Figure 7.9: Beta prior for the pipe 3.

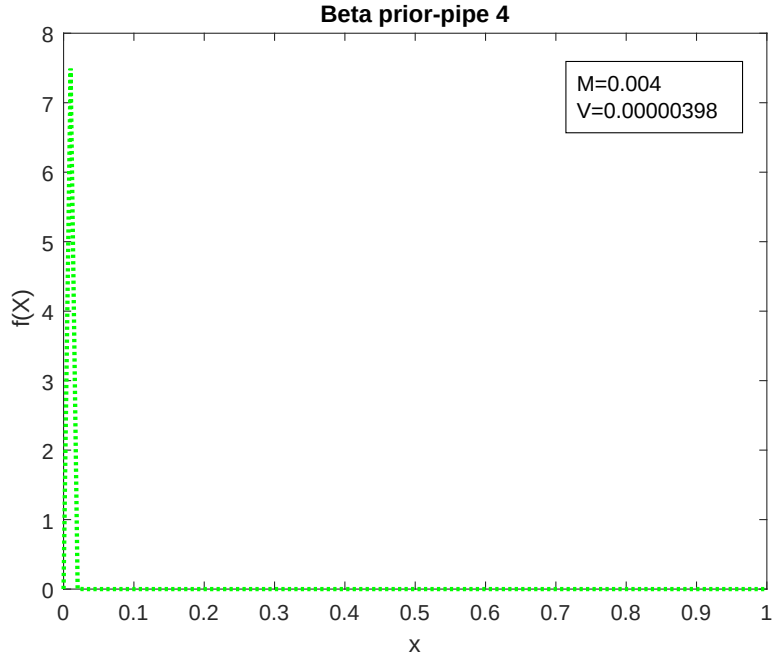

Figure 7.10: Beta prior for the pipe4.

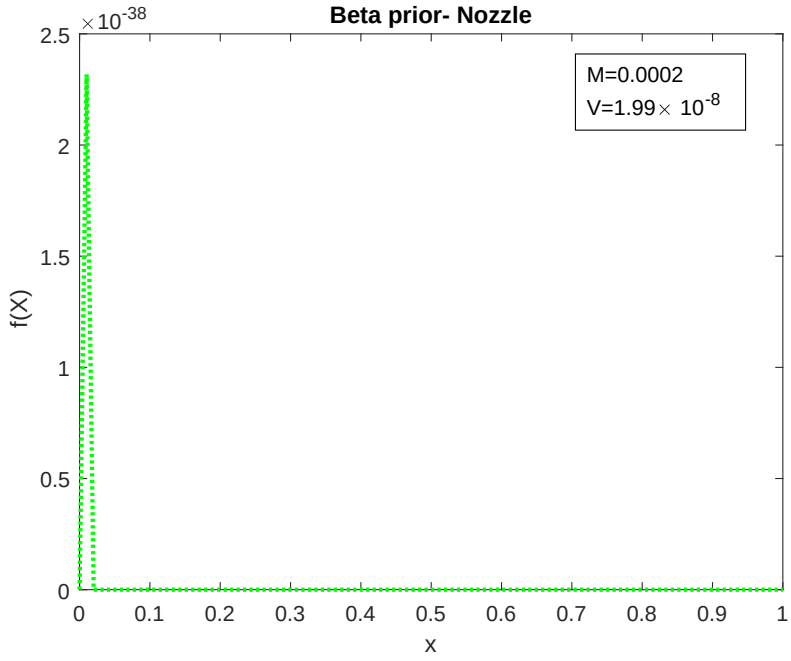

Figure 7.11: Beta prior for the nozzle.

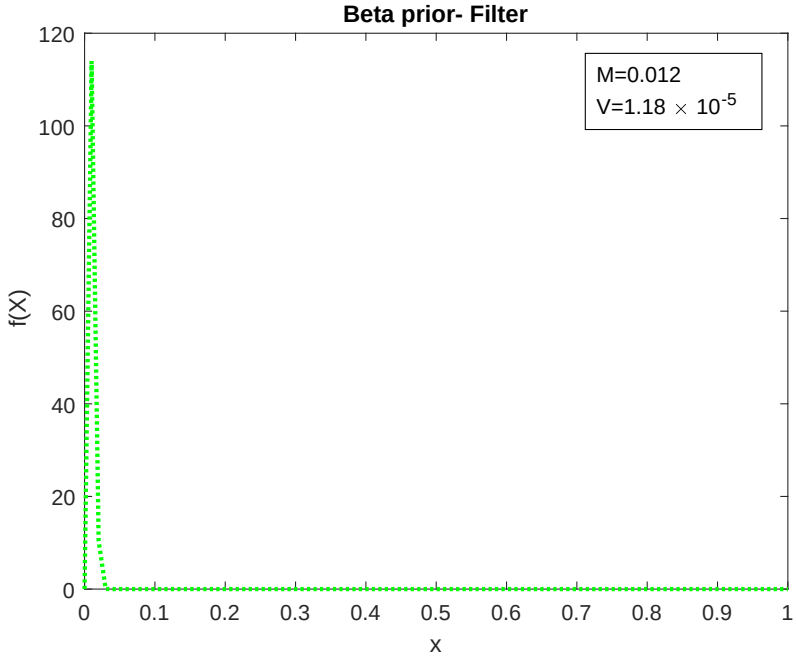

Figure 7.12: Beta prior for the filter.

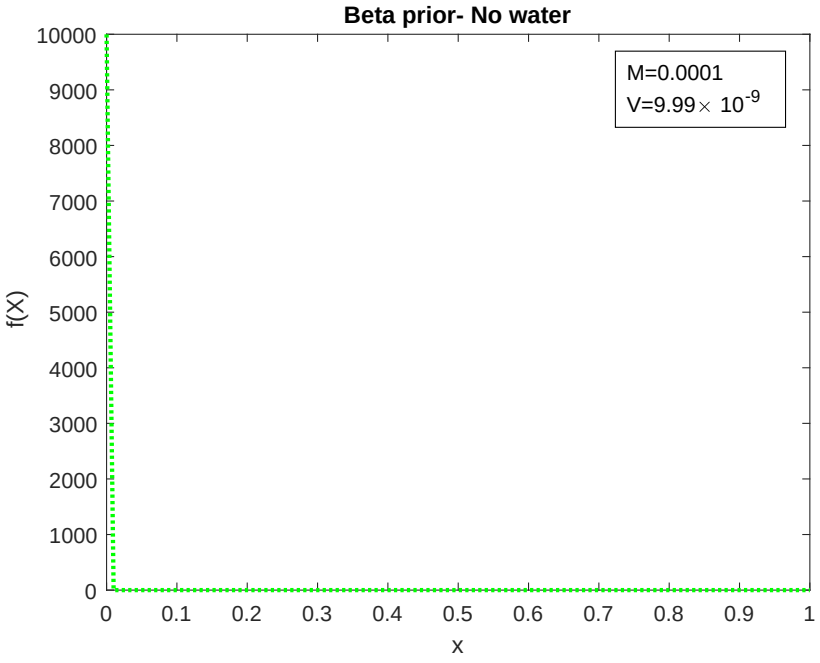

Figure 7.13: Beta prior for the no water.

7.5.2 Posterior probabilities using SBN

By assigning the prior probabilities to the SBN, (Figure 7.4), and propagated evidence, PS1-PS5, the posterior probabilities of each node is obtained.

Figures (7.14-7.18) show the posterior probability of the root nodes when PS1-PS5 were propagated to the network respectively. These figures demonstrate how the posterior probabilities of each root node were changing when different evidence was spread to the network.

Figures (7.14) and (7.18) identify the pipe 4 with higher posterior probability as the main root cause of intermittent fault when PS1 and PS5 are the evidence to the system. Although, Figures (7.15), (7.16) and (7.17) present S-O-V as the main root cause of intermittent fault when PS2, PS3 and PS4 are the evidence to the network respectively.

The results obtained from Figures (7.14-7.18) are summarized in Table (7.3). This Table shows that the sensors which were closer to the fault location, (S-O-V), show the greater changes rather than others. Moreover, the posterior probability of the S-O-V was significantly changing when PS3, PS4 and PS5 were evidenced to the system. However, PS1 and PS5 which were far from the S-O-V location detected pipe 4 as the main root cause of intermittent fault. These results also demonstrate that the sensors which were closer to the fault position have shown more changes.

Moreover, Figures (7.19-7.25) compare the posterior and prior probabilities of the root nodes obtain for all evidence.

Figures (7.19) and (7.20) show that the posterior probabilities of S-O-V and gear pump with evidence PS2, PS4 and PS5 has been increased, despite PS1 and PS5. However, Figures (7.21), (7.22) and (7.24) show that the posterior probabilities of the pipe 3, pipe 4 and filter for all the evidence have been increased. It means that with every evidence propagated to the SBN, pipe 3, pipe 4 and filter are detected as the main root

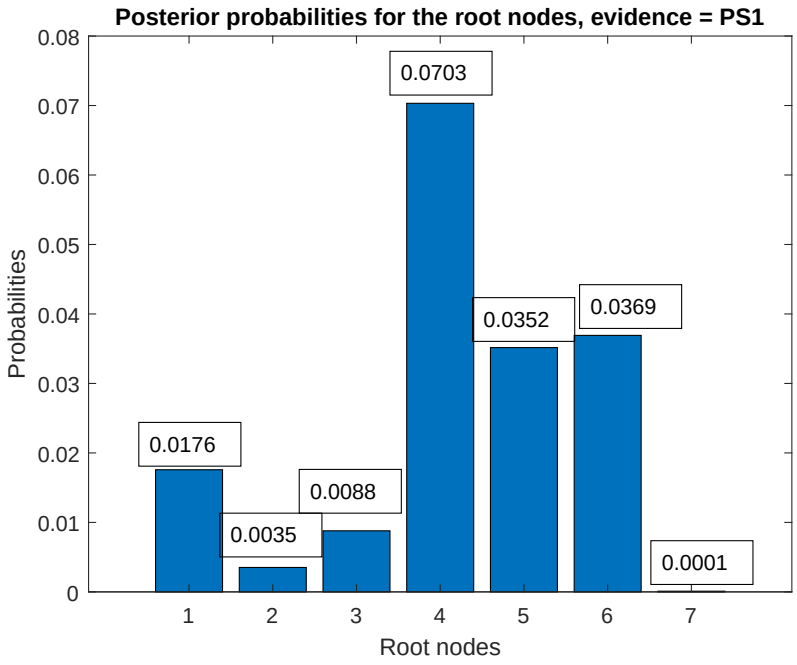

Figure 7.14: Posterior probabilities for the root nodes when pressure sensor 1 is the considered evidence to the network. In this figure, 1 represents the S-O-V, 2 represents the gear pump, 3 represents the pipe 3, 4 represents the pipe 4, 5 represents the nozzle, 6 represents the filter and 7 represents no water in the tank.

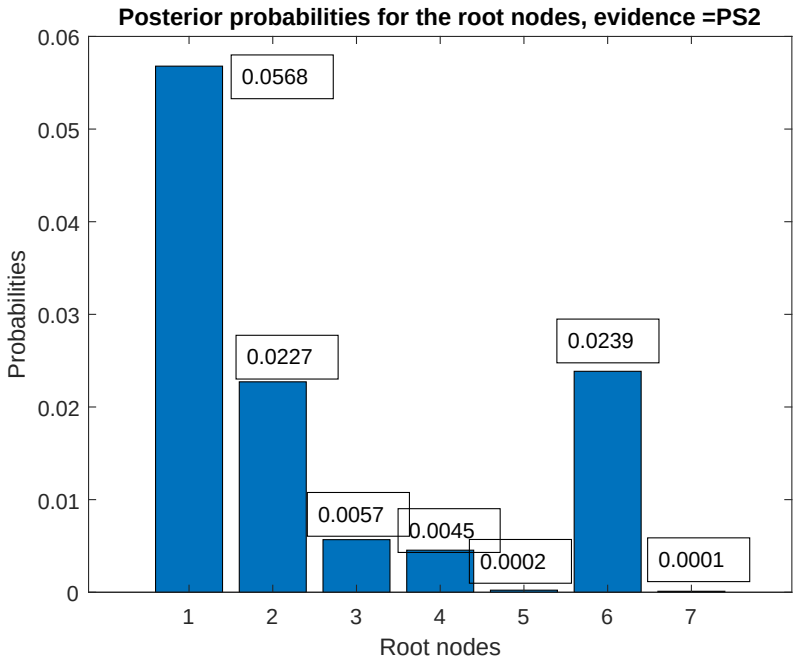

Figure 7.15: Posterior probabilities for the root nodes when pressure sensor 2 is the considered evidence to the network. In this figure 1 represents the S-O-V, 2 represents the gear pump, 3 represents the pipe 3, 4 represents the pipe 4, 5 represents the nozzle, 6 represents the filter and 7 represents no water in the tank.

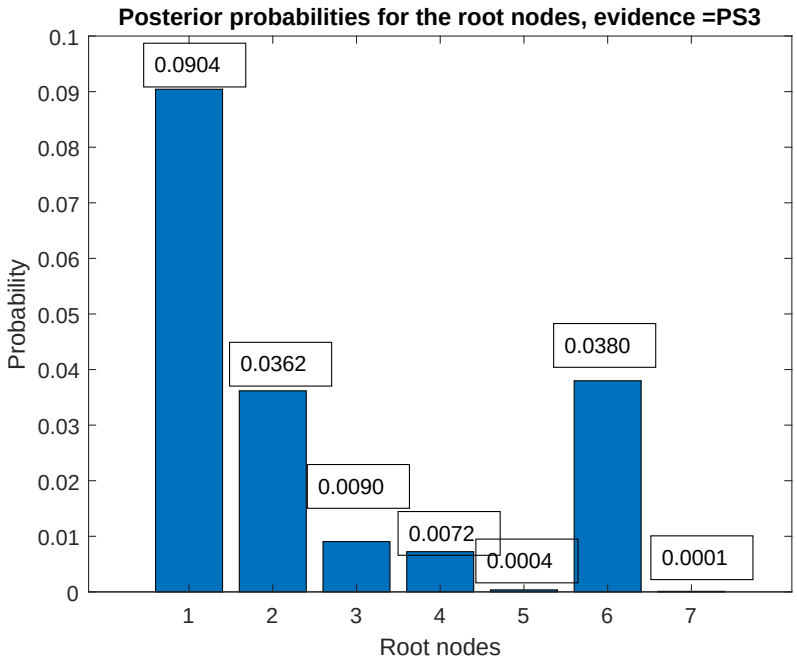

Figure 7.16: Posterior probabilities for the root nodes when pressure sensor 3 is the considered evidence to the network. In this figure 1 represents the S-O-V, 2 represents the gear pump, 3 represents the pipe 3, 4 represents the pipe 4, 5 represents the nozzle, 6 represents the filter and 7 represents no water in the tank.

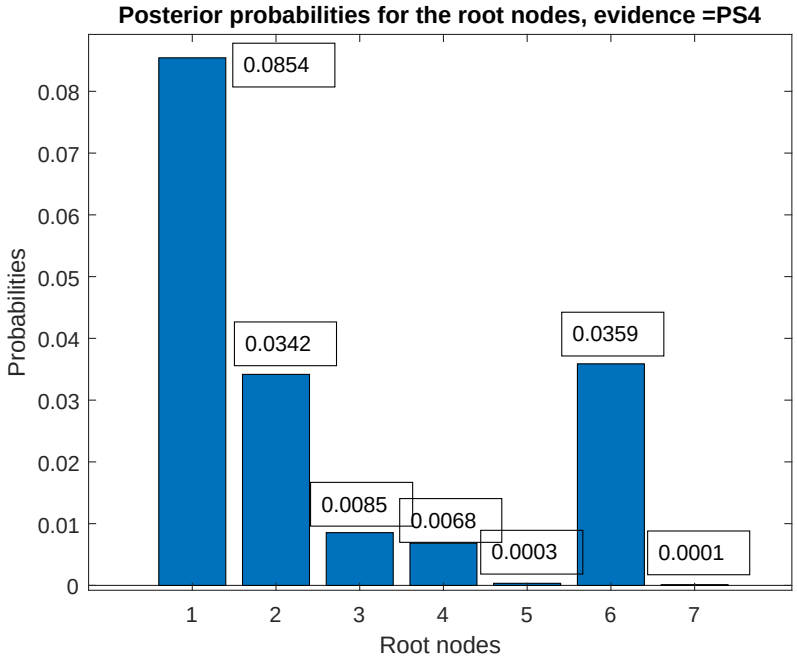

Figure 7.17: Posterior probabilities for the root nodes when pressure sensor 4 is the considered evidence to the network. In this figure 1 represents the S-O-V, 2 represents the gear pump, 3 represents the pipe 3, 4 represents the pipe 4, 5 represents the nozzle, 6 represents the filter and 7 represents no water in the tank.

Components (root nodes)	Prior probabilities	Posterior probabilities (ev:PS1)	Posterior probabilities (ev:PS2)	Posterior probabilities (ev:PS3)	Posterior probabilities (ev:PS4)	Posterior probabilities (ev:PS5)
Shut-off-valve	0.05	0.0176	0.0568	0.0904	0.0854	0.0183
Gear pump	0.02	0.0035	0.0227	0.0362	0.0342	0.0037
Pipe 3 (pipe before the shut-off valve)	0.005	0.0088	0.0057	0.0090	0.0085	0.0091
Pipe 4 (pipe after the shut-off valve)	0.004	0.0703	0.0045	0.0072	0.0068	0.0731
Nozzle	0.0002	0.0352	0.0002	0.0004	0.0003	0.0365
Filter	0.012	0.0369	0.0239	0.0380	0.0350	0.0384
No Water (no water in the tank)	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

Table 7.3: Root nodes, their prior probabilities and their posterior probabilities using SBN. In this table ev:PS1 indicates that the evidence is pressure sensor 1, ev:PS2 indicates that the evidence is pressure sensor 2, ev:PS3 indicates that the evidence is pressure sensor 3, ev:PS4 indicates that the evidence is pressure sensor 4, ev:PS5 indicates that the evidence is pressure sensor 5.

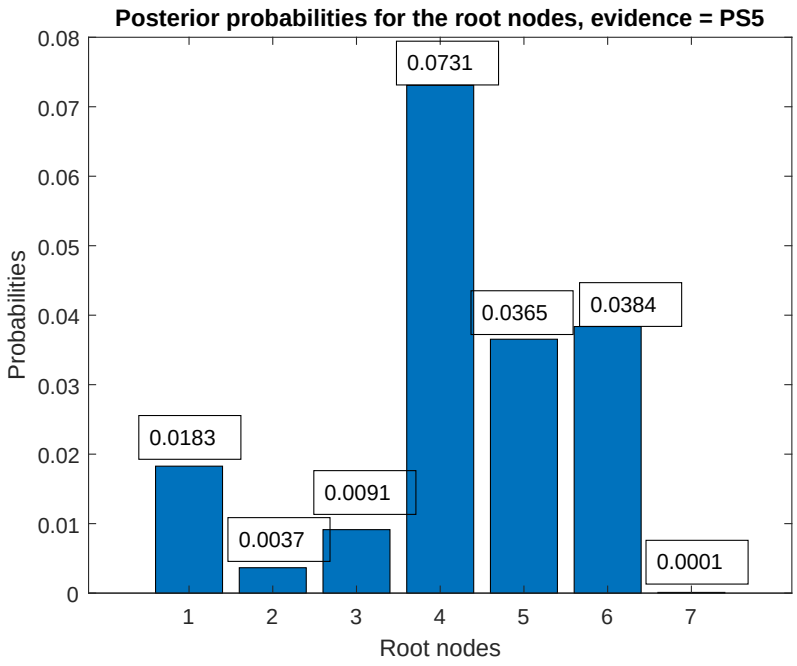

Figure 7.18: Posterior probabilities for the root nodes when pressure sensor 5 is the considered evidence to the network. In this figure 1 represents the S-O-V, 2 represents the gear pump, 3 represents the pipe 3, 4 represents the pipe 4, 5 represents the nozzle, 6 represents the filter and 7 represents no water in the tank.

causes of the intermittent fault. Although, the S-O-V and gear pump are detected with PS2, PS3 and PS4 only.

Furthermore, Figure (7.23) shows that the prior probability of the nozzle was increased with all evidence except PS2. However, the probability of the nozzle being the root cause of intermittent fault was increase significantly with PS1 and PS5.

Finally, Figure (7.25) shows that the ″No ware″ prior probability didn't change with any of the evidence. This figure demonstrates that this node does not have a direct relation to intermittent fault and it cannot increase the posterior probability of this node.

Finally, the presented simulation results show that the proposed SBN is not successful to detect intermittent fault and when the propagated evidence or sensor reading in the network is far from the fault location, then it may detect false root causes for intermittent fault.

Hence, although SBN is computationally faster than DBN or HDBN and perhaps are easier to develop and maintain due to significantly fewer parameters involved, because of the nature of the intermittent faults, for the better results, a continuous detection algorithm, HDBN, is also presented in this chapter.

7.5.3 Posterior probabilities using HDBN

The posterior probabilities of the root nodes, S-O-V, gear pump, pipe 3, pipe 4, nozzle, filter and intermittent fault when pressure sensors are given as the evidence to the HDBN are presented in Figures (7.26-7.30).

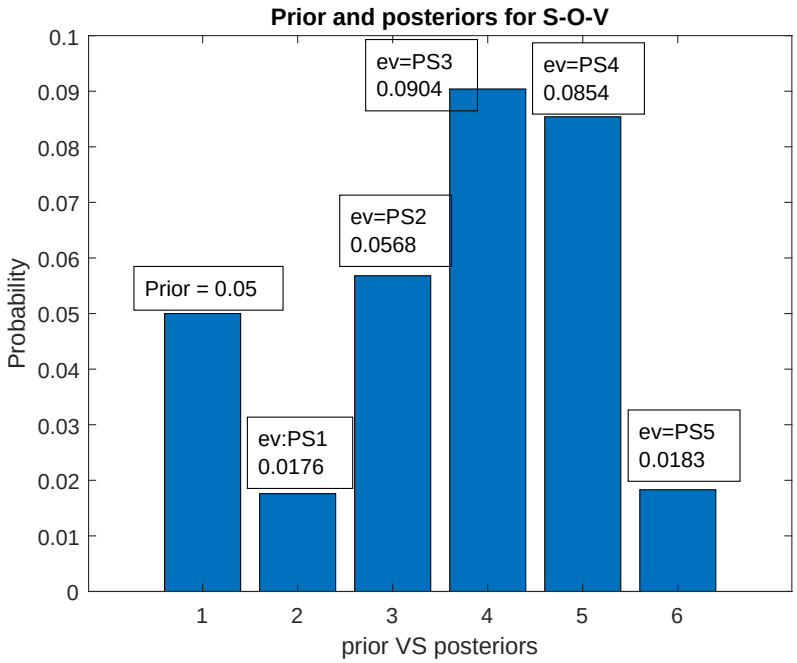

Figure 7.19: Prior and posterior probabilities for the S-O-V. In this figure 1 indicates the prior probability of the S-O-V, 2 indicates the posterior probability of the S-O-V when PS1 is the evidence to the SBN, 3 indicates the posterior probability of the S-O-V when PS2 is the evidence to the SBN, 4 indicates the posterior probability of the S-O-V when PS3 is the evidence to the SBN, 5 indicates the posterior probability of the S-O-V when PS4 is the evidence to the SBN and 6 indicates the posterior probability of the S-O-V when PS5 is the evidence to the SBN,

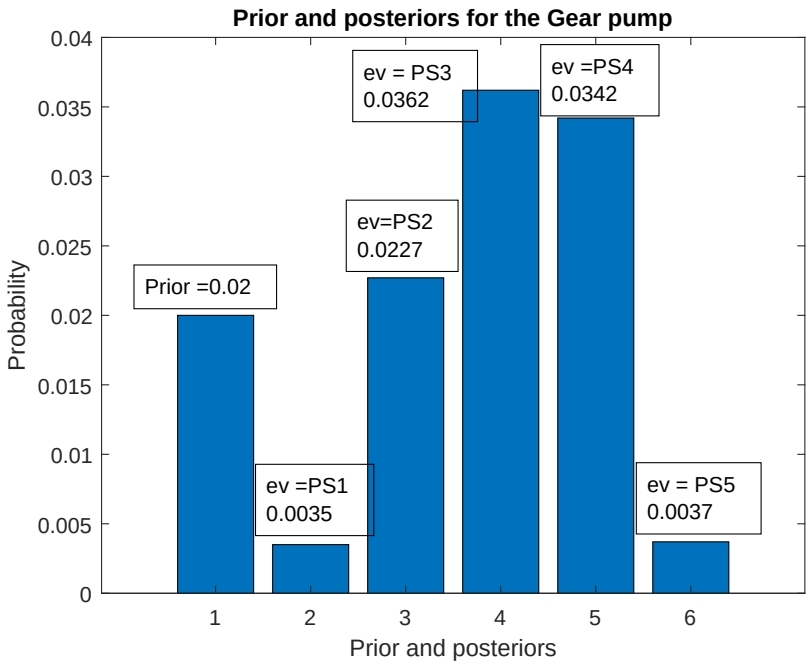

Figure 7.20: Prior and posterior probabilities for the gear pump. In this figure 1 indicates the prior probability of the gear pump, 2 indicates the posterior probability of the gear pump when PS1 is the evidence to the SBN, 3 indicates the posterior probability of the gear pump when PS2 is the evidence to the SBN, 4 indicates the posterior probability of the gear pump when PS3 is the evidence to the SBN, 5 indicates the posterior probability of the gear pump when PS4 is the evidence to the SBN and 6 indicates the posterior probability of the gear pump when PS5 is the evidence to the SBN,

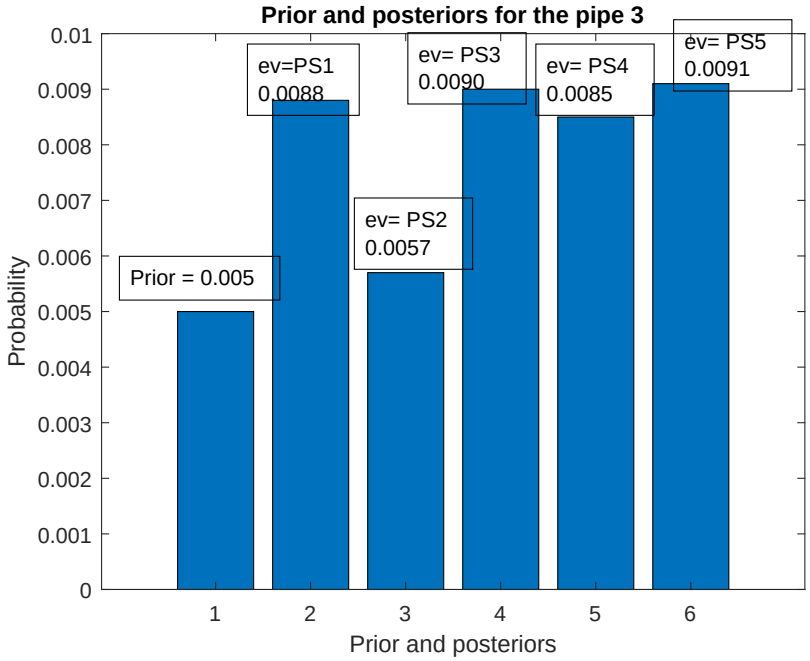

Figure 7.21: Prior and posterior probabilities for the pipe 3. In this figure 1 indicates the prior probability of the pipe 3, 2 indicates the posterior probability of the pipe 3 when PS1 is the evidence to the SBN, 3 indicates the posterior probability of the pipe 3 when PS2 is the evidence to the SBN, 4 indicates the posterior probability of the pipe 3 when PS3 is the evidence to the SBN, 5 indicates the posterior probability of the pipe 3 when PS4 is the evidence to the SBN and 6 indicates the posterior probability of the pipe 3 when PS5 is the evidence to the SBN,

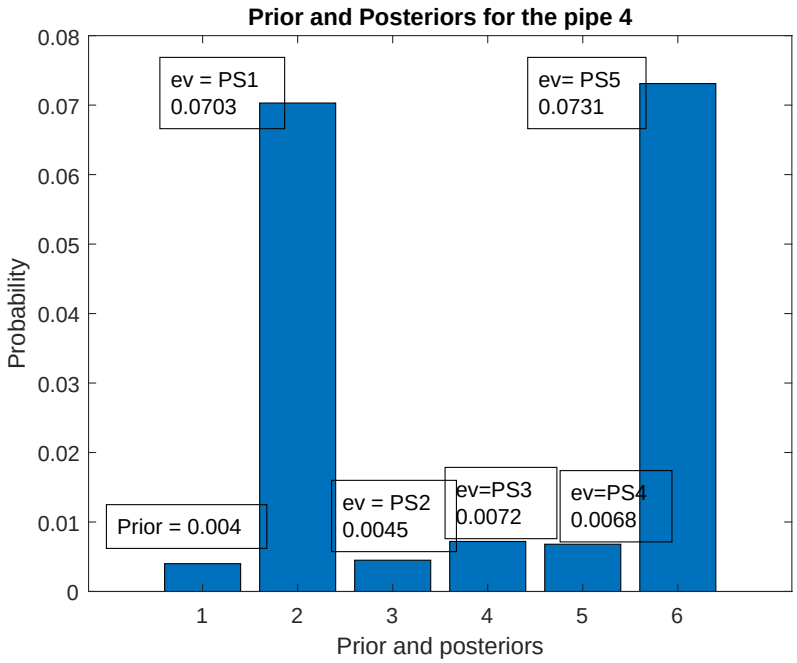

Figure 7.22: Prior and posterior probabilities for the pipe 4. In this figure 1 indicates the prior probability of the pipe 4, 2 indicates the posterior probability of the pipe 4 when PS1 is the evidence to the SBN, 4 indicates the posterior probability of the pipe 4 when PS2 is the evidence to the SBN, 4 indicates the posterior probability of the pipe 4 when PS3 is the evidence to the SBN, 5 indicates the posterior probability of the pipe 4 when PS4 is the evidence to the SBN and 6 indicates the posterior probability of the pipe 4 when PS5 is the evidence to the SBN,

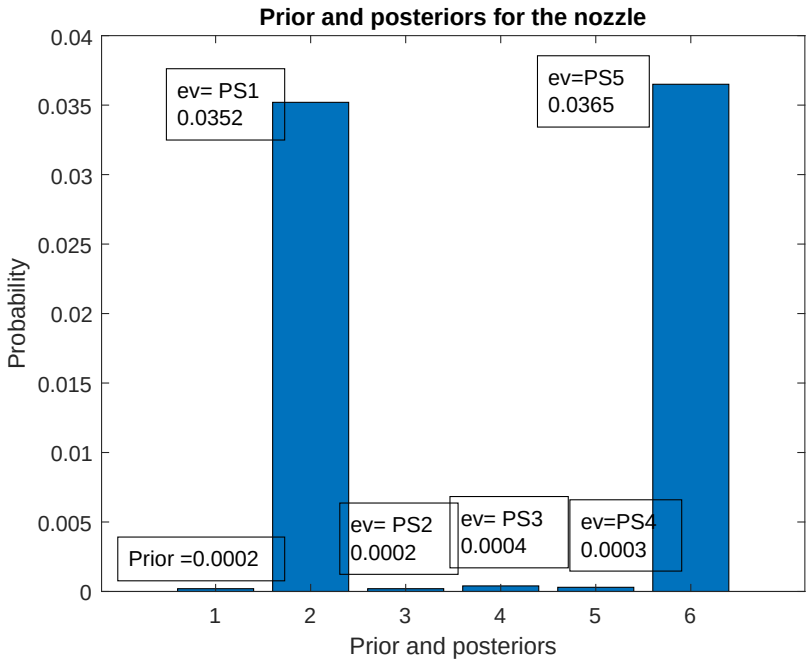

Figure 7.23: Prior and posterior probabilities for the nozzle. In this figure 1 indicates the prior probability of the nozzle, 2 indicates the posterior probability of the nozzle when PS1 is the evidence to the SBN, 3 indicates the posterior probability of the nozzle when PS2 is the evidence to the SBN, 4 indicates the posterior probability of the nozzle when PS3 is the evidence to the SBN, 5 indicates the posterior probability of the nozzle when PS4 is the evidence to the SBN and 6 indicates the posterior probability of the nozzle when PS5 is the evidence to the SBN,

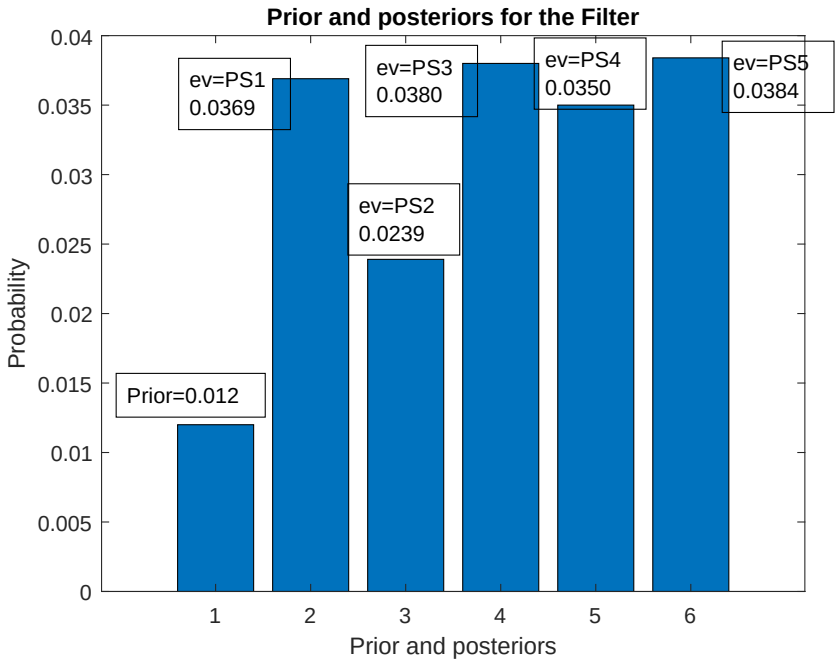

Figure 7.24: Prior and posterior probabilities for the filter. In this figure 1 indicates the prior probability of the filter, 2 indicates the posterior probability of the filter when PS1 is the evidence to the SBN, 4 indicates the posterior probability of the filter when PS2 is the evidence to the SBN, 4 indicates the posterior probability of the filter when PS3 is the evidence to the SBN, 5 indicates the posterior probability of the filter when PS4 is the evidence to the SBN and 6 indicates the posterior probability of the filter when PS5 is the evidence to the SBN,

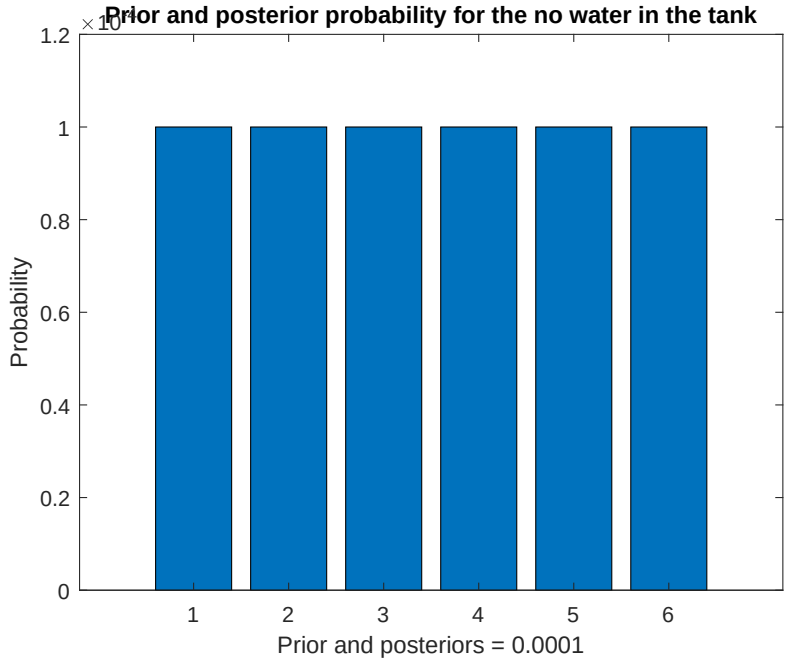

Figure 7.25: Prior and posterior probabilities for the no water in the tank. In this figure 1 indicates the prior probability of the no water in the tank, 2 indicates the posterior probability of the no water in the tank when PS1 is the evidence to the SBN, 4 indicates the posterior probability of the no water in the tank when PS2 is the evidence to the SBN, 4 indicates the posterior probability of the no water in the tank when PS3 is the evidence to the SBN, 5 indicates the posterior probability of the no water in the tank when PS4 is the evidence to the SBN and 6 indicates the posterior probability of the no water in the tank when PS5 is the evidence to the SBN,

Sensors	S-O-V	Gear pump	Pipe 3	Pipe 4	Nozzle	Filter
PS5	9.11×10^{-6}	4.9×10^{-7}	1.7×10^{-6}	6.67×10^{-6}	3.08×10^{-7}	8.31×10^{-7}
PS4	1.52×10^{-5}	7.19×10^{-7}	1.7×10^{-6}	6.3×10^{-6}	3×10^{-7}	6.7×10^{-7}
PS3	1.6×10^{-5}	8.3×10^{-7}	1.9×10^{-6}	8.3×10^{-6}	3.6×10^{-7}	8.3×10^{-7}
PS2	8.7×10^{-6}	4.8×10^{-7}	1.3×10^{-6}	8.3×10^{-6}	3×10^{-7}	4.6×10^{-7}
PS1	7.9×10^{-6}	4×10^{-7}	9.9×10^{-7}	4×10^{-6}	2×10^{-7}	4.2×10^{-7}

Table 7.4: Root nodes and intermittent fault prior probabilities using HDBN. In this table, PS1 indicates the pressure sensor 1, PS2 indicates the pressure sensor 2, PS3 indicates the pressure sensor 3, PS4 indicates the pressure sensor 4, and PS5 indicates the pressure sensor 5. The prior probabilities are learned from training data.

These prior and the posterior probabilities are also summarized in Tables (7.4) and (7.5) respectively.

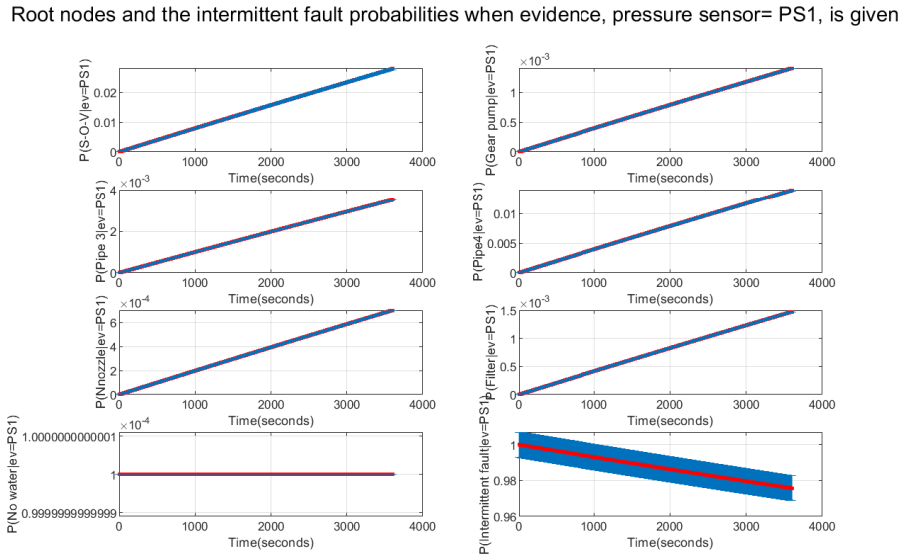

Figure 7.26: The posterior probability of the root nodes and intermittent fault when the pressure sensor 1 is the evidence to the HDBN.

These figures clearly show that for all five evidences propagated to the HDBN, the S-O-V has been detected as the main root cause of intermittent fault. However, the evidences PS3 and PS4 will increase the posterior probability of S-O-V significantly, because they are closer to the fault location.

Moreover, intermittent fault has higher probability when evidences PS3, PS4 and PS5 were propagated to the HDBN. Theses sensors are closer to the fault location and hence, show more changes.

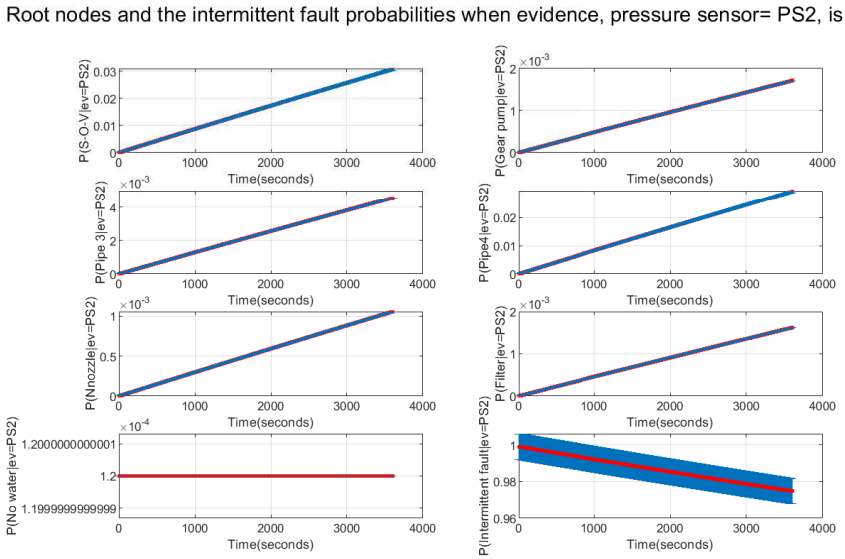

Figure 7.27: The posterior probability of the root nodes and intermittent fault when the pressure sensor 2 is the evidence to the HDBN.

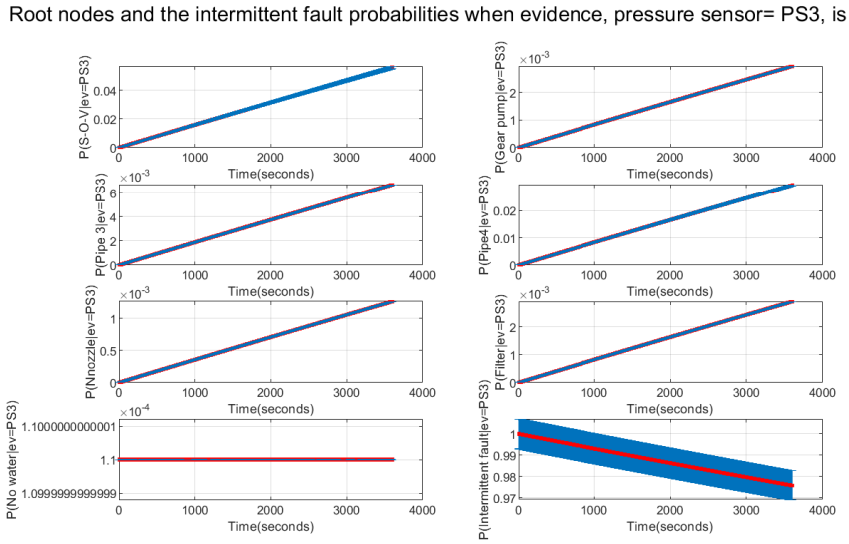

Figure 7.28: The posterior probability of the root nodes and intermittent fault when the pressure sensor 3 is the evidence to the HDBN.

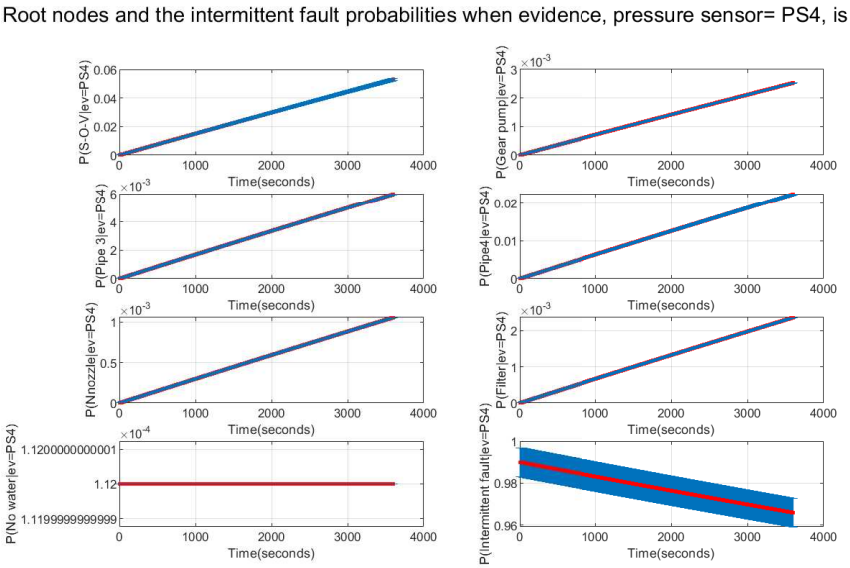

Figure 7.29: The posterior probability of the root nodes and intermittent fault when the pressure sensor 4 is the evidence to the HDBN.

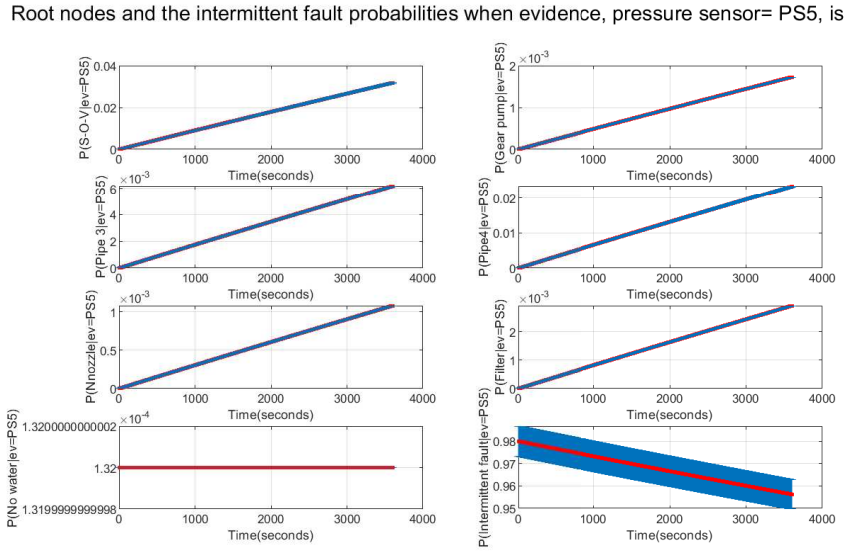

Figure 7.30: The posterior probability of the root nodes and intermittent fault when the pressure sensor 5 is the evidence to the HDBN.

Sensors	S-O-V	Gear pump	Pipe 3	Pipe 4	Nozzle	Filter	Intermittent fault
PS5	0.03176	0.001725	0.006176	0.02336	0.001082	0.002916	0.9562
PS4	0.05294	0.002525	0.005966	0.02239	0.001054	0.002356	0.9659
PS3	0.05578	0.002949	0.006668	0.0294	0.00126	0.002915	0.9757
PS2	0.3069	0.001712	0.0045	0.0294	0.001054	0.001622	0.9747
PS1	0.0279	0.001404	0.003509	0.014	0.0007024	0.001475	0.9757

Table 7.5: Root nodes and intermittent fault posterior probabilities. In this table, PS1 indicates the pressure sensor 1, PS2 indicates the pressure sensor 2, PS3 indicates the pressure sensor 3, PS4 indicates the pressure sensor 4, and PS5 indicates the pressure sensor 5.

7.5.4 Sensitivity analysis

A combination of the data-driven and knowledge-based approaches were utilised in this study for the construction and evaluation of the proposed HDBN model (the HDBN structure) to identify the main root causes of intermittent fault.

The emerging root causes were detected using the bottom-up inference in the proposed HDBN (Sedighi, 2019b). Theses probabilities explore the relationships between different nodes and their roles individually in spreading intermittent fault, particularly to identify those which have the highest impact.

Next, to provide a measure of the severity of intermittent fault the top-bottom inference was performed (sensitivity analysis).

Outcomes are then given to show the applicability of HDBN to diagnose the high-impact root-causes of intermittent fault in the fuel rig system.

The following steps were performed in order to achieve the presented results for the proposed sensitivity analysis:

- After that the HDBN graph was constructed and the prior probabilities were elicited from the available literature and data, the HDBN updated the prior probabilities every time that new information (evidence) was propagated to the network.

- Each node has the upper, estimated and lower bounds. So the proposed HDBN was running several times to compute the posterior probabilities of the interested variables for all three bounds of each node and different node combinations (different scenarios). This computation allows handling the uncertainty due to the lack of information in selecting the priors (See Tables 7.6).

There are seven root nodes in the proposed HDBN (Figure 7.5). The root node, no water, is always constant with the very low probability of 0.0001, hence, it was neglected for sensitivity analysis. The prior probabilities of the other root nodes are a function of the time and have been changed in a bounded range (higher bound, estimated, lower bound). The possible combinations for all the six root nodes in these three different ranges are equal to $6^3 = 216$. Hence, the HDBN has been run for 216 times and at each time the components with the higher posterior probabilities have been ranked. Finally, the components which have been selected the most are considered as the sensitivity set.

Because of the lack of expert knowledge and reliable information and enough historical data for intermittent fault in this experimental fuel rig system, the lower bound has been considered as (estimated priors $*$ 0.1) and the higher bound is considered as (estimated priors $*$ 10) for all nodes. Table (7.6) shows the different combinations of priors for all six root nodes.

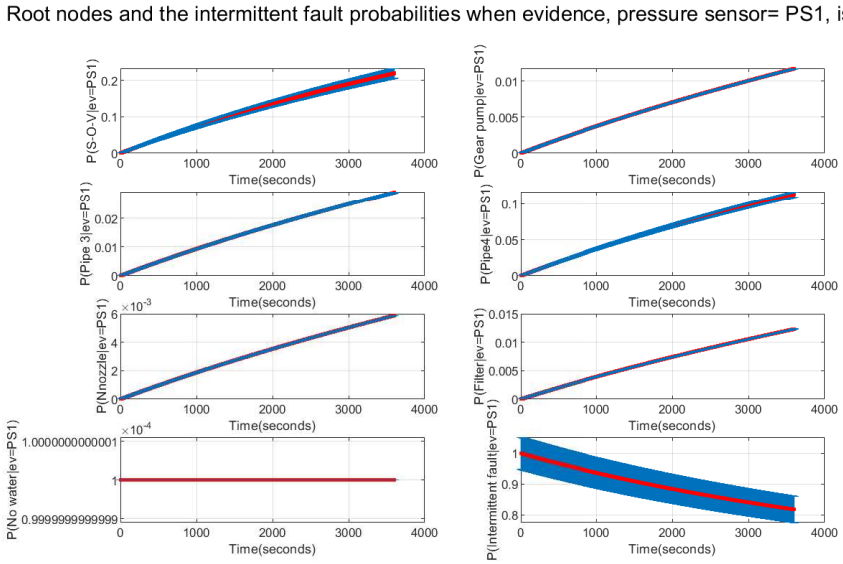

Figure 7.31: The probability of the root nodes and intermittent fault when the pressure sensor 1 is the evidence to the HDBN and the root nodes prior probabilities are in their higher bound.

The posterior probabilities of the root nodes, S-O-V, gear pump, pipe 3, pipe 4, nozzle, filter and intermittent faults when the pressure sensors data were propagated to the proposed HDBN as evidences and their prior probabilities were in their higher and lower bounds are presented in Figures (7.31-7.35) and (7.36-7.40) respectively.

These results are also summarized in Tables (7.7) and (7.8) respectively.

The obtained results from Figures (7.31-7.40) and Tables (7.7-7.8) demonstrate that the sensitivity set for this system is:

$$Sensitivity\ \ set = [S-O-V,\ \ Pipe3,\ \ Pipe4], \tag{7.2}$$

because these nodes had the highest posterior probabilities with regard to their priors in most of the combinations presented in Table (7.6).

Figure (7.41) and Table (7.10) show that if the S-O-V as the first component identified with the sensitivity analysis, become intermittently faulty, how the different pressure sensors will react. In fact, in this figure the top-bottom inference of the proposed HDBN is performed and the figure shows that the sensors closer to the S-O-V in the experimental fuel rig system will show greater effect rather than other sensors.

Figure (7.42) and Table (7.42) show the pressure sensors' reactions when S-O-V, pipe 3 and pipe 4 are intermittently faulty at the same time.

S-O-V	G Pump	Pipe 3	Pipe 4	Nozzle	Filter
E	E	E	E	E	E
H	E	E	E	E	E
H	H	E	E	E	E
H	H	H	E	E	E
H	H	H	H	E	E
H	H	H	H	H	E
H	H	H	H	H	H
L	E	E	E	E	E
L	L	E	E	E	E
L	L	L	E	E	E
L	L	L	L	E	E
L	L	L	L	L	E
E	H	E	E	E	E
E	H	H	E	E	E
E	H	H	H	E	E
E	H	H	H	H	E
E	H	H	H	H	H
E	L	E	E	E	E
E	L	L	E	E	E
E	L	L	L	E	E
E	L	L	L	L	E
E	L	L	L	L	L
E	E	L	L	L	L
E	E	H	H	H	H
E	E	L	L	L	E
E	E	H	H	H	E
E	E	L	L	E	E
E	E	H	H	E	E
E	E	L	E	E	E
E	E	H	E	E	E
E	E	E	E	E	L
E	E	E	E	E	H
⋮	⋮	⋮	⋮	⋮	⋮
L	L	L	L	L	L

Table 7.6: Different combinations for the prior probabilities of the root nodes to define the sensitivity set. In this table, E indicates the estimated prior probability, H indicates the higher bound of the prior probability and L indicates the lower bound of the prior probability.

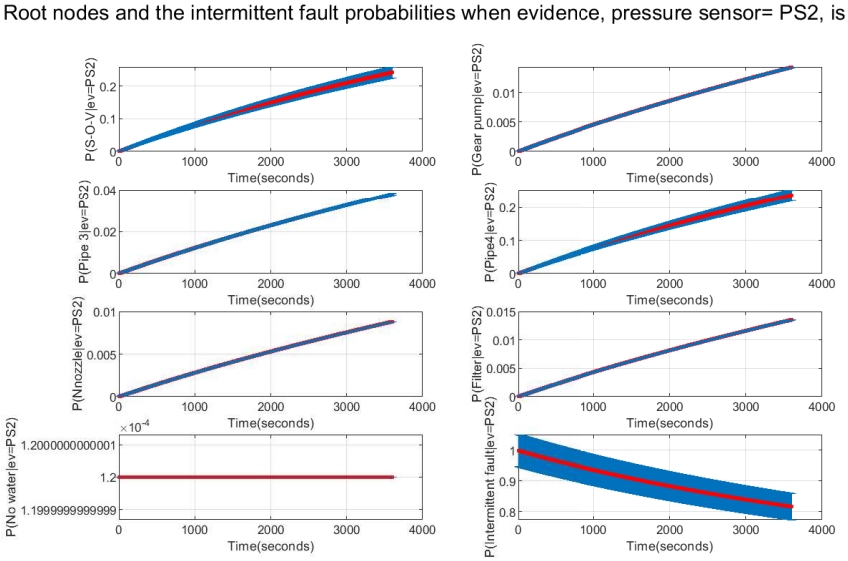

Figure 7.32: The probability of thereto nodes and intermittent fault when the pressure sensor 2 is the evidence to the HDBN and the root nodes prior probabilities are in their higher bound.

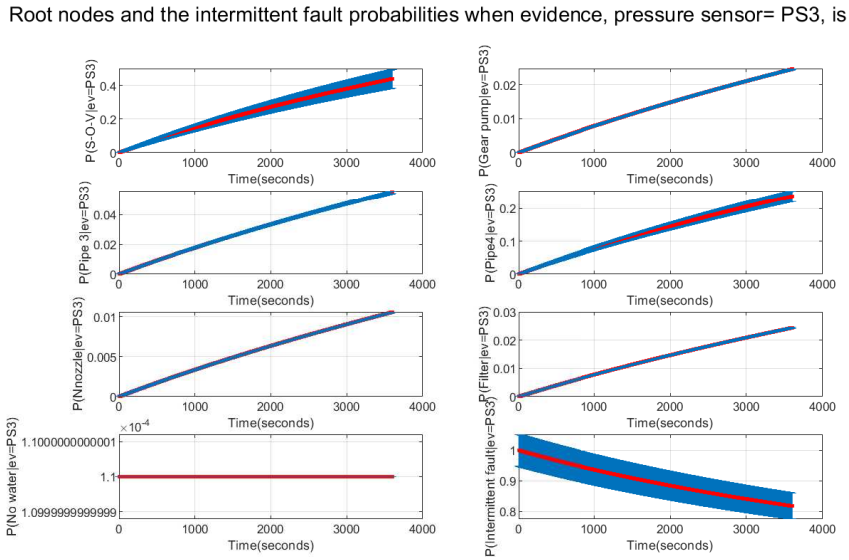

Figure 7.33: The probability of the root nodes and intermittent fault when the pressure sensor 3 is the evidence to the HDBN and the root nodes prior probabilities are in their higher bound.

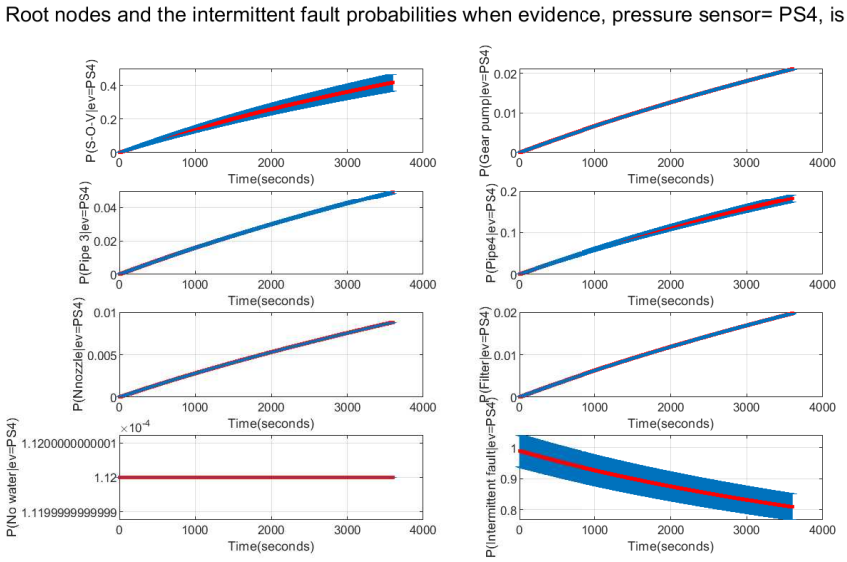

Figure 7.34: The probability of the root nodes and intermittent fault when the pressure sensor 4 is the evidence to the HDBN and the root nodes prior probabilities are in their higher bound.

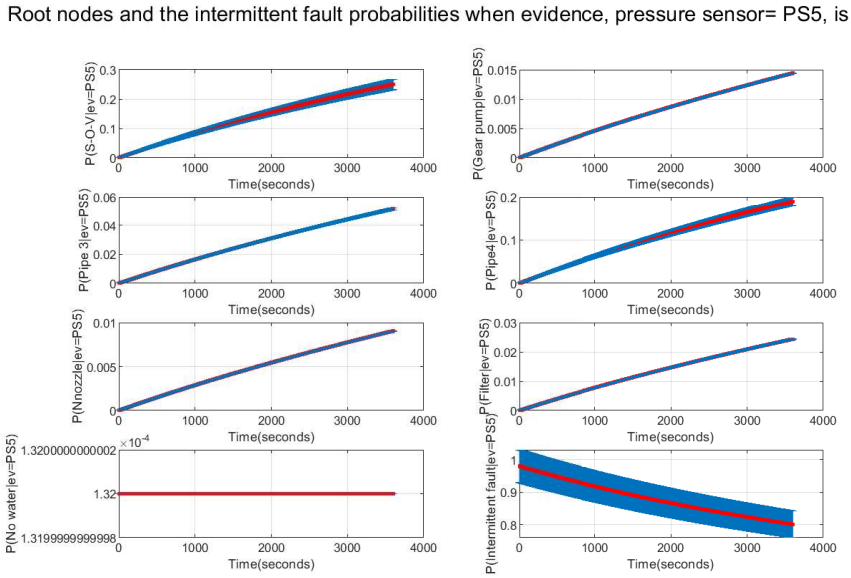

Figure 7.35: The probability of the root nodes and intermittent fault when the pressure sensor 5 is the evidence to the HDBN and the root nodes prior probabilities are in their higher bound.

Sensors	S-O-V	Gear pump	Pipe 3	Pipe 4	Nozzle	Filter	Intermittent fault
PS5	0.25	0.0144	0.0513	0.1897	0.0090	0.0244	0.8013
PS4	0.4167	0.0211	0.0496	0.1818	0.0088	0.0197	0.8095
PS3	0.4386	0.0246	0.0554	02386	0.0106	0.0244	0.8176
PS2	0.2412	0.0143	0.0379	0.2386	0.0088	0.0135	0.8168
PS1	0.2193	0.0117	0.0292	0.1136	0.0059	0.0123	0.8176

Table 7.7: Root nodes and intermittent fault posterior probabilities when their priors are in their higher bounds. In this table PS1 indicates the pressure sensor 1, PS2 indicates the pressure sensor 2, PS3 indicates the pressure sensor 3, PS4 indicates the pressure sensor 4, and PS5 indicates the pressure sensor 5,

Sensors	S-O-V	Gear pump	Pipe 3	Pipe 4	Nozzle	Filter	Intermittent fault
PS5	0.0033	0.0002	0.0006	0.0024	0.0001	0.0003	0.9775
PS4	0.0055	0.0003	0.0006	0.0023	0.0001	0.0002	0.9875
PS3	0.0057	0.0003	0.0007	0.0030	0.0001	0.0003	0.9975
PS2	0.0032	0.0002	0.0005	0.0030	0.0001	0.0002	0.9965
PS1	0.0029	0.0001	0.0004	0.0014	0.0001	0.0002	0.9975

Table 7.8: Root nodes and intermittent fault posterior probabilities when their priors are in their lower bounds. In this table PS1 indicates the pressure sensor 1, PS2 indicates the pressure sensor 2, PS3 indicates the pressure sensor 3, PS4 indicates the pressure sensor 4, and PS5 indicates the pressure sensor 5.

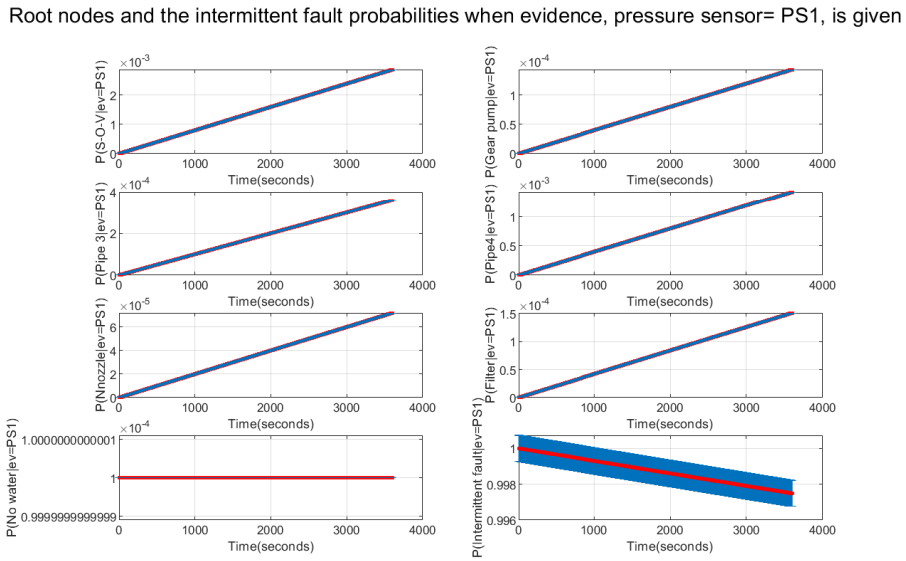

Figure 7.36: The probability of the root nodes and intermittent fault when the pressure sensor 1 is the evidence to the HDBN and the root nodes prior probabilities are in their lower bound.

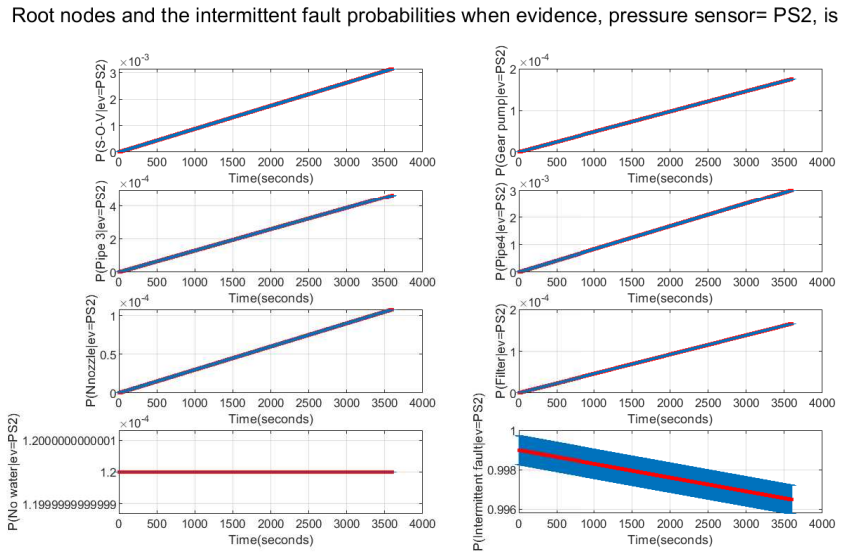

Figure 7.37: The probability of the root nodes and intermittent fault when the pressure sensor 2 is the evidence to the HDBN and the root nodes prior probabilities are in their lower bound.

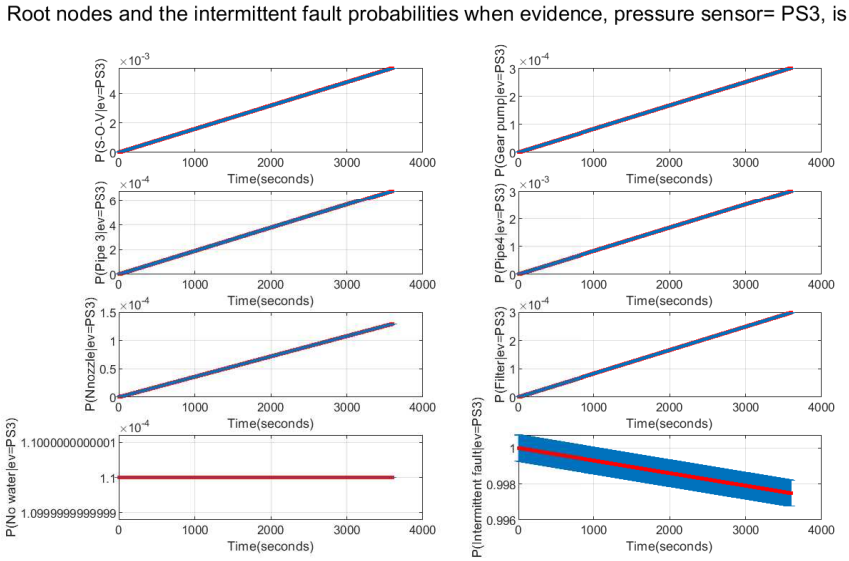

Figure 7.38: The probability of the root nodes and intermittent fault when the pressure sensor 3 is the evidence to the HDBN and the root nodes prior probabilities are in their lower bound.

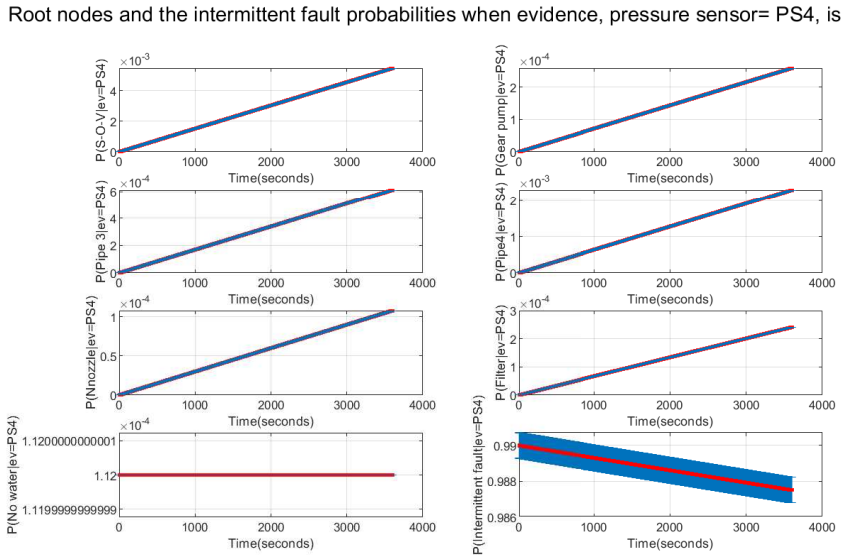

Figure 7.39: The probability of the root nodes and intermittent fault when the pressure sensor 4 is the evidence to the HDBN and the root nodes prior probabilities are in their lower bound.

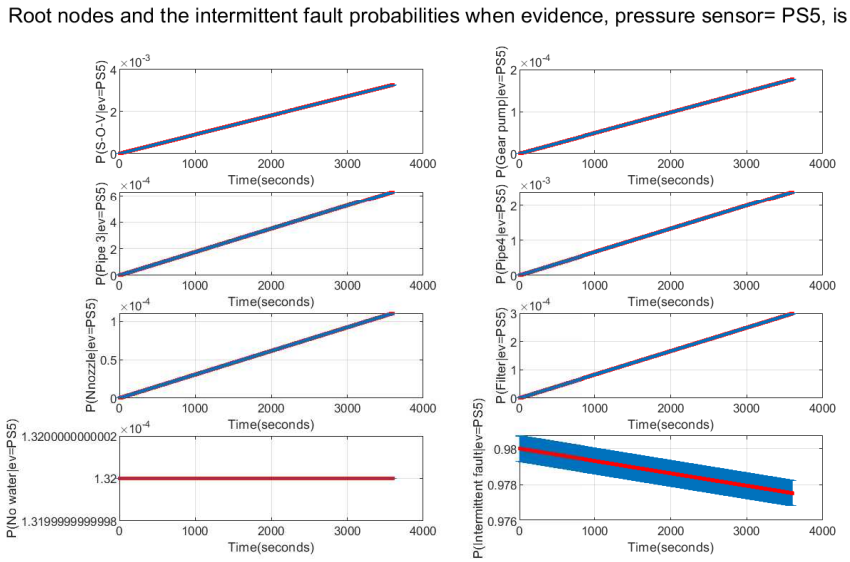

Figure 7.40: The probability of the root nodes and intermittent fault when the pressure sensor 5 is the evidence of to the HDBN and the root nodes prior probabilities are in their lower bound.

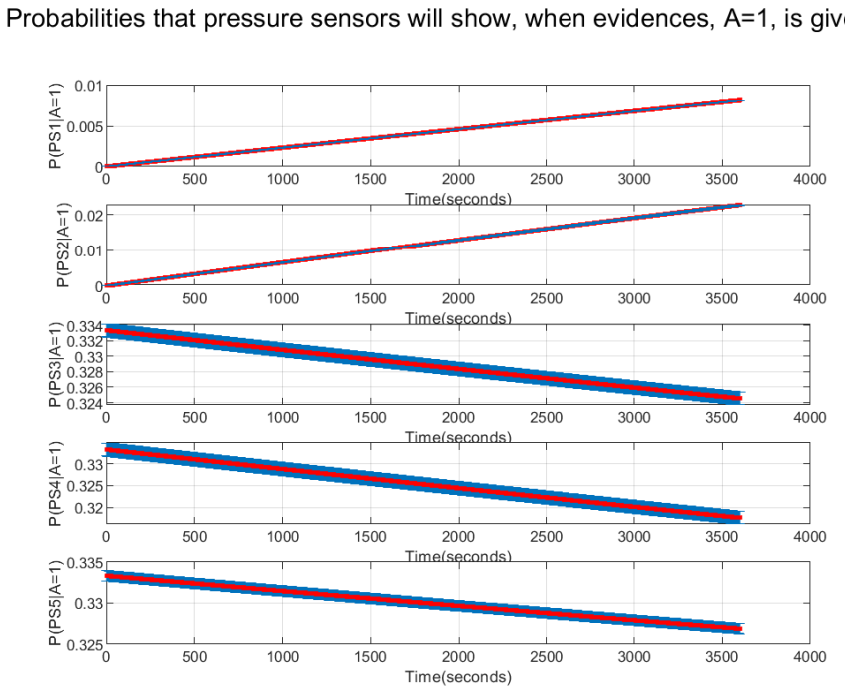

Figure 7.41: The posterior probability of each pressure sensor when the A: S-O-V, on its first state (being faulty), is the evidence to the HDBN.

Condition(s)	PS1	PS2	PS3	PS4	PS5
S-O-V: 1	0.008164	0.02279	0.3245	0.3176	0.3269

Table 7.9: The posterior probability of all five pressure sensors when the S-O-V is in its first state (being faulty). In this table PS1 indicates the pressure sensor 1, PS2 indicates the pressure sensor 2, PS3 indicates the pressure sensor 3, PS4 indicates the pressure sensor 4, and PS5 indicates the pressure sensor 5.

Condition(s)	PS1	PS2	PS3	PS4	PS5
S-O-V: 1 Pipe 3 =1 Pipe 4 =1	0.008164	0.02266	0.3246	0.3176	0.3269

Table 7.10: The posterior probability of all five pressure sensors when the S-O-V, pipe 3 and pipe 4 are in their first state (being faulty). In this table PS1 indicates the pressure sensor 1, PS2 indicates the pressure sensor 2, PS3 indicates the pressure sensor 3, PS4 indicates the pressure sensor 4, and PS5 indicates the pressure sensor 5.

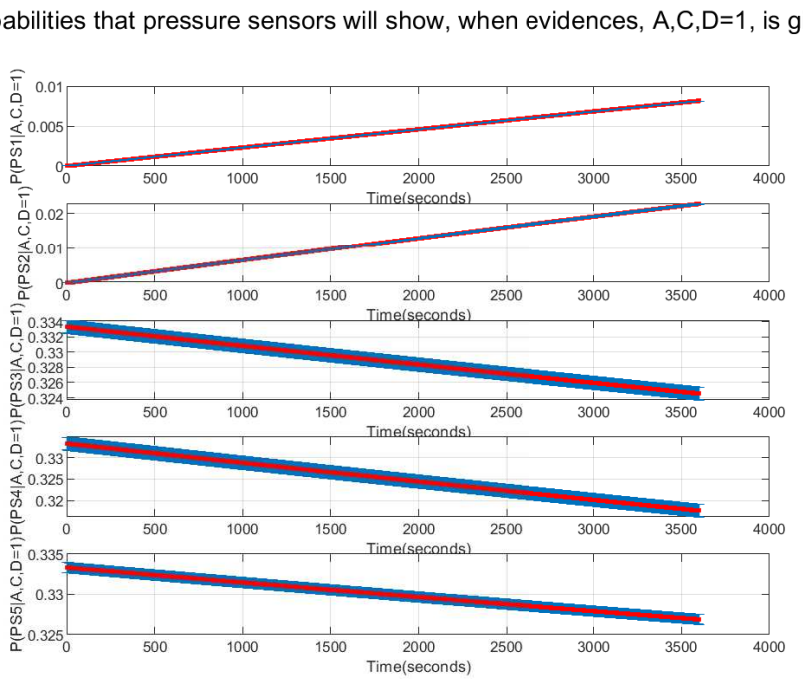

Figure 7.42: The posterior probability of each pressure sensor when the sensitivity set, A: S-O-V, C: pipe 3 and D: pipe 4, on their first state (being faulty), are the evidences to the HDBN.

Moreover, Figure (7.11) and Table (7.11) show the posterior probabilities of the pressure sensors when S-O-V, pipe 3 and pipe 4 are in their second states (not faulty) and other root nodes, gear pump, nozzle and the filter are in their first states (are faulty).

Although, in these figures, the posterior probability of the PS1, PS2 and PS3 decreased, the value of their probabilities are still significantly considerable.

Figure (7.44) shows the posterior probability of the pressure sensors when all the root nodes are in their second states and they are healthy. This figure shows that the probability that any of the sensors detect intermittent fault in the system under investigation is almost zero. As it is shown in Figure (7.44), when all the root nodes are in their healthy states and there is no fault in the systems, then all the pressure sensors show no evidence of intermittent fault in the system.

7.6 Conclusions

The presented model in this chapter is able to detect and isolate intermittent fault in a system. The model demonstrates the influence of key components upon the risk of intermittent fault in the fuel rig system.

The presented results from SBN and HDBN, suggest that S-O-V along with pipe 3 and pipe 4 are the main risk factors of intermittent fault.

Moreover, it is concluded that applying the HDBN modelling approach that endogenous uncertainty, can act as a better decision support tool for intermittent fault detection

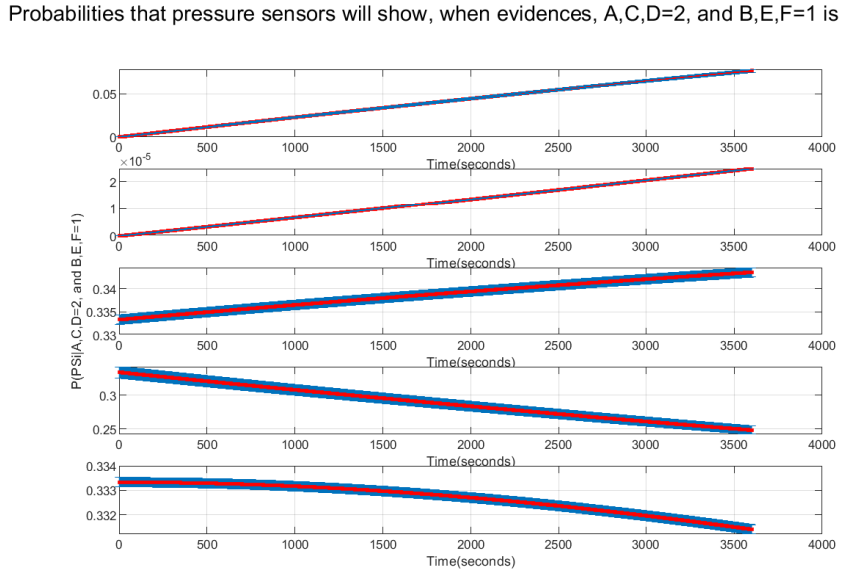

Figure 7.43: The posterior probability of all five pressure sensors when, A: S-O-V, C: pipe 3 and D: pipe 4 are on their second state (not faulty) and the root nodes B: gear pump, E; nozzle and F: filter are on their first state (being faulty) are the evidence to the HDBN.

condition(s)	PS1	PS2	PS3	PS4	PS5
S-O-V: 2 Pipe 3 =2 Pipe 4 =2 G Pump =1 Nozzle =1 Filter=1	0.07653	2.4×10^{-5}	0.3435	0.2485	0.3314

Table 7.11: The posterior probability of all five pressure sensors when S-O-V, pipe 3 and pipe 4 are in their second states (not faulty) and other root nodes, gear pump, nozzle and the filter are in their first states (are faulty). In this table PS1 indicates the pressure sensor 1, PS2 indicates the pressure sensor 2, PS3 indicates the pressure sensor 3, PS4 indicates the pressure sensor 4, and PS5 indicates the pressure sensor 5.

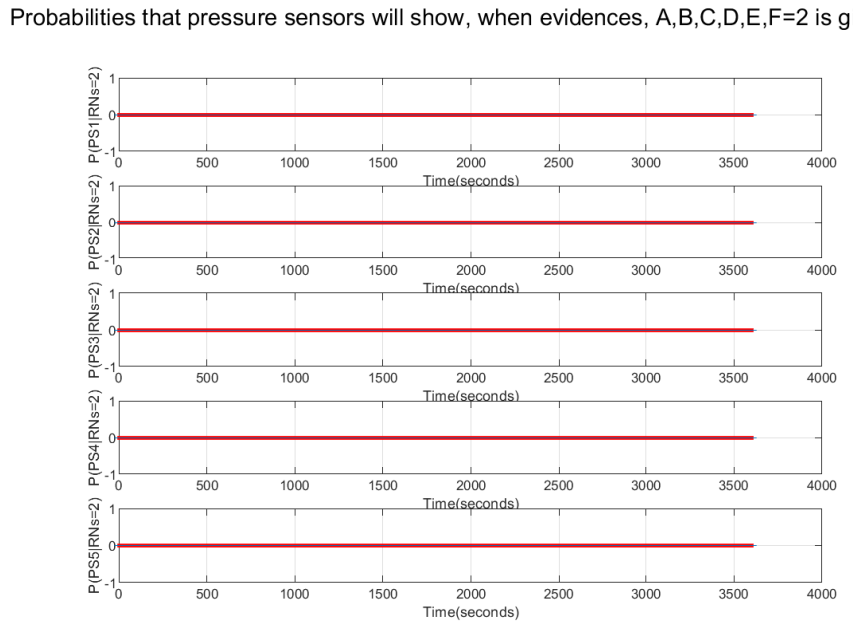

Figure 7.44: The posterior probability of each pressure sensor when all the root nodes are on their second states (no fault) are the evidence to the HDBN.

and isolation than SBN.

Some of the characteristics of the HDBN method over other methods are listed as follows:

- Solution confidence: Since the HDBN method deals with the probability distributions it has the ability to perform multiple solutions in the solution space. Moreover, the proposed HDBN explored and displayed casual and complex relationships between key factors and final outcomes in a straight-forward and understandable manner.
- Nonlinearity: The HDBN can model nonlinear dependence between nodes. So, the nonlinear relations between the pressure sensor readings, intermittent faults and root causes are modelled by the proposed HDBN method.
- Feedback loop and complexity: By modelling time dynamic, HDBN can accommodate feedback loops which makes it capable of capturing complex relationships which is useful for explaining unobserved nodes, may exist in the pathway between observed inner and outer nodes.
- No discretization: The HDBN, as well as DBN, does not require the discretization of continuous variables which can cause several issues including missing the intermittent fault during the discretization process.
- Data from different sources: This approach works with data from the variety of sources and handles a mix of subjective and objective data and can incorporate variables which differ across the contexts.

- Uncertainty: The proposed HDBN is also used to calculate the effectiveness of the interventions where the uncertainties associated with these casual relationships at the same time.
- Hypothesis updating: HDBNs are a reasonable supplement to traditional experimental and statistical methods since, traditional reliability analysis methods cannot update hypothesises with new information added in, while HDBNs can update the system reliability when evidence is added during the analysis.

7.6.1 Limitations

In spite of the remarkable power of BNs there are some inherent limitations that some of them are listed below:

- Poorly defined states and variables may mask the impact of a particular scenario or decision (Woodberry *et al.*, 2004).
- Poor representation of initial conditions and structured networks will not provide an insight into the issue of concern (Woodberry *et al.*, 2004).
- In the absence of empirical data, causal relationships can be specified based on expert opinion. The sheer volume of questions to be answered and probabilities to be filled in will pose a considerable cognitive barrier for the expert (Woodberry *et al.*, 2004).

Furthermore, various assumptions are made in the proposed HDBN model which are a consequence of either lack of complete knowledge on how intermittent fault really works and/or poor data (uncertainty about the accuracy, provenance, method of collection, are other features of the data). However, the model could be developed to address some of these assumptions. This is likely to result in longer computational times and more resources to collect data and theories on intermittent fault mechanisms.

Nevertheless, the current model represents satisfying results in detecting and isolating intermittent faults.

Chapter 8

Gaussian Process Regression-Based Intermittent Fault Prediction

8.1 Introduction

The early and precise fault prediction in a system and knowing where faults are expected to arise play an important role in determining the health of a system, reducing maintenance costs and improving the quality of system performance.

The main challenge of fault prediction is to minimize the uncertainty on available information about the system such as operating environment and loading conditions, computational resources, and time horizon. Furthermore, the involvement of physics of failure models is also important in fault prediction.

In this chapter, the focus is on introducing a kernel-based fully Bayesian regression algorithm, known as Gaussian Process Regression (GPR) for the intermittent fault prediction in a system along with its simulation results. The main reason for selecting GPR rather than other Bayesian methods is it's faster response especially when the size of data is huge and it is a nonparametric method (Das *et al.*, 2018).

This chapter is organized as follows: in Section 8.2, GPR along with the selection of the covariance function and hyperparameters are introduced. The prediction using GPR is explained in Section 8.3 and intermittent fault prediction in the experimental fuel rig system using GPR along with their simulation results, and discussions are presented in Sections 8.4 and 8.5 respectively. Finally, conclusions and the limitations of the proposed method are presented in Section 8.6.

8.2 Gaussian process modelling

The Gaussian Process (GP) is a kernel-based probabilistic model and is described as a group of random variables that any finite subset have a joint Gaussian distribution. A

GP is defined as a random process with Gaussian correlated noise:

$$f(x) \sim gp(m(x), k(x,x')) \tag{8.1}$$

where the mean function, $m(x)$, and the kernel function $k(x,x')$ denote the expectation value $E[f(x)]$ and the covariance (kernel) $cov(f(x), f(x'))$ respectively. Moreover, the mean function $m(x)$ could be considered as a priori (Shen *et al.*, 2006; Williams & Rasmussen, 2006; Ebden *et al.*, 2008).

8.2.1 Covariance function

Consider n_m observations in an arbitrary data set, $y = y_1, \cdots, y_{n_m}$. This data set can be partnered with a GP. Very often, it is assumed that the mean of this partner GP is zero everywhere. Also the data are usually noisy, from the measurement errors. Hence, each observation y can be related to an underlying function, $f(x)$, through a Gaussian noise model:

$$y = f(x) + N(0, \sigma_{n_m}^2), \tag{8.2}$$

and regression is the search for this function, $f(x)$ where $f(x) \sim GP(0, k(.,.))$ is the corresponding underlying function. It means that y related to x nonlinearly through an unknown function $f(x)$, which, in turn, it is being approximated by a GP.

The covariance (kernel) function, $k(x,x')$ is what relates one observation to another observation.

The most commonly-used covariance (kernel) function in GPR is the squared exponential (Guo, 2011; Williams & Rasmussen, 2006; Ebden *et al.*, 2008):

$$k(x,x') = \sigma_f^2 \exp([\sum_{dim=1}^{D} \frac{-(x-x')^2}{2l^2}]) \tag{8.3}$$

where dim indicates the dimension, σ_f^2 denotes the signal variance (the maximum allowable covariance) and l presented the typical distance between two peaks (length distance). This should be high for functions which cover a broad range on the y axis. If $x \approx x'$, then $k(x,x')$ approaches this maximum, meaning $f(x)$ is nearly perfectly correlated with $f(x')$. Now if x is distant from x' then instead $k(x,x') \approx 0$, means that the two points cannot see each other. How much effect this separation has will depend on the length parameter, l, so there is much flexibility built into (8.3).

For simplicity, the error function (noise), $e_f = \sigma_n^2 \delta(x,x')$, can be folded into the covariance function, $k(x,x')$, (Williams & Rasmussen, 2006) ,

$$k(x,x') = \sigma_f^2 \exp[\frac{-(x-x')^2}{2l^2}] + \sigma_n^2 \delta(x,x'), \tag{8.4}$$

where $\delta(x,x')$ is the Kronecker delta function and σ_n presents the noise, which usually

keeps separate from $k(x,x')$.

In more complicated cases where a long-term downward trend, has some fluctuations, so a more sophisticated covariance function may present (Melkumyan & Ramos, 2009; Williams & Rasmussen, 2006; Ebden *et al.*, 2008):

$$k(x,x') = \sigma_{f_1}^2 \exp\frac{-(x-x')^2}{2l_1^2} + \sigma_{f_2}^2 \exp\frac{-(x-x')^2}{2l_2^2} + \sigma_n^2 \delta(x,x'), \tag{8.5}$$

when the first term takes into account the small vicissitudes of the dependent variable, and the second term has a longer length parameter $(l_2 \approx 6l_1)$ to represent its long-term trend.

Sometimes the function might looks as if it contain a periodic element, hence, it is needed to consider another covariance function, with a periodic element:

$$k(x,x') = \sigma_f^2 \exp(\frac{-(x-x')^2}{2l^2}) + \exp(-2\sin^2[\nu\pi(x-x')]) + \sigma_n^2 \delta(x,x'), \tag{8.6}$$

where the first term represents the hill-like trend over the long term, and the second term gives periodicity with frequency ν. In this case x and x' can be distant and yet still see each other.

Since the dependent variable might have other dynamics, hence there is no limit to how complicated $k(x,x')$ can be chosen. Covariance functions can be grown in this way to suit the complexity of the considered data, although, the reliability of the proposed regression is dependent on how well the covariance function is selected (Haranadh & Sekhar, 2008; Williams & Rasmussen, 2006; Ebden *et al.*, 2008).

General properties of covariances are controlled by the small number of hyperparameters.

Covariance structures: There are different covariance structure in mixed model analysis such as Gaussian process including variance components, autoregressive, component symmetry, unstructured and structured or Toeplitz. In the structured or Toeplitz all measurements next to each other have the same correlation, measurements two apart have the same correlation different from the first, measurements three apart have the same correlation different from the first two, etc (Kincaid, 2005).

In this thesis the defult choice for the covariance matrix is structured or Toeplitz. Some of the properties of this matrices are:

- Structured or Toeplitz matrices are positive definite.
- Structured or Toeplitz matrices commute asymptotically. This means they diagonalize in the same basis when the row and column dimension tends to infinity.
- The inverse of the nonsingular symmetric structured or Toeplitz has the representation.
- Structured or Toeplitz matrices are ubiquitous and are one of the most well-studied and understood classes of structured matrices.

- Structured or Toeplitz matrices have some of the most attractive computational properties and are amenable to a wide range of disparate algorithms.

Although, The structured matrix is not the only choice for the covariance matrices and the covariance functions can grow as much as complex as possible, in this chapter , the exponential squared covariance function has been chosen because this covariance function is infinitely differentiable, which means that the GP with this covariance function has mean square derivatives of all orders, and is thus very smooth (Williams & Rasmussen, 2006).

8.2.2 Hyperparameters

The free parameters l, σ_f, and, σ_n in covariance function, are the hyperparameters, Θ, and are important in the performance of the selected covariance function. The hyperparameters can be varied during the regression process and if they were not chosen carefully, then the outcomes are not accurate enough (Wilson & Adams, 2013; Williams & Rasmussen, 2006; Ebden *et al.*, 2008).

The length-scale, l, characterizes the distance in input space before the function value can change significantly. Short length-scales mean that the predictive variance, σ_f, can grow rapidly away from the data points. The noise, σ_n, that affects the process is supposed to be random, and so no correlation between different inputs are expected, and usually is only present on the diagonals of the covariance matrix (Sang & Huang, 2012; Williams & Rasmussen, 2006; Ebden *et al.*, 2008).

The advantage of the probabilistic GP is that the hyperparameters and covariances can be chosen directly from the training data unlike other models such as splines which requires cross-validation.

Bayes theorem tells that, assuming there is little prior knowledge about what Θ should be, this corresponds to maximizing $\log p(y|x,\Theta)$, given by:

$$\log p(y|x,\Theta) = -\frac{1}{2}y^T K^{-1} y - \frac{1}{2} log|K| - \frac{n}{2}\log 2\pi. \tag{8.7}$$

Then by running an appropriate multivariate optimization algorithm (e.g. conjugate gradients (Malandain *et al.*, 2013), Nelder-Mead simplex (Dennis & Woods, 1987), etc.) on the equation (8.7) a pretty good choice for Θ will be found. In (Williams & Rasmussen, 2006), the detailed methodologies to estimate the hyperparameters Θ are explained in more details.

8.3 Prediction using GPR

Assume that there are n_m noisy observations, y, then the objective is to predict y_*, not the actual f_* by GPR. Then the GP is used as Bayesian prior expressing beliefs about underlying function, f.

The key assumption in GPR modelling is that the data can be represented as a sample from a multivariate Gaussian distribution, then

$$\begin{pmatrix} y \\ y_* \end{pmatrix} \sim N(0, \begin{pmatrix} K & K_*^T \\ K_* & K_{**} \end{pmatrix}), \tag{8.8}$$

where K presents the covariance function and T indicates matrix transposition. From (8.8), the steps to prepare a GPR are as follows:

- First step is the covariance function selection.
- Then among all possible combinations, the findings are summarized in three matrices:

$$K = \begin{pmatrix} k(x_1,x_1) & k(x_1,x_2) & \cdots & k(x_1,x_n) \\ k(x_2,x_1) & k(x_2,x_1) & \cdots & k(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n,x_1) & k(x_n,x_2) & \cdots & k(x_n,x_n) \end{pmatrix} \tag{8.9}$$

$$K_* = \begin{pmatrix} k(x_*,x_1) & k(x_*,x_2) & \cdots & k(x_*,x_n) \end{pmatrix} \tag{8.10}$$

  and

$$K_{**} = k(x_*,x_*). \tag{8.11}$$

- Next the conditional probability $p(y_*|y)$ is calculated: given the data, how likely is a certain prediction for y_*, while the probability follows a Gaussian distribution:

$$y_*|y \sim N(K_*K^{-1}y, K_{**} - K_*K^{-1}K_*^T) \tag{8.12}$$

  Where the best estimate for y_* is the mean of the above distribution:

$$\bar{y_*} = K_*K^{-1}y \tag{8.13}$$

- Finaly, the uncertainty is estimated as follows:

$$var(y_*) = K_{**} - K_*K^{-1}K_*^{-T}. \tag{8.14}$$

  In this process, $X = [x_1, x_2, \cdots, x_{n_m}]$, x_*, y and y_* are the input training data, test point, output training data and predicted output respectively.

8.3.1 Numerical example

Assume that there are $n_m = 6$ observations, at

$$X = [-1.55 \quad -1.00 \quad -0.75 \quad -0.40 \quad -0.25 \quad 0.00]$$

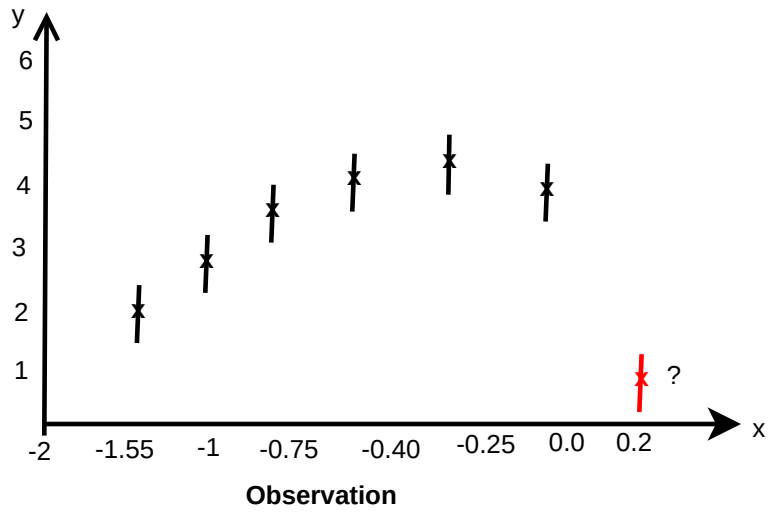

Figure 8.1: Given six noisy data points (error bars are indicated with vertical lines), the interest is in estimating the seventh at $x_* = 0.2$.

and the noise is known $\sigma_n = 0.3$. Then σ_f and l are selected as ($l = 1$ and $\sigma_f = 1.27$) and the covariance matrix is designed using (8.9):

$$K = \begin{pmatrix} 1.70 & 1.42 & 1.21 & 0.87 & 0.72 & 0.51 \\ 1.42 & 1.70 & 1.56 & 1.34 & 1.21 & 0.97 \\ 1.21 & 1.56 & 1.70 & 1.51 & 1.42 & 1.21 \\ 0.87 & 1.34 & 1.51 & 1.70 & 1.59 & 1.48 \\ 0.72 & 1.21 & 1.42 & 1.59 & 1.70 & 1.56 \\ 0.51 & 0.97 & 1.21 & 1.48 & 1.56 & 1.70 \end{pmatrix}$$

For simplicity, the error function (noise), e_f, is folded into the covariance function, $k(x,x')$, see equation (8.4).

From (8.10) and (8.11):

$$K_* = \begin{pmatrix} 0.38 & 0.79 & 1.03 & 1.35 & 1.46 & 1.58 \end{pmatrix}, \quad K_{**} = 1.70.$$

Next from (8.13) and (8.14), $\bar{y_*} = 0.95$ and $var(y_*) = 0.21$, which show the best estimation and the variance of the dependent variable at $x_* = 0.2$, the 7^{th} observation in the data set X (Figure 8.1). This was generated from a GPR with the square exponential covariance function with hyperparameters $\Theta = (1, 1.27, 0.3)$.

This method can be repeated for other points. Although, it could avoid the repetition by performing the above procedure once with suitably larger K_* and K_{**} matrices.

The obtained results will give us the estimated mean value and error bars of the predicted points for the whole period of the prediction time.

8.4 Intermittent fault prediction in the experimental fuel rig system

Fault prediction can be initiated at any time in the life of a system based on the last available state estimate (Williams & Rasmussen, 2006; Ebden *et al.*, 2008). In general, the fault prediction framework first uses the system model for fault detection to obtain information about the current state of the system and then predicts the future states of the system (Vanhatalo *et al.*, 2009; Williams & Rasmussen, 2006; Ebden *et al.*, 2008).

As explained earlier, because of the nature of IFs, their available historical data, notably when a system is newly introduced to the market, may not contain enough information to help the voted prognostic system to predict the upcoming IFs. Hence, the chosen prognostic method should be

(i) Non-parametric which will give a higher degree of freedom to the prognostic system,

(ii) Capable of dealing with lack of information, missing data and/or small data sets with the help of available reliable information such as expert knowledge,

(iii) Capable of quantifying the prediction uncertainty to make a better-informed decision.

Moreover, if a prognostic system, whether Bayesian methods or others can make use of historical data, heuristic information, and sensor readings, etc., then the prognostic system can make a better-informed decision than if the data/ information was not available.

In response to the mentioned challenges, in this research, a novel hybrid approach is designed for IF detection and prediction in a complex system (see Figure 8.2).

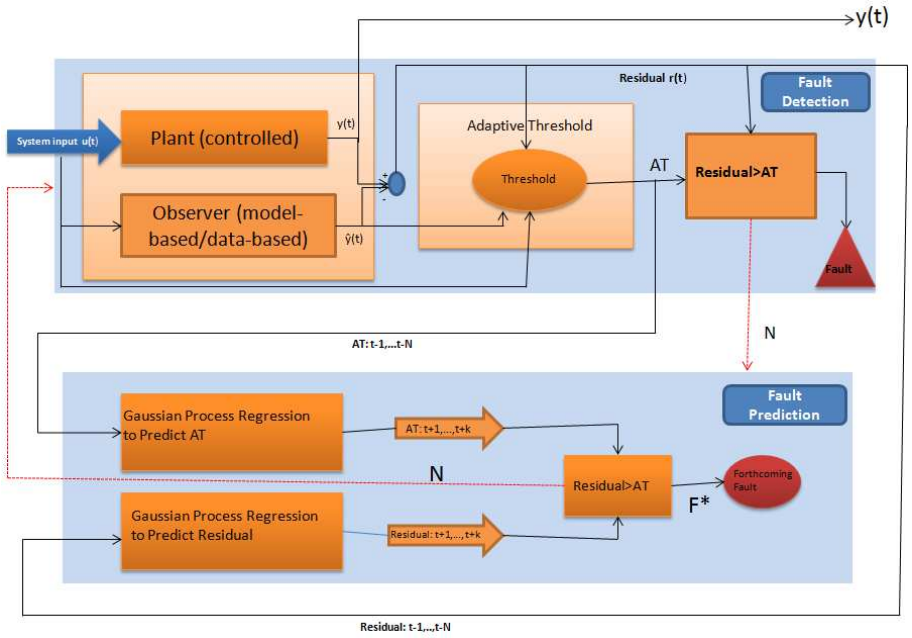

Figure 8.2: The two-step intermittent fault detection and prediction process.

In the first step the suitable NUI model-based FD model is designed to obtain the actual system response, generate the residual errors, and the corresponding adaptive threshold for on-line fault detection where t indicates the current time. The types of fault captured in this research are IFs which are difficult to detect and could provide a great deviation of the residual signals.

Moreover, when the mathematical modelling of the system is difficult and expensive due to its complexity or when the needed information is not available, the data driven-based methods such as Bayesian network is used to detect the intermittent fault in the first step.

Then, in the second step, a non-parametric Bayesian method is employed to predict the forthcoming residual error and adaptive threshold, with the estimation of the likelihood of system failure over some future time interval. So that the fault is predicted anytime that the predicted forthcoming residual has reached or passed the forthcoming predicted adaptive threshold. Next, the obtained results will be injected to the first step as new data sets to improve the real-time detection process. (Figure 8.2) .

The development of such diagnostic and prognostic system along with the demonstration of its performance in a real-world testbed, an aircraft fuel system simulation rig which simulates by hardware similarity the components of an aircraft fuel system, are presented in this chapter.

The proposed method was also verified by the presented simulation results.

8.4.1 GPR model for the experimental fuel rig system

In this chapter to support the theory of the two-step IF detection and prediction approach, the fuel rig introduced in Chapters 4 is considered as a case study. According to the presented results in Chapter 6, the mathematical modelling of the fuel rig system has been verified against its real-world system with an acceptable error and the proposed results have demonstrated that the shortest intermittent fault activation period in the fuel rig system lasts for a few seconds.

In the experimental fuel rig system, the probability density function associated with the data provided by each sensor can be characterized by a Gaussian distribution. The standard deviation of these distributions (σ) is between $0.001 - 0.03$ bar where the gear pump speed is 400 rpm at steady-state stage (Table 8.1).

Table 8.1: The standard deviations of the sensors distribution at 400rpm, (Niculita *et al.*, 2013).

Sensor	Standard Deviation (σ)
Pressure Sensor 1 (PS1)	0.014
Pressure Sensor 2 (PS2)	0.015
Pressure Sensor 3 (PS3)	0.0036
Pressure Sensor 4 (PS4)	0.02
Pressure Sensor 5 (PS5)	0.03

To keep the gear pump speed at specific speed, $400rmp$, the system is controlled by an appropriate adaptive controller.

Data	σ_f	l	σ_n
residual 1	0.11251	0.090441	0.00967
threshold 1	0.102	0.0540	0.0500
residual 2	0.2003	0.0067	0.0300
threshold 2	0.437	0.0078	0.00995
residual 3	0.437	0.3548	0.68
threshold 3	0.0342	0.00548	0.0097
residual 4	0.422	o.o485	0.85
threshold 4	0.0868	0.00654	0.076
residual 5	0.765	0.2387	0.0654
threshold 5	0.0863	0.000786	0.091

Table 8.2: The hyperparameters for residual and adaptive threshold of all five pressure sensors obtained by GPML toolbox in Matlab.

To design the GPR for the experimental fuel rig system the following steps are performed,

- Identify the number of observed sampling, n_m.

- Define the suitable covariance (Kernel) function: the smoothness of the GPR function is defined by the selected covariance function. In this chapter, the squared exponential kernel function has been selected because it is one of the most commonly used covariance functions and is the default option for GP fitting in GPML toolbox. The squared exponential kernel function is defined as:

$$k = \sigma_f^2 \exp\left(\frac{|x_1 - x_2|^2}{2l^2}\right) + e_f. \tag{8.15}$$

- Define the hyperparameters (Kernel parameters) from the training data: as mentioned earlier, the covariance function is normally parameterized by a set of kernel parameters known as hyperparameters. In this chapter, the hyperparameters for each sensor reading has been obtained using GPML toolbox in Matlab (Table 8.2) by optimization over the selected kernel function. The optimization has been repeated several times for different initializations. For example, for the residual and adaptive threshold 5, the effect of two different sets of hyperparameters are shown in Figures (8.3) and (8.33). These figures clearly show that the selected hyperparameters for the residual and adaptive threshold 5 in figure (8.33) are producing the GP Mean which is closer to the ground truth. The same process has been repeated several times to obtain the final hyperparameters presented in Table (8.2).

  These parameters along with the noise variance estimated from the training data during the GPR model training.

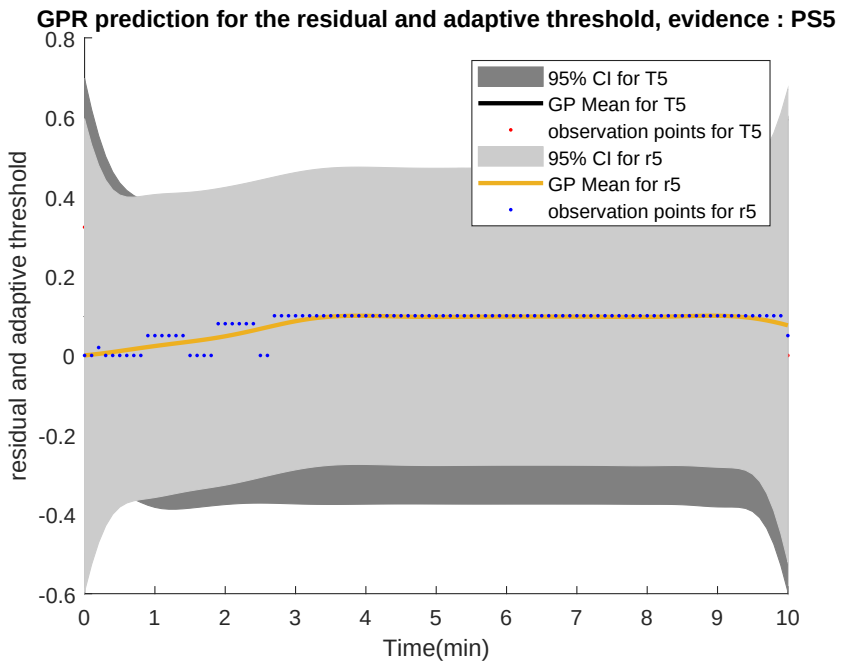

Figure 8.3: The GPR predictions for the residual and threshold 5. In this figure the grey areas are the confidence intervals, the yellow and the black solid lines are the GP mean functions for the residual and threshold respectively. The blue and the red points are the residual 5 and threshold 5 observation points. In this figure, r5 indicates the residual 5 and T5 indicates threshold 5. The hyperparameters for the residual and adaptive threshold are $\Theta_r = \{\sigma_f = 0.862, l = 0.9877, \sigma_n = 0.654\}$ and $\Theta_t = \{\sigma_f = 0.0291, l = 0.0068, \sigma_n = 0.0.051\}$ respectively.

- Identify the error function e_f,

$$e_f = \sigma_n^2 \delta_n(x_1, x_2) \tag{8.16}$$

  which is a Gaussian noise with mean 0 and variance δ_n as mentioned earlier in Section (8.2.1) .

- Define the length of the prediction, x_p. The original simulation time for the fuel rig under investigation was 6 minutes or 360 seconds. For prediction by GPR, the obtained data during the 6 minutes simulation was divided into training and validating data sets. Traditionally, 70% of the data which is around 4 minutes of simulation has been used for training and the rest (30% equal to 2 minutes) were used for validation. The prediction time then was continued to 10 minutes, $x_p = 4$ minutes.

- Calculate the mean and variance and the corresponding kernel function at each point for the obtained x_p. Because the GPR is a probabilistic model, so, it is capable of calculating the prediction intervals (variances) and Mean values of GP function from the trained model.

  The covariance matrix along with the selected Mean function which produces the expected value of $f(x)$ in (8.2), defines the GP.

- Finally, define the confidence intervals or uncertainty as 95% confidence intervals,

$$\bar{y}_* = 1.96\sqrt{var(y_*)}. \tag{8.17}$$

  where y_* is the prediction of y.

Figures (8.4)-(8.6) demonstrate the way that GPR predicting the future steps. Figure (8.4) shows the five samplings from GP priors for Ps3 and the observation points are presented in Figure (8.5). Then Figure (8.6) shows the effectiveness of the proposed GPR to predict the PS3 in the fault-free case where the grey area indicates the 95% confidence intervals.

The same method is carried out to predict the residual and adaptive threshold for each pressure sensor and the results are presented in the next section.

8.5 Simulation results and discussions

The following simulation results have demonstrated the effectiveness of the proposed GPR intermittent fault prediction approach obtained from the two-step intermittent fault detection and prediction process mentioned in Figure (**??**). In this approach, two GPRs were designed for the residual and adaptive threshold of the nominated pressure sensor.

Then each time that the predicted residual goes over the predicted adaptive threshold, an intermittent fault is predicted.

Figure (8.7) shows the five GP sampling for the residual 1. Then the residual 1 is predicted, Figure (8.8), using the GPR with the selected squared exponential covariance function and its hyperparameters (summarized in Table 8.2).

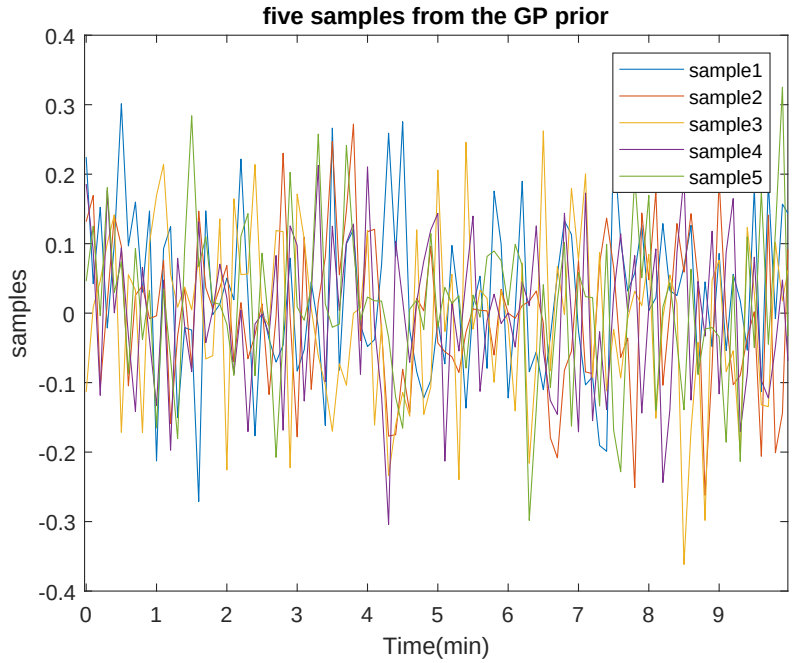

Figure 8.4: Five GP sampling for the pressure 3.

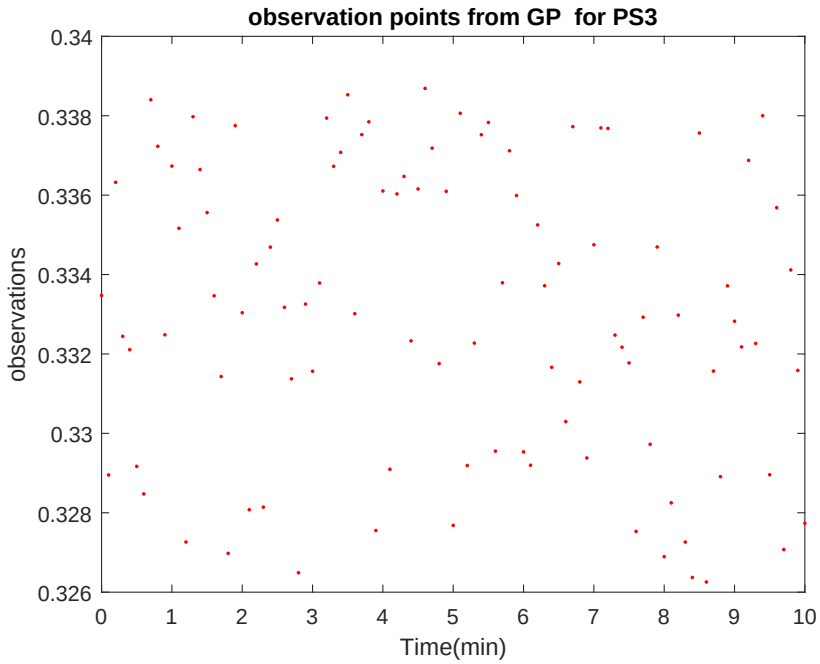

Figure 8.5: The observation points for pressure 3. In this figure, PS3 indicates the pressure sensor 3.

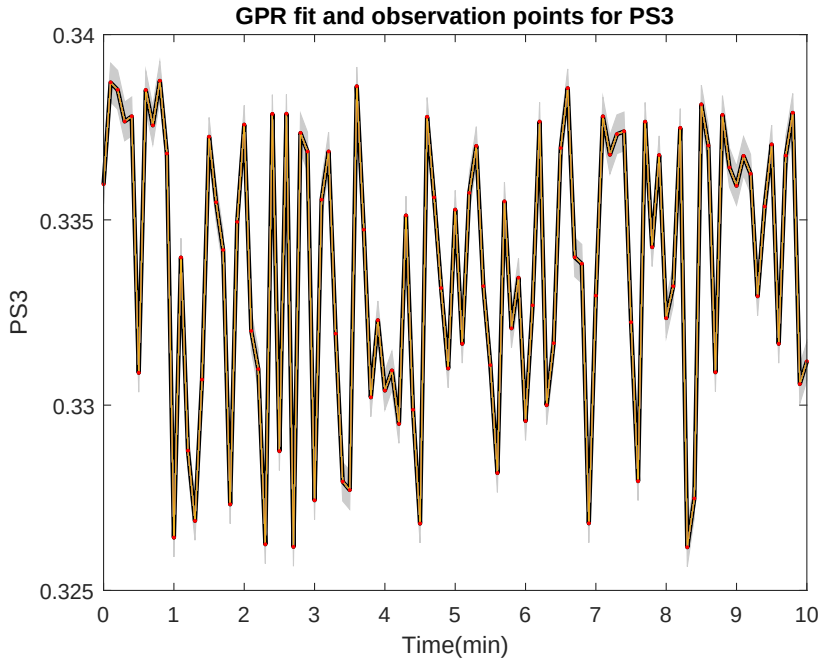

Figure 8.6: the GPR prediction for pressure 3 when there is no fault in the system. In this figure, PS3 indicates the pressure sensor 3.

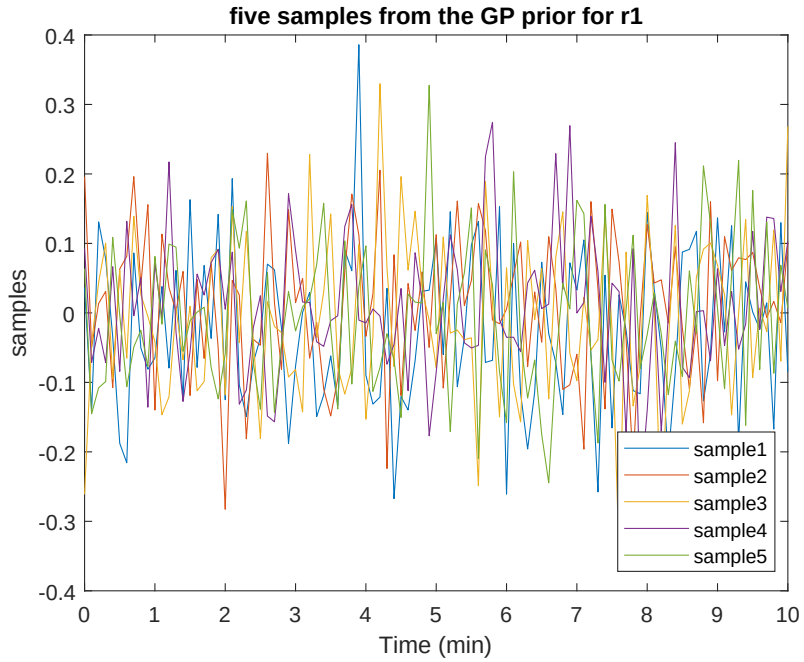

Figure 8.7: Five GP sampling for the residual 1. The sampling points are selected from these sampling for residual 1. In this figure, r1 indicates the residual 1.

The same process has been repeated for all five residuals and adaptive thresholds captured from the intermittent fault detection sections. For each residual and adaptive threshold a suitable GPR along with its covariance function, hyperparameters and error function has been designed. Their simulations are presented in Figures (8.9)-(8.26).

Moreover, Figures (8.27)-(8.33) demonstrate the prediction of intermittent fault while the adaptive thresholds and the residuals were presented in one figure. These figures clearly show that the intermittent fault in the system was successfully detected and when the GPR was trained based on available historical data, it was able to predict the next steps.

However, the simulation results show that anytime that the pressure sensors are closer to the fault location (shut-off valve, between PS3 and PS4) the uncertainty in prediction will increase significantly.

Figure (8.12) which presented residual 2 (related to PS2) shows a greater uncertainty in prediction in compare with residual 1 in Figure (8.12). The reason is that PS2 is closer to the fault location and has been more affected by the fault than PS1. Furthermore, the residuals 3 and 4 in Figures (8.16) and (8.20) show even more uncertainties in compare with residual 2 because PS3 and PS4 are the closest sensors to the fault location. Finally, in Figure (8.24) the residual r5 has been predicted by GPR and as it shows the GP Mean is closer to the observations because the PS5 is not as close as PS3 and PS4 to the fault location.

Also, Figures (8.29) and (8.31) show the highest uncertainty in predicting the intermittent fault and in both the figures the confidence intervals of the adaptive thresholds have been masked by the confidence intervals of the residuals. However, Figures (8.30) and (8.32) present the adaptive threshold and their 95% confidence intervals for Ps3 and PS4 respectively.

Therefore, the simulation results demonstrate that the location of the fault is very effective in the quality of the observed data and the prediction process (GPR fit and prediction intervals). Because the observed values are not the exact function values, but a noisy realization of them. So, whenever the observations are very noisy like PS3 and PS4, then the GPR fit is far from the observations and the standard deviation of the predicted responses are very large with huge prediction intervals around them. Although, when the observations are less noisy like PS1, then the GPR fit is very close to the observation or may cross the observations, and the prediction intervals are very small around these values.

Finally, all the presented results in this section have been demonstrated the effectiveness of the proposed GPR in predicting intermittent fault in the fuel rig system under investigation.

8.6 Conclusions

This chapter presents a two-step model-based/GPR intermittent fault prediction approach. In the first step, a NUI model (observer)-based FD model was designed to generate the residual errors and the adaptive threshold to detect the intermittent fault in the system under investigation (results have been presented in previous chapters). Then, the forthcoming residual and adaptive threshold were predicted iteratively by GPR. Anytime

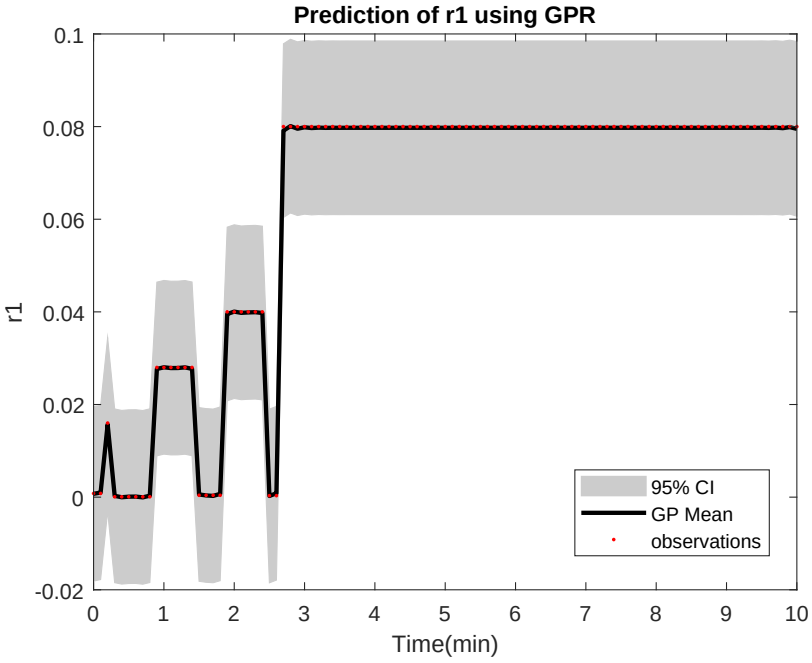

Figure 8.8: The GPR prediction for the residual 1. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the residual 1 observation points. In this figure, r1 indicates the residual 1.

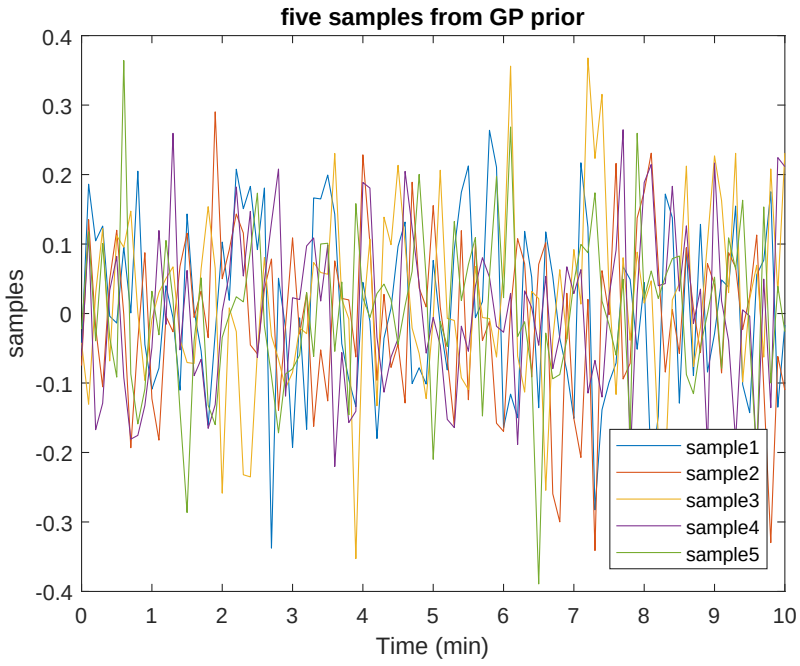

Figure 8.9: Five GP sampling for the threshold 1. The sampling points are selected from these GP sampling for threshold1. In this figure, T1 indicates the threshold 1.

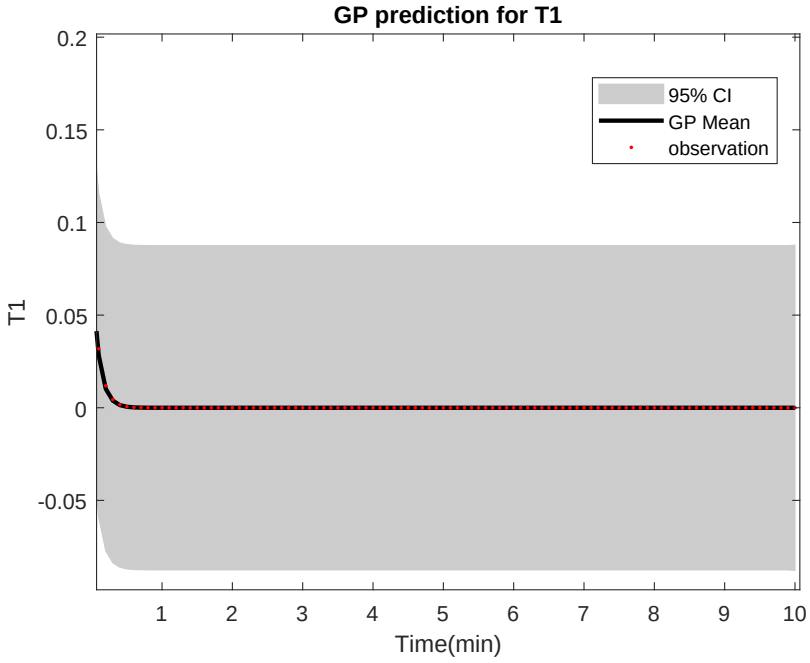

Figure 8.10: The GPR prediction for the threshold 1. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the threshold 1 observation points. In this figure, T1 indicates the threshold 1.

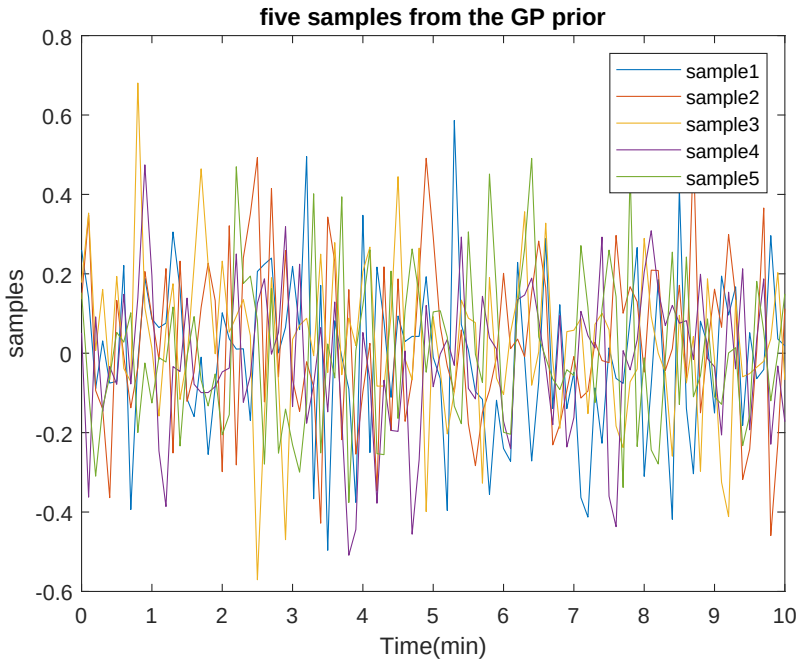

Figure 8.11: Five GP sampling for the residual 2. The sampling points are selected from these GP sampling for residual 2. In this figure, r2 indicates the residual 2.

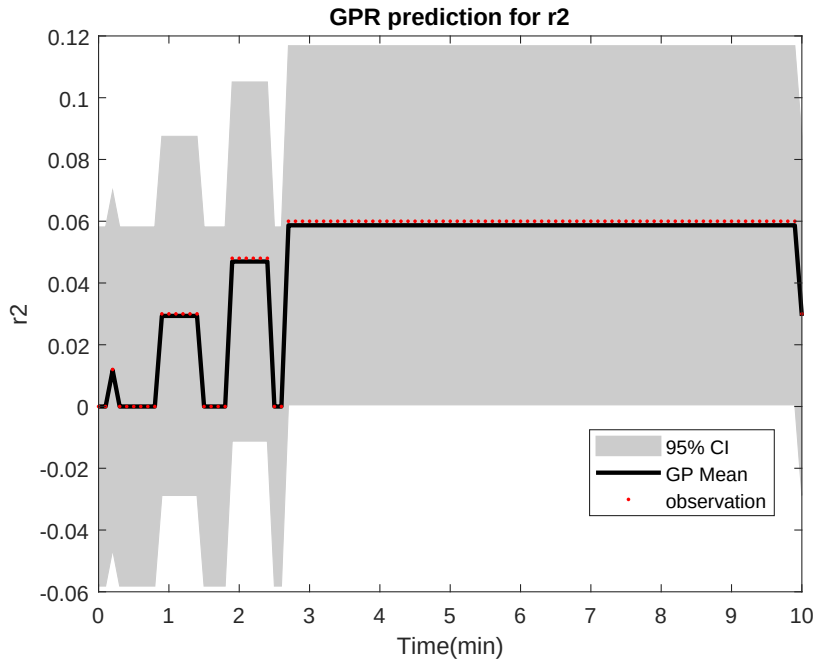

Figure 8.12: The GPR prediction for the residual 2. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the residual 2 observation points. In this figure, r2 indicates the residual 2.

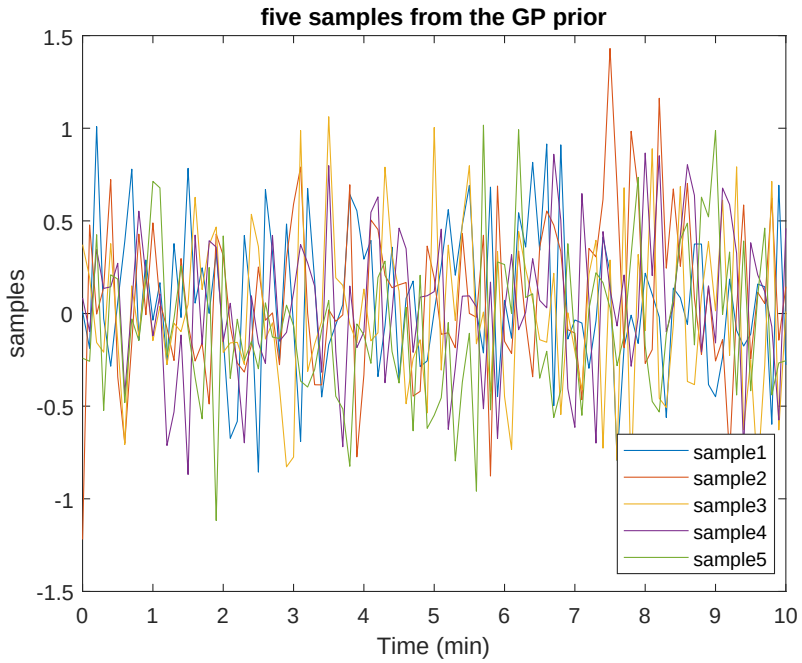

Figure 8.13: Five GP sampling for threshold 2. The sampling points are selected from these GP sampling for threshold 2. In this figure, T2 indicates the threshold 2.

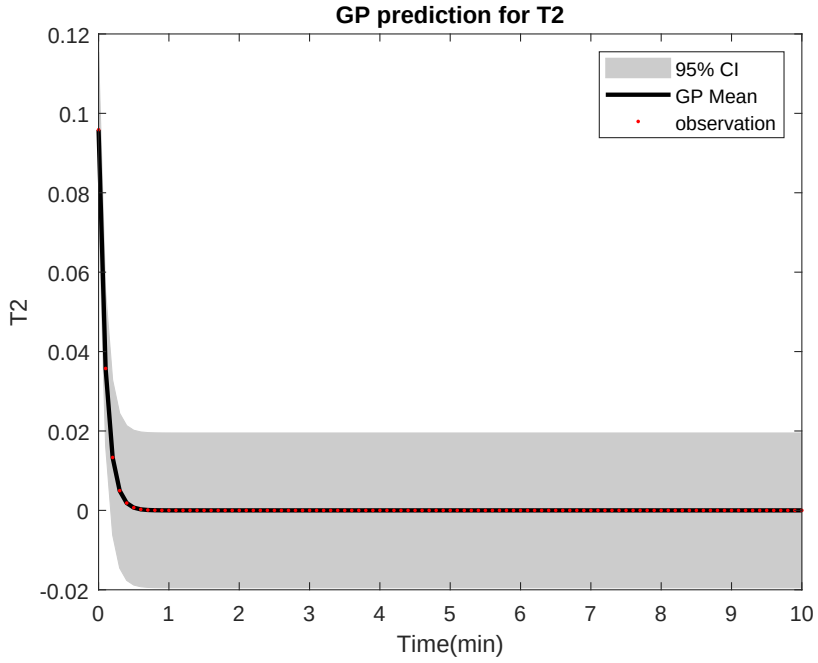

Figure 8.14: The GPR prediction for threshold 2. In this figure the grey areas are the confidence intervasl, the black solid line is the GP mean function and the red points are the threshold 2 observation points. In this figure, T2 indicates the threshold 2.

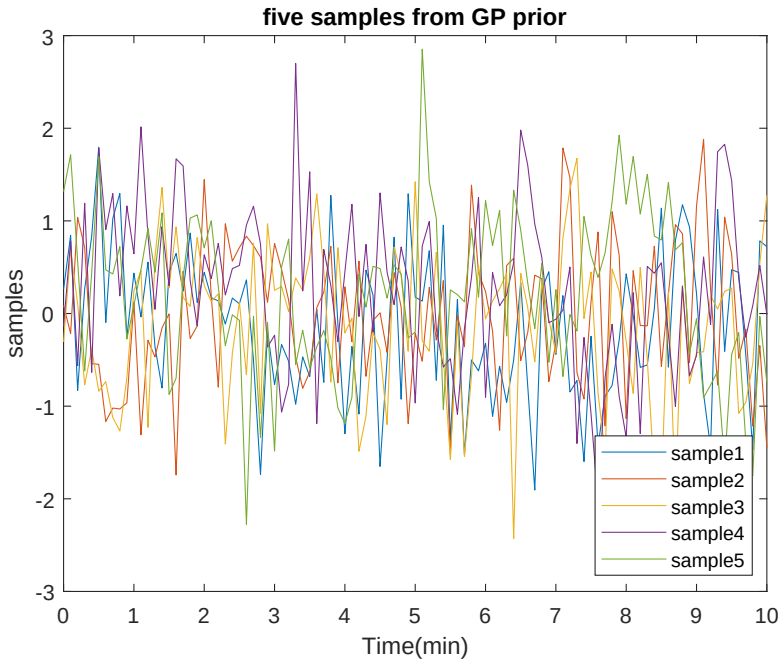

Figure 8.15: Five GP sampling for the residual 3. The sampling points are selected from these GP sampling for residual 3. In this figure, r3 indicates the residual 3.

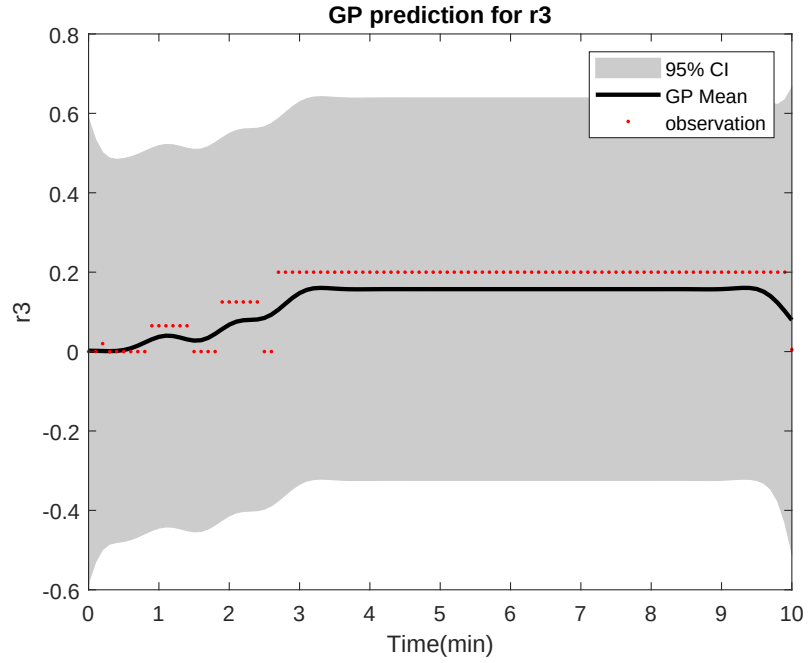

Figure 8.16: The GPR prediction for the residual 3. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the residual 3 observation points. In this figure, r3 indicates the residual 3.

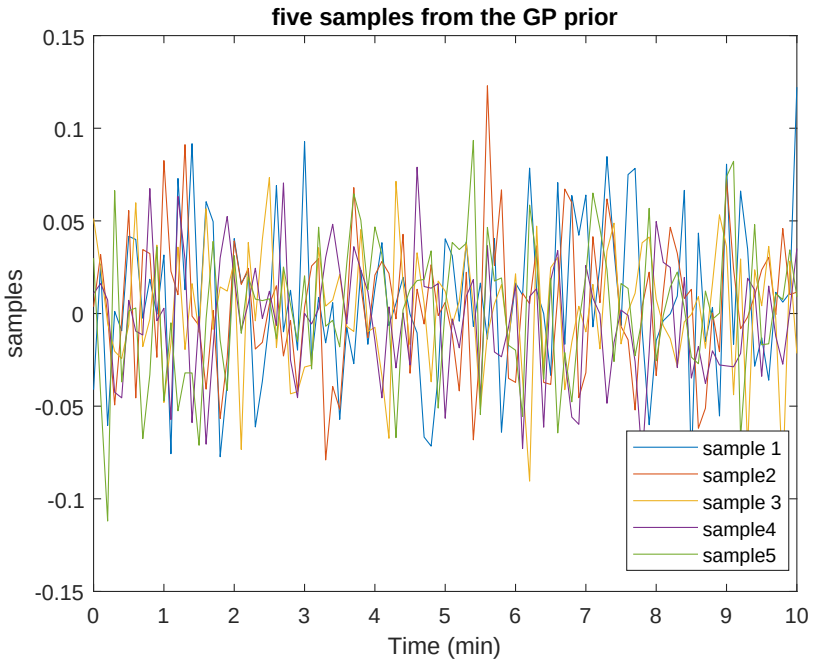

Figure 8.17: Five GP sampling for the threshold 3. The sampling points are selected from these GP sampling for threshold3. In this figure, T3 indicates the threshold 3.

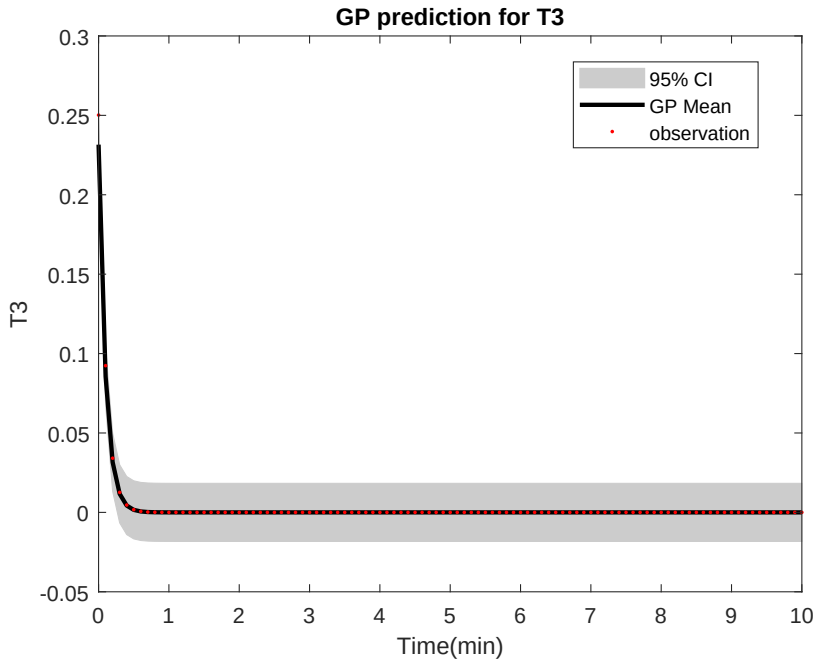

Figure 8.18: The GPR prediction for the threshold 3. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the threshold 3observation points. In this figure, T3 indicates the threshold3.

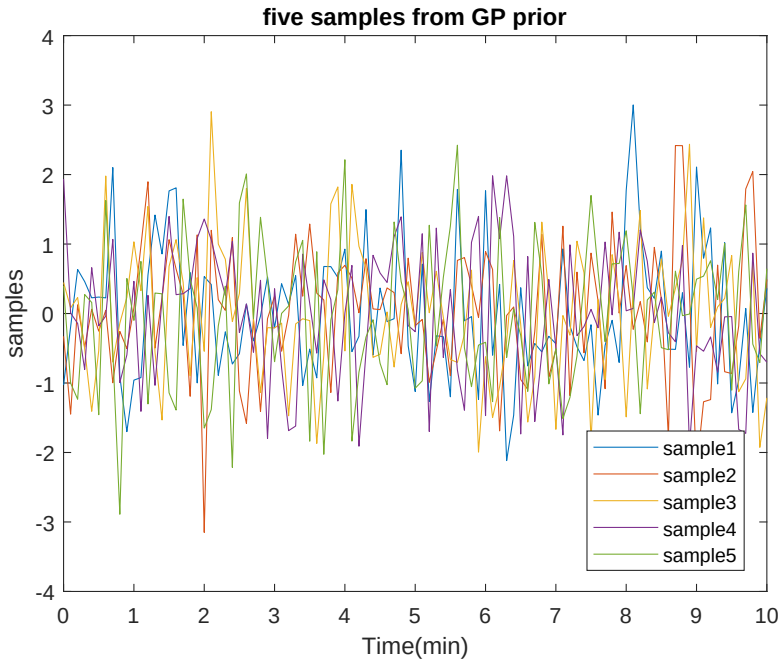

Figure 8.19: Five GP sampling for the residual 4. The sampling points are selected from these GP sampling for residual 4. In this figure, r4 indicates the residual 4.

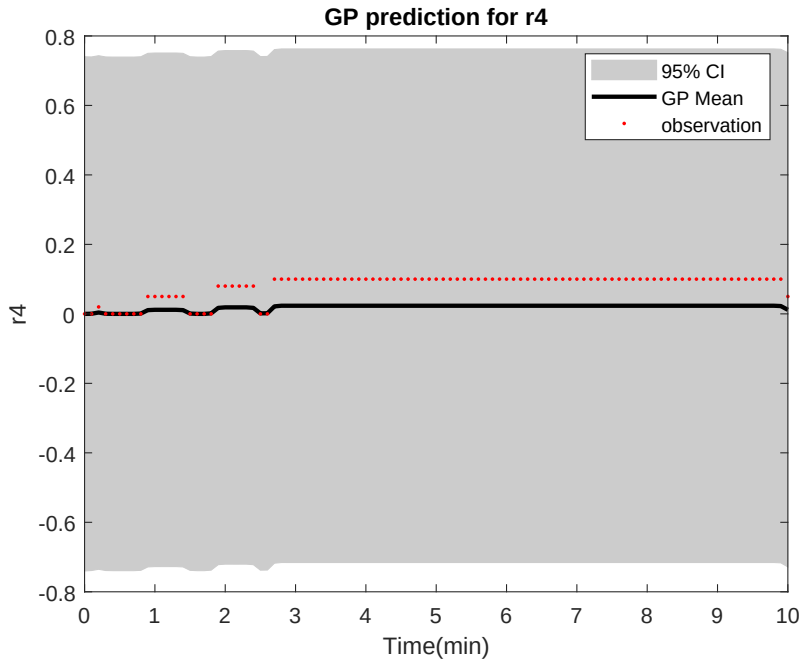

Figure 8.20: The GPR prediction for the residual 4. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the residual 4 observation points. In this figure, r4 indicates the residual 4.

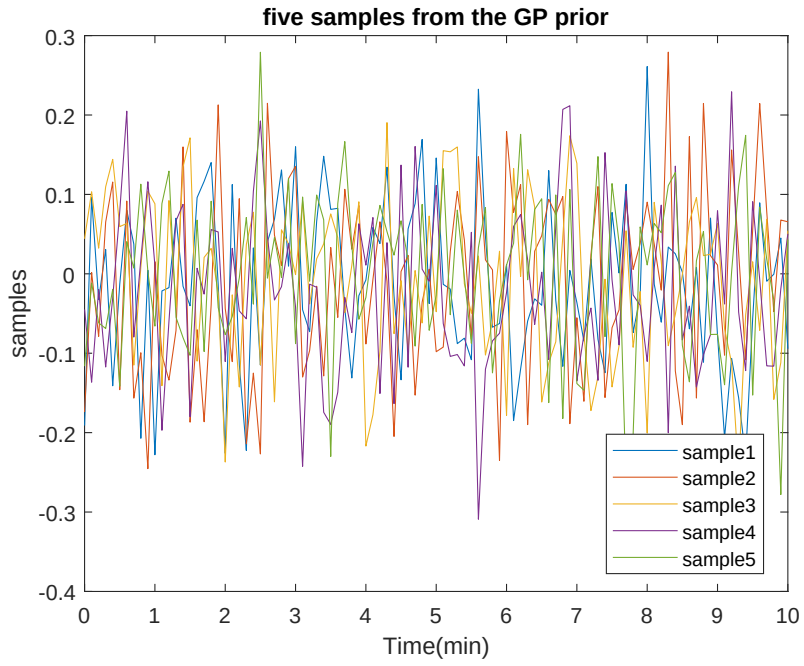

Figure 8.21: Five GP sampling for the threshold 4. The sampling points are selected from these GP sampling for threshold 4. In this figure, T4 indicates the threshold 4.

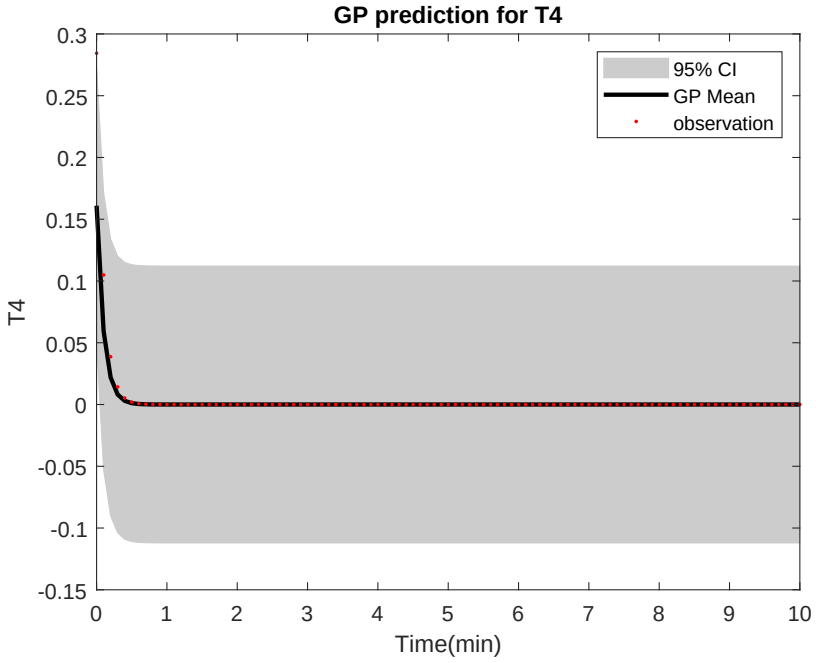

Figure 8.22: The GPR prediction for threshold 4. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the threshold 4 observation points. In this figure, T4 indicates the threshold 4.

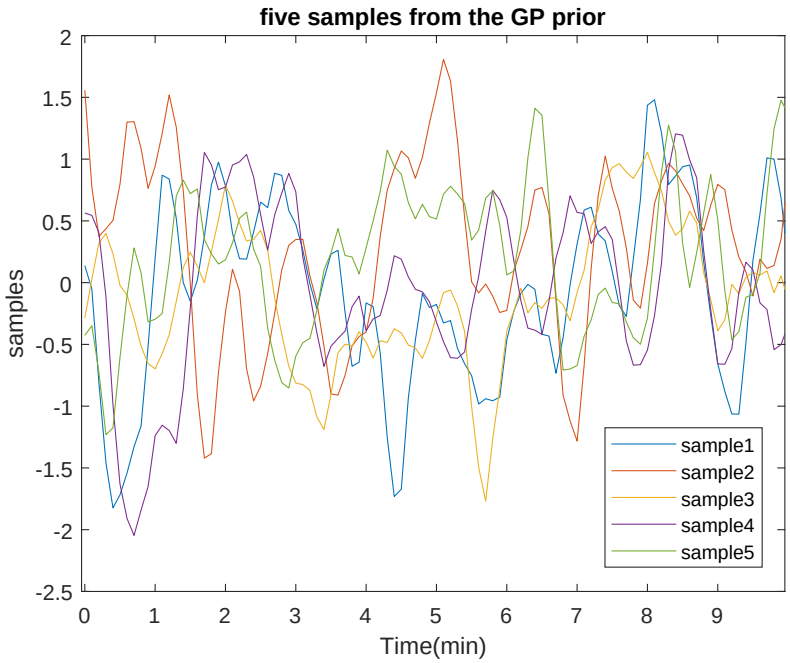

Figure 8.23: Five GP sampling for the residual 5. The sampling points are selected from these GP sampling for residual 5. In this figure, r5 indicates the residual 5.

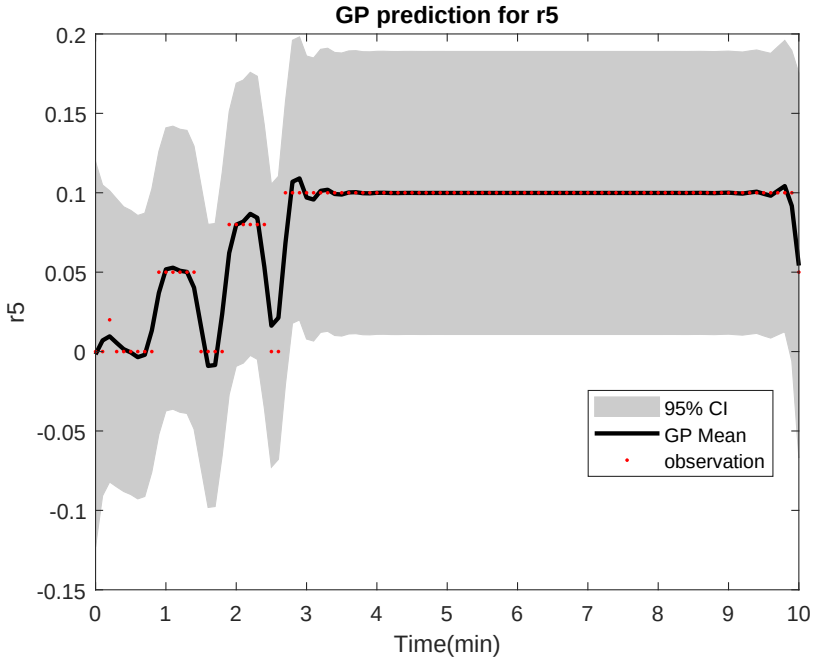

Figure 8.24: The GPR prediction for the residual 5. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the residual 5 observation points. In this figure, r5 indicates the residual 5.

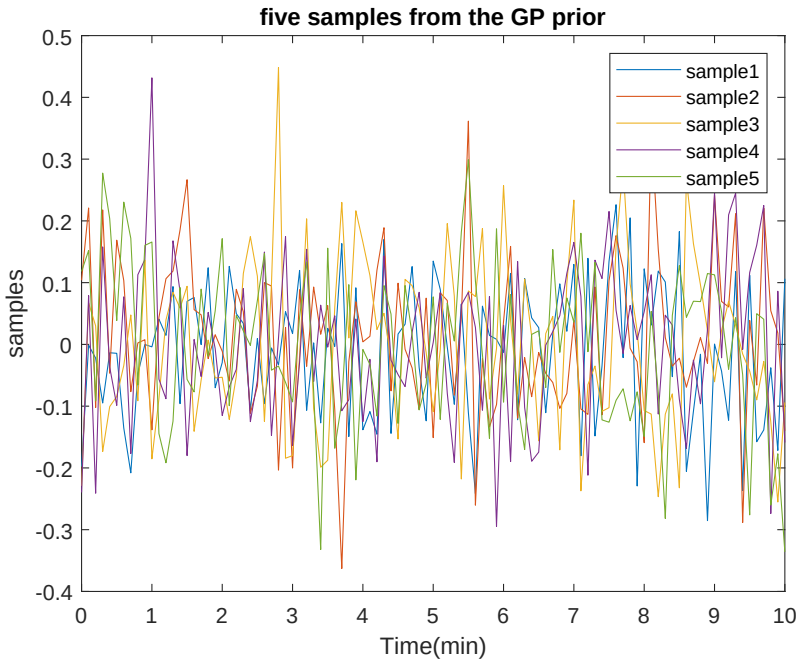

Figure 8.25: Five GP sampling for threshold 5. The sampling points are selected from these GP sampling for threshold 5. In this figure, T5 indicates the threshold 5.

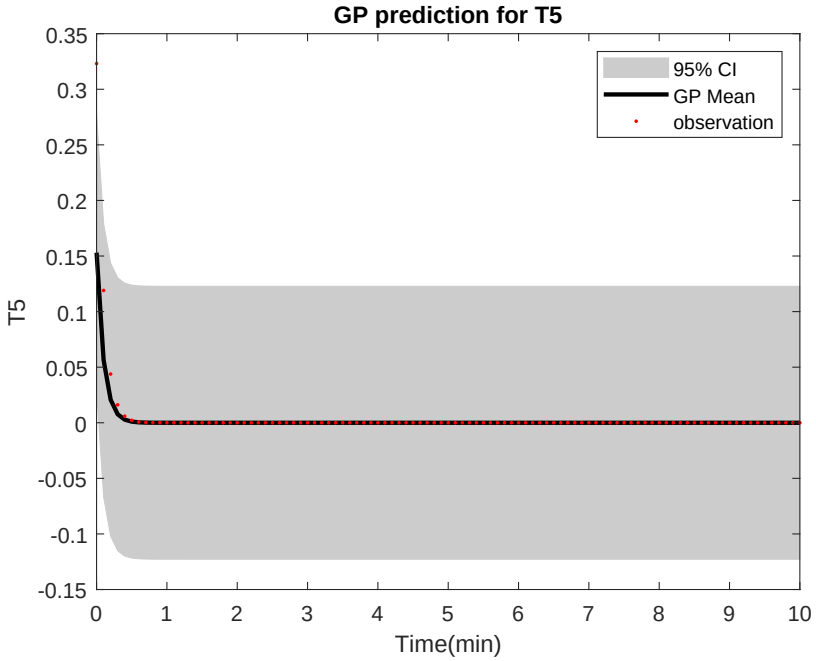

Figure 8.26: The GPR prediction for the threshold 5. In this figure the grey areas are the confidence intervals, the black solid line is the GP mean function and the red points are the threshold 5 observation points. In this figure, T5 indicates the threshold 5.

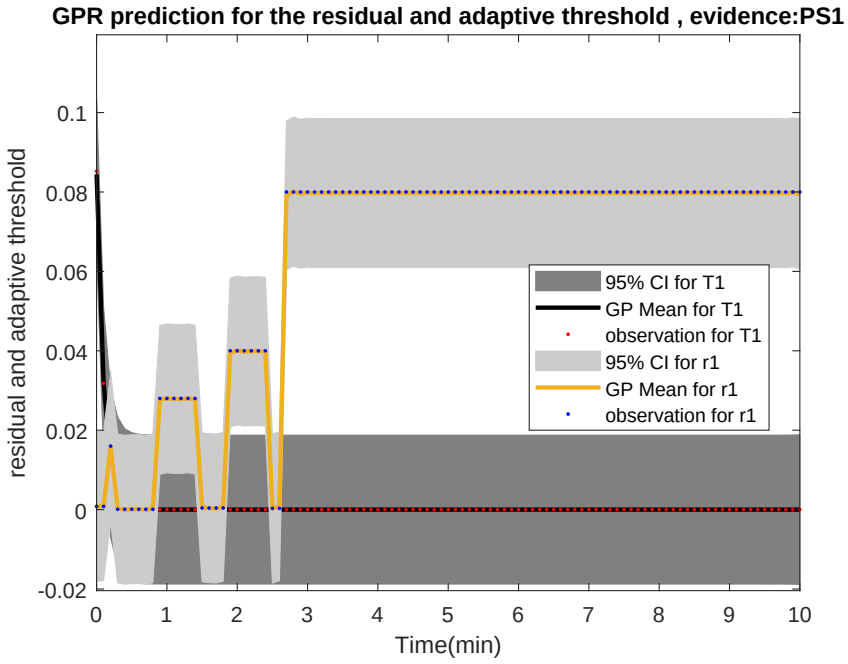

Figure 8.27: The GPR predictions for the residual and threshold 1. In this figure the grey areas are the confidence intervals, the yellow and the black solid lines are the GP mean functions for the residual and threshold respectively. The blue and the red points are the residual 1 and threshold 1 observation points. In this figure, r1 indicates the residual 1 and T1 indicates the threshold 1.

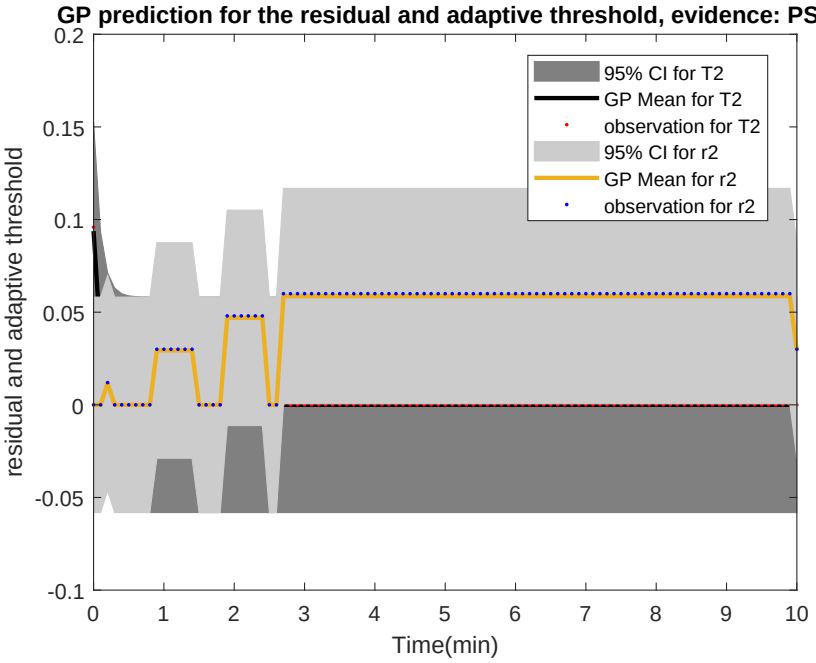

Figure 8.28: The GPR predictions for the residual and threshold 2. In this figure the grey areas are the confidence intervals, the yellow and the black solid lines are the GP mean functions for the residual and threshold respectively. The blue and the red points are the residual 2 and threshold 2 observation points. In this figure, r2 indicates the residual 2 and T2 indicates the threshold 2.

that the predicted residual exceeds the predicted adaptive threshold, the forthcoming intermittent faults were predicted.

The results in this chapter show that the reliability of the GPR depends on how well the covariance (kernel) function, and its hyperparameters, Θ, were selected. If the hyperparameters were not chosen carefully, then the results are not accurate and sensible.

Moreover, the selected covariance (kernel) function can also grow to suit the complexity of our data since there is no limit to how complicated the kernel function can be. In this chapter based on the prior knowledge about the physical process under investigation, the best covariance (kernel) function among alternative covariance functions was chosen.

8.6.1 Limitations

Despite the noticeable advantages of GPR, there are some limitations that one of the important ones is listed below:

- The amount of data is important since too few data makes it difficult to extract accurate insight, whilst too much data make the model unnecessarily complex. Moreover, the efficiency of the prediction model is critically dependent on the quality of the data. Collecting good quality data containing intermittent faults is very challenging. Collecting data is made more challenging because large

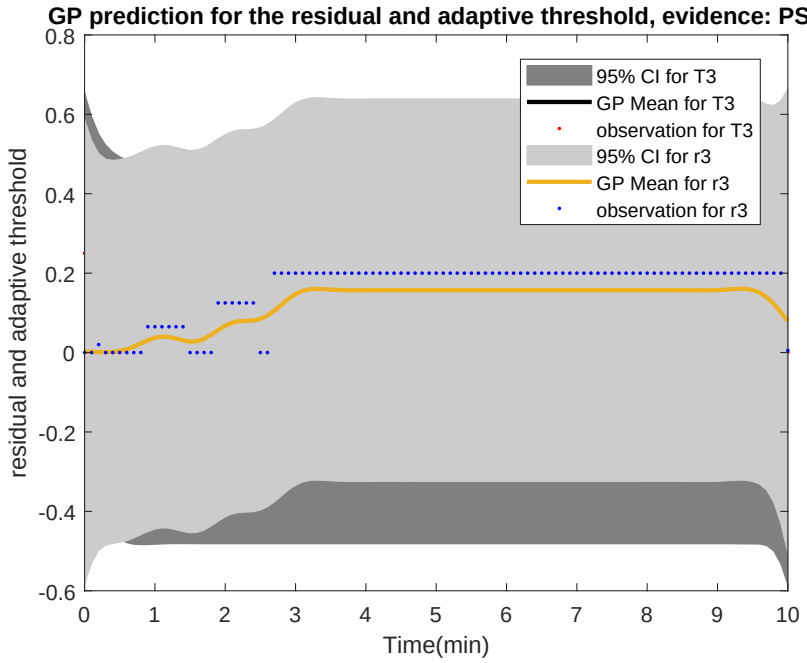

Figure 8.29: The GPR predictions for the residual and threshold 3. In this figure the grey areas are the confidence intervals, the yellow and the black solid lines are the GP mean functions for the residual and threshold respectively. The blue and the red points are the residual 3 and threshold 3 observation points. In this figure, r3 indicates the residual 3 and T3 indicates the threshold 3. In this figure, the threshold3 is masked by the confidence intervals of the residual 3.

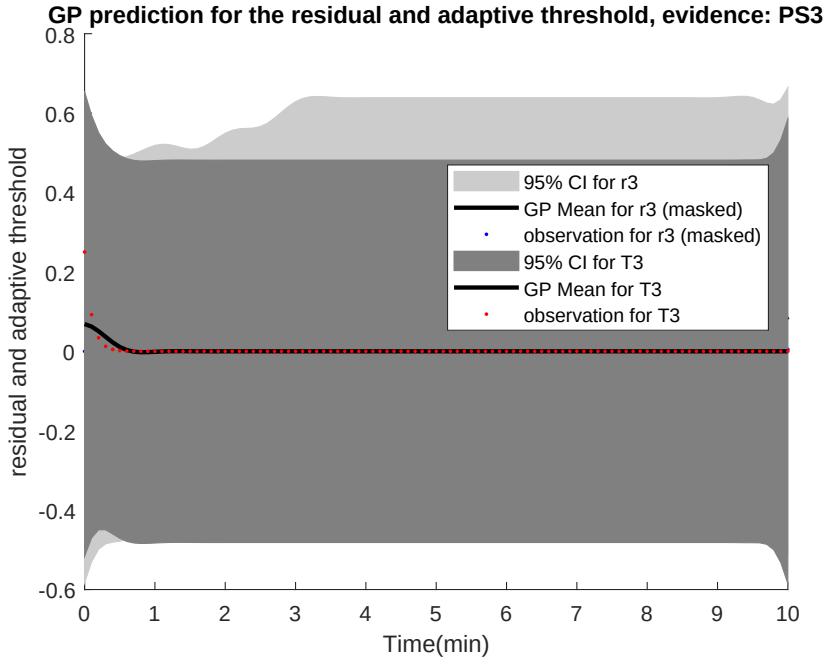

Figure 8.30: The GPR predictions for the residual and threshold 3. In this figure the grey areas are the confidence intervals, the yellow and the black solid lines are the GP mean functions for the residual and threshold respectively. The blue and the red points are the residual 3 and threshold 3 observation points. In this figure, r3 indicates the residual 3 and T3 indicates the threshold 3. In this figure, the masked threshold 3 in Figure (8.29) is presented.

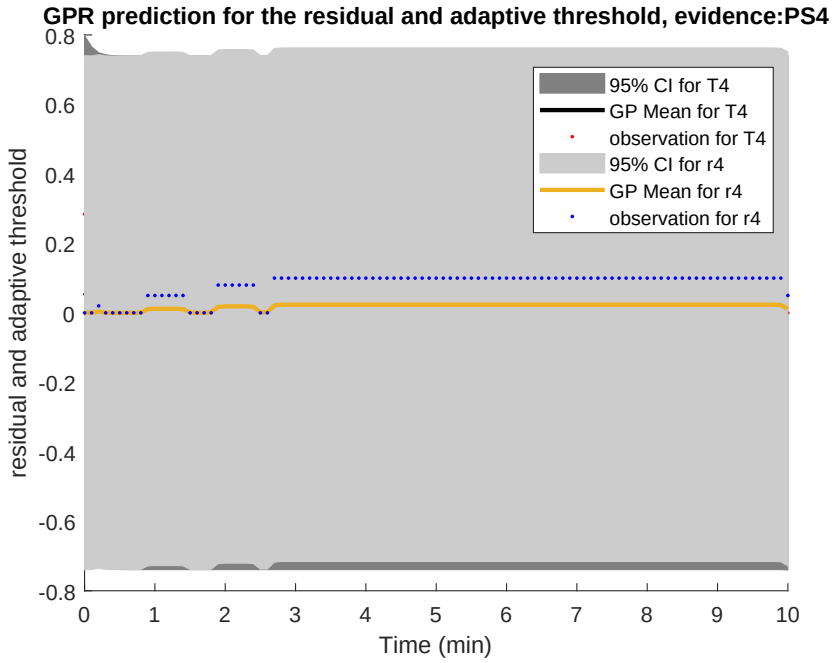

Figure 8.31: The GPR predictions for the residual and threshold 4. In this figure the grey areas are the confidence intervals, the yellow and the black solid lines are the GP mean functions for the residual and threshold respectively. The blue and the red points are the residual 4 and threshold 4 observation points. In this figure, r4 indicates the residual 4 and T4 indicates the threshold 4. In this figure, the threshold3 is masked by the confidence intervals of the residual 4.

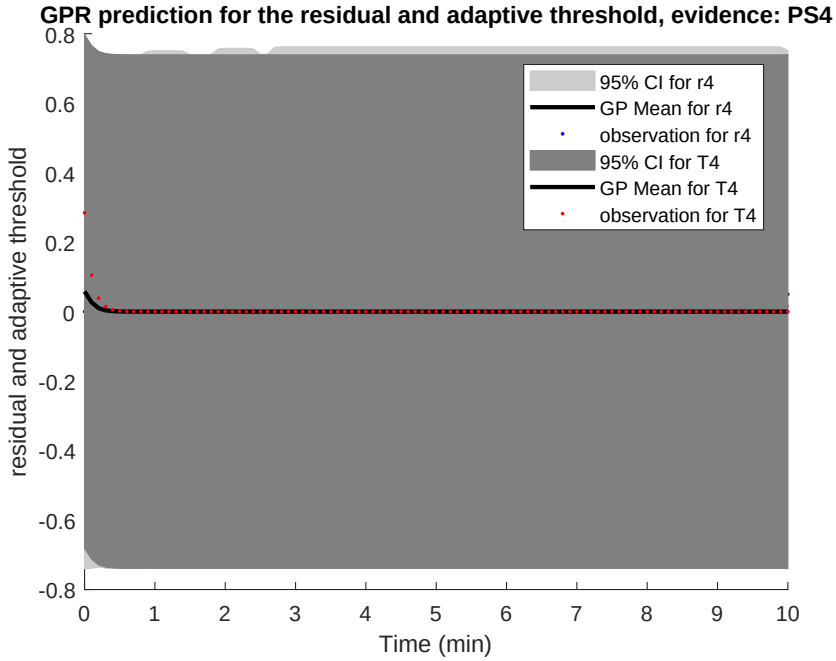

Figure 8.32: The GPR predictions for the residual and threshold 4. In this figure the grey areas are the confidence intervals, the yellow and the black solid lines are the GP mean functions for the residual and threshold respectively. The blue and the red points are the residual 4 and threshold 4 observation points. In this figure, r4 indicates the residual 4 and T4 indicates the threshold 4. In this figure, the masked threshold 4 in Figure (8.31) is presented.

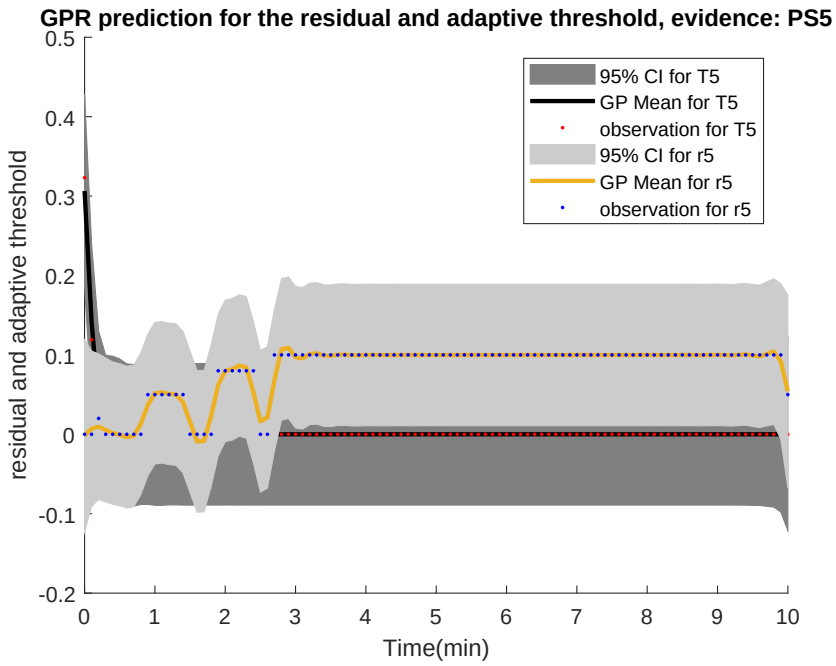

Figure 8.33: The GPR predictions for the residual and threshold 5. In this figure the grey areas are the confidence intervals, the yellow and the black solid lines are the GP mean functions for the residual and threshold respectively. The blue and the red points are the residual 5 and threshold 5 observation points. In this figure, r5 indicates the residual 5 and T5 indicates the threshold 5.

datasets are usually necessary for reliable fault prediction (Malhotra & Bansal, 2012; Schneider, 2015).

Chapter 9

Conclusions and Future Work

This chapter offers a brief summary of the achievements and conclusions obtained during this research. It also summarises the main contributions described in previous chapters, some of which have already been published already. Finally, some further research directions are presented.

9.1 Review of research objectives

The methodology of this thesis was divided into the tasks of intermittent fault detection, isolation and prediction. In this study, a model-based fault detection filter and model-free fault detection and prediction systems were developed to detect, isolate and predict IFs in complex dynamic systems characterised uncertainty.

In Chapter 1, three objectives were stated, and attempt were made to achieve them. This section revisits each research objective with a discussion of how they were achieved.

The first objective stated in Chapter 1 was to:

1) Identify, test and validate novel model-based methods that can detect faults in the presence of uncertainties, with a special focus on intermittency. This objective was addressed in Chapters 5 and 6.

   Consistency checking in model (observer)-based methods is normally achieved using residual quantities, which are computed as differences between the measured signals and the corresponding signals generated by the mathematical model. In practice, the most frequently used diagnosis method is to monitor the level (or trend) of the residual and take action when the signal reaches a given threshold. Although finding a systematic technique to deal with the unknown input (disturbance) and the intermittent faults in a complex nonlinear system is challenging, in this thesis, a method has been presented to overcome these limitations by designing appropriate observers:

   (a) Two novel observers have been presented in this thesis:

- The NUIO is usually designed such that its state estimation errors approach zero asymptotically, regardless of the presence of the unknown input (disturbance) in the system. To design such an observer, normally the rank condition must be satisfied, which means that the measurements are directly dependent on the states affected by the unknown inputs (disturbances). However, for the system under investigation, this rank condition is not satisfied, and none of the measurements is directly dependent on the states affected by the unknown inputs (disturbances). Hence, to minimise the effect of the unmatched unknown inputs/disturbances on the error estimation of the system, the novel NUIOs are presented in Chapter 5.
  To design this observer:
  - First, a dummy full rank matrix is designed to satisfy the rank condition of the observer design,
  - Then, the observer gain is selected,
  - Next, sufficient conditions for designing the NUI observer are derived,
  - Finally, the stability of the errors is ensured.

  The proposed NUIO has been claimed as a novel observer because the existing restriction of the NUIO rank condition has been avoided without losing the design convenience by introducing the auxiliary disturbance's distribution matrix. Moreover, this observer is not only applicable to the limited class of nonlinear systems with Lipschitz nonlinearity. Instead, the proposed observer has been designed for the more general form of the nonlinear systems, while the asymptotic stability of the system error estimation (5.17) and the existence of the NUIO given by (5.9) guarantee the applicability of this statement.
  Moreover, another advantage of the proposed NUIO is the possibility of converting the complicated nonlinear system to the linear system by applying a linear equality (5.41) mentioned in the design procedure, which makes the difficult NUIO design problem an easy task for the considered class of nonlinear systems.
- The feed-forward observer, in addition to the NUIO, is another novel observer designed in Chapter 5. This observer also neglects the rank condition assumption and hence will give more degrees of freedom to the designer. The designed feed-forward observer is also effective in detecting IFs in nonlinear complex systems. These observers are then used to obtain the system's state-space equations. These equations are then used to design the model-based fault detection filter.

(b) State–space equations are a set of differential equations that make use of the notions of state, inputs, outputs and dynamics to describe a system's behaviour. To formulate these equations, the mathematical modelling of the system is needed. The main case study investigated in this thesis to test the effectiveness of the NUI observer-based fault detection filter was the experimental fuel rig system. The experimental fuel rig in this research was a SIMO system because it has only one capacitor (one main tank) in the

system. Moreover, the mathematical model of a Multi Input Multi Output (MIMO) car suspension system was also designed for the purpose of IF detection using a feed-forward observer.

The main contribution of this part of the objective is that the obtained state-space equations from the mathematical modelling of the systems provide a fully functioning system that can be used to test the effectiveness of the proposed fault detection filters. This has been addressed in Chapter 6.

Furthermore, the main idea of the NUI and feed-forward observers design was to make the residual signal decouple from the external unknown inputs. Usually, to achieve this, attempts are made to make the residual signal insensitive to unknown inputs but sensitive to faults.
However, a problem may arise when the intermittent fault lies in the same subspace as the disturbances; then, upon decoupling the disturbances, the residual will also be insensitive to the fault, which is obviously not the objective of a detection filter. Moreover, the existence conditions for decoupling the unknown inputs become quite strict, which limits the use of these approaches. How to deal with these problems is explained by introducing a novel adaptive threshold.

(c) A novel adaptive threshold has been designed in this research that can adapt itself to the system's response. The designed adaptive threshold is claimed to be novel because it was designed with regards to the effect of nonlinearity of the system, input and output disturbances while remaining sensitive to the faults. This has been addressed in Chapters 5 and 6.

Finally, the effectiveness of the proposed methods has been tested and validated against two physical real systems – an experimental fuel rig and a car suspension system– and their simulation results were presented in Chapters 5 and 6.

The second objective stated in Chapter 1 was to:

2) Employ, compare and validate model-free approaches to detect and isolate faults in the presence of uncertainties, with a special focus on intermittency.
This objective has been addressed in Chapter 7 .

The chosen model-free method for detecting and isolating faults in this study was the probabilistic data-driven, Bayesian network method. The effectiveness of the Bayesian network in detecting and isolating well-known types of faults, such as hard or abrupt faults, has been already tested. However, the novelty of employing Bayesian networks in this research is the methodology of using them to detect and isolate intermittent faults.

To achieve this objective, two types of BNs were designed to detect and isolate the intermittent faults in the experimental fuel rig system with the presence of uncertainty:

- Static BN, and
- Hybrid dynamic BN.

The effectiveness of the static BN has been demonstrated by its detecting and isolating the faults. However, when the faults are intermittent, this network is not powerful enough to detect the fault and its possible root causes. The reason is that the fuel rig system contains a few continuous variables that may lose their effectiveness as a result of the discretisation process while using a static Bayesian network.

Hence, a HDBN was designed to detect and isolate IF on the experimental fuel rig system. The reason for selecting a HDBN over a dynamic BN is that the system under investigation contains both dynamic variables, such as pressure sensor readings, and static or Boolean variables such as faults. Hence, a HDBN is more suitable since unlike static and dynamic BNs, it can handle both discrete and continuous variables.

The effectiveness of the designed HDBN in detecting IF and isolating their root causes has been tested and validated in Chapter 7, and its performance was compared with that of static BNs.

Moreover, the reliability of the HDBN in detecting and isolating IFs was tested and validated, and positive results were achieved using sensitivity analysis.

When an IF is diagnosed, it is necessary to determine how that fault will affect system performance in the future. Hence the third objective stated in Chapter 1 was to:

3) Propose, test and validate model-free approaches to predict faults in the presence of uncertainties, with a special focus on intermittency. This objective has been addressed in Chapter 8.

The sequence of development leading to this objective is listed below:

- Apply an appropriate statistical scheme (on the fuel rig system as an example system) to predict the forthcoming intermittent faults for the steps ahead.
- Combine the intermittent fault detection and prediction techniques to create the intermittent fault detection and prediction system for the system under investigation.

To predict the intermittent fault in a complex nonlinear system, a unified and reliable method is needed. The chosen method in this thesis was the Bayesian Gaussian process regression method, which is a non-parametric probabilistic data-driven prediction technique that produces a mean estimate along with an indication of the uncertainty therein. The novelty of the proposed method was its capability to predict the intermittent fault in the experimental fuel rig system and quantify its prediction uncertainties.

Moreover, finding an appropriate covariance matrix for better prediction with less uncertainty was one of the challenges of this task. The simulation results in Chapter 8 show how different covariance matrices can change the prediction results and their uncertainty quantification significantly.

In this thesis, to predict IF, a novel hybrid approach for intermittent fault detection and prediction in a complex system was presented (Figure (8.2) in Chapter 8). The first stage of this approach required the development of a mathematical or HDBN model of the nonlinear system under investigation. This nonlinear model was then used to create a fault detection filter with a suitable adaptive threshold to detect the possible faults in the system. Then in the second stage, a probabilistic data-driven fault prediction method, Gaussian process regression, was implemented to predict the intermittent faults using all the information captured from the first stage.

The effectiveness of the proposed methodology for the experimental fuel rig system was demonstrated and validated using simulation results.

Finally, these objectives have been achieved to honour the main aim of this research which was:

- **To achieve a robust intermittent fault detection and prediction approach for nonlinear complex systems.**

All the work addressed in this thesis incorporates the development of the path to achieving a fault detection and prediction system capable of reliably detecting and predicting the intermittent faults as they occur and identifying the location of the fault.

9.1.1 Shortcoming of this research

Although the aim of this research was to design a novel method for intermittent fault detection and prediction, there are some shortcomings to this work of which the author is aware, as follows:

1) In this thesis, intermittent fault was infused into the system artificially; however, in the real world intermittent faults may be more inchoate and vague.

2) The obtained historical data generated for the prediction process are very rich with the intermittency included. Nonetheless, this is not always achievable in the real world.

3) The considered case study in this thesis was a mechanical-hydraulic system; however, intermittent faults are more likely to be developed in electrical or electronic systems.

9.2 Summary of contributions

A summary of the research contributions is presented in the list below:

1) The novel design of model-based fault detection filters in the presence of uncertainty and intermittency along with the novel design of adaptive threshold to improve the fault detection process (Chapter 5).

2) The employment of two model-free, data-driven, fault detection methods for intermittent fault detection and isolation. The two methods were compared and validated using sensitivity analysis (Chapter 7). In this chapter, the system modelling approach was used to construct a HDBN model of the system that incorporates expert opinion, reliability data, mathematical models, operational data and laboratory data. The probabilistic nature of HDBNs allows them to handle uncertainty in the information used to build them.

   Another contribution was, the ability to update the distributions of the variables when new information injected which makes HDBNs capable of detecting intermittent faults. Finally, the designed HDBN was able to detect the intermittent fault and its possible root causes in the experimental fuel rig successfully.

3) Introduction of a novel methodology for intermittent fault prediction in the system under investigation. The achieved positive results were tested and validated against a physical system. The simulation results show that the two-step process performs successfully for intermittent fault detection and prediction of intermittent clogged valve in the experimental fuel rig system (Chapter 8).

Consequently, the work addressed in this thesis incorporates the development of the path to achieving a fault detection and prediction system capable of reliably detecting and predicting intermittent faults as they occur and classifying their location. Moreover, the presented achievements in this thesis apply to any system with the same conditions.

9.3 Future work

The methods proposed in this research thesis addressed the aim and objectives mentioned in Chapter 1. However, there is much room for improvement in many directions in both long-term and short-term future works.

9.3.1 Short-term future work

A number of short-term future works could be carried out to improve the problems discussed in this thesis, as follows:

- The first part of this thesis, focused on model-based fault detection methods. One extension to this part of the research could be to study the model-based isolation techniques and compare their isolation capability against the model-free-based isolation techniques.

- Moreover, the designed Bayesian networks can be extended to a decision-making-based Bayesian network. Hence, the user can make the best decision under the prevailing uncertainty for each action or scenario by comparing their maximum expected values.

- A further development of this work could be to evaluate the designed fault detection and prediction method with a piloted in-service aircraft fuel rig system. The

interest of this work is that a human, the pilot, is also in the loop along with other reliable available information to detect and predict the fault.

9.3.2 Long-term future work

The possible interesting long-term directions to extend and improve the current research are such as:

- The implementation and improvement of other intermittent fault detection and prediction methods in comparison to this work.
- Another possible future objective which would be the most important extension to this research, is the improvement of the intermittent fault prediction methods. In this research GPR, which is a non-parametric kernel-based probabilistic method, was used for the purpose of intermittent fault prediction. However, if the appropriate historical data are not available or do not contain any form of intermittent fault, then the GPR prediction is not accurate. To remediate this problem, one of the suggestions is to combine the GPR method with the Bayesian optimization method. A Bayesian optimization method is applied to all the Gaussian process regressors to maximize their prediction by finding the global optimum in a minimum number of steps. In other words, the Bayesian optimization method can be used to provide more sampling data points from the system's function to improve the data sets used by the GPR for prediction.

Appendix A

Lipschitz Nonlinear Systems

Consider the class of nonlinear systems described by the following equations:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + g(x,u,t) \\ y(t) &= Cx(t) \end{aligned} \tag{A.1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$ present state, input and output vectors, respectively. Moreover, $A \in \mathbb{R}^{n\times n}$, and $C \in \mathbb{R}^{p\times n}$ are known linear control matrices of the dynamic system and the pair (A,C) is assumed to be observable.

A nonlinear function $g(x,u,t)$ is said to be Lipschitz in a region D_s enclosing the origin if there exists a scalar such $\kappa \in \mathbb{R}$ that the relation

$$\|g(x,u,t) - g(\hat{x},u,t)\| \leq \kappa \|x - \hat{x}\|$$

holds $\forall x, \hat{x} \in D_s$, where $\kappa \geq 0$ is the Lipschitz constant (Li, 2017).

Appendix B

Probability Theory

In statistics the probability is the likelihood that an event will occur. In Bayesian probability, the mathematical theory of probability is applied to the degree to which a belief is considered probable. This subsection gives a brief summary of the most important concepts of probability theory considered under this approach (Maher, 2010). In order to understanding of information that permits to compute the probability the following axioms are defined (Chung, 2001):

- $Pr(A)$ indicates the unconditional probability of an event occurring and is non-negative for all real A, $Pr(Re(A)) \geq 0$.
- Probability of an event A, may range from zero to one, $0 \leq Pr(A) \leq 1$.
- The probabilities of all possible outcomes (denoted by Ω) must sum to one,

$$Pr(\Omega) = \sum_i Pr(A_i) = 1.$$

- $Pr(A) + Pr(\bar{A}) = 1$, where $\bar{A}$ indicates the complement of event A.
- $Pr(A|B)$ indicates conditional probability. The probability of event A occurring given that event B occurs.
- $Pr(A \cap B)$ or $Pr(A \textit{ and } B) = Pr(A).Pr(B)$ indicates the joint probability (independent events). It is the probability of the intersection of two or more events.
- $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A).Pr(B)$ where A and B are mutually exclusive propositions (addition rule).
- $Pr_C(A) = \frac{n_A}{n}$ indicates classic probability where n_A is the number of repetition of A, and n is the total number of the data.
- $Pr_F(A)$ indicates the frequencies probability is defined as follows

$$Pr_F(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

The value of $Pr_F(A)$ will finally converge to a number which is the probability.

- For a discrete function probability at a single point is, $Pr(A) = f(A)$.
  The random outcomes are countable and values between these counts cannot occur.

- For continuous function, probability at a single point is zero, hence expressed in terms of an integral between two points, $Pr[a \leq A \leq b] = \int_a^b f(A)dA$.
  The outcomes and related probabilities are not defined at specific values, but rather over an interval of values.

B.1 Bayesian probability

Independent events:

If an event occurring, A, does not alter the probability of another event occurring, B, then these events are independent (Ross, 2003).

Dependent events:

If an event occurring, A, changes the probability of another event occurring, B, then the probability of event B is dependent on event A, (Ross, 2003).

Exact inference

The task of computing the probability of each variable when other variable's values are known. That means once some evidence about variable's states are asserted into the network, the effect of evidences will be propagated through the network and in every propagation the probabilities of adjacent nodes are updated (Butts, 2003).

There are many inference algorithms, such as: junction tree, variable elimination, Monte Carlo (MC), Gibbs sampling, etc., and each has its own advantages over others in different situations.

Bayes' inferences

The Bayes' theorem gives a criterion for updating belief when new knowledge is introduced (Albert, 2011).

Diagnostic inferences

Obtaining the conditional probability when some variable values are known (priors). If some state of nodes is not available, the propagation of priors finds the best hypotheses consistent with the actual data (Reggia *et al.*, 1985).

Independencies

Two random variables A and B are independent for all values if,

- $Pr(A,B) = Pr(A)Pr(B)$, which means knowing B provide no change in the probability for A, or B contains no information about A.

- $Pr(A|B) = Pr(A)$ or $Pr(A|B) = Pr(B)$, which means knowing A provide no change in the probability for B, or A contains no information about B.

These can generalized to more than two random variables (Butz *et al.*, 2016). In practice true independence is very rare. Independence is an assumption which is impose on the model.

Conditional independencies

Two random variables A and B are conditionally independent given C for all value A, B, C if:

$$Pr(A,B,C) = Pr(A|C)Pr(B|C)$$

or

$$Pr(A|B,C) = Pr(A|C) \;\; or \;\; Pr(B|A,C) = Pr(B|C). \tag{B.1}$$

Equation $B.1$ implies that learning about B, given that we already know C, provide no change in our probability for A, or B contains no information about A beyond what C provides (Butz, 2001).
Conditional independencies can generalized to more than two variables:

$$Pr(A|B,C,D) = Pr(A|D).$$

The conditional independence statements can be used to simplify the joint probability of a system under investigation.

Conditional independencies versus independencies

Conditional independencies does not imply independencies. For illustration for

$$Pr(B|C,A) = Pr(B|C)$$

and

$$Pr(A|B,C) = Pr(A|C)$$

A and B are dependent but are conditionally independent given C.

Chain rules:

There are two main components in BN: the graph structure which show the conditional independencies and assumptions between the variables, and the numerical probabilities for each variable given it's parents known as conditional probability table (Wang & Zhao, 2015). Assume that there are n random variables,
$(X_1, X_2 \cdots, X_n)$, then in general, the Joint Probability Distribution (JPD), for any network, given nodes $\mathbf{X} = (X1, \ldots, Xn)$, is presented as:

$$Pr(\mathbf{X}) = \prod_{i=1}^{n} Pr(X_i | parents(X_i)) \tag{B.2}$$

where $parents(X_i)$ denote the parent set of node X_i. This joint probability is also known as "Chain rule".

Bayes' rule:

The CPT of variable X_i where all its parents are given, known as posterior probability or "Bayes' rule", (Kong *et al.*, 2006). For variable X and its evidence Y, the Baye's rule is defined as:

$$postrior\ probability = \frac{joint\ probability}{prior\ probability} \tag{B.3}$$

or

$$Pr(X|Y) = \frac{Pr(X_E|X).Pr(X)}{Pr(X_E)} \tag{B.4}$$

where

- $Pr(X|X_E)$ is the posterior probability which is the probability of the variable X, given the evidence Y.

- $Pr(X_E|X) = L(X)$ is the likelihood function. It qualifies the likelihood that the observed data would have been observed as a function of the unknown model parameters.

- $Pr(X)$ is the prior probability. It shows what is known before performing the experiments.

The number of entries in CPTs of a BNs can easily become large and grows exponentially with the number of parents, (i.e., for a variable with m states and n parents, the CPTs entries are m^n), however, some simplification in assumption can help in reducing it. Also in BN one can use of "prior domain knowledge" to come up with a BN that requires fewer probabilities (Slezak, 2009).

Markov Chain Monte Carlo (MCMC) (MArkov Property):

A Markov chain is a sequence of random variables or vectors, X_i, for $i = (0, 1, \cdots, N)$, with the property that the transition probability

$$Pr(X_{N+1}|X_0, \cdots, X_N) = Pr(X_{N+1}|X_N). \tag{B.5}$$

Equation (B.5) means that the future state of the chain does not depend on the entire past states, but only on the present state of the process (MCMC).
MCMC method is easily applicable to model with a large number of parameters (Balan (2014)).

Appendix C

Probability Distribution

C.0.1 Gaussian(Normal, Bell-curve) Distribution $\mathcal{N}(\mu, \sigma^2)$

The probability density of the Gaussian distribution (shown in Figure (C.1)) is defined by,

$$Pr(x) = f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\mu}} e^{-} \frac{(x-\mu)^2}{2\sigma^2}$$

where $\mu \in \mathbb{R}$ indicates the mean (expectation) of the distribution, σ indicates the standard deviation and $\sigma^2 > 0$ indicates the variance (Ebden *et al.*, 2008).

If the given variable is discrete, then the corresponding node is also discrete and is uniformly distributed unless other information is given.

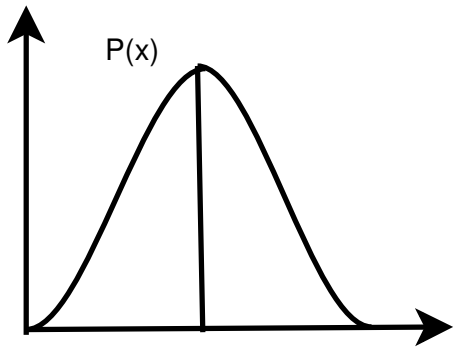

Figure C.1: The probability of the Gaussian distribution.

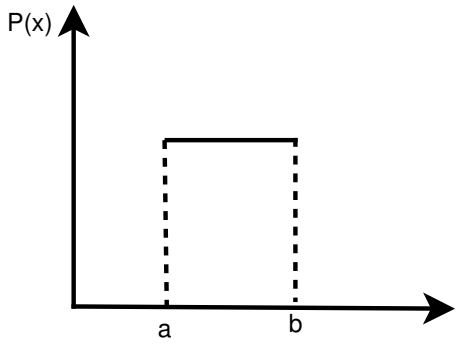

Figure C.2: The probability of the Uniform distribution.

C.0.2 Uniform Distribution $U(a,b)/Unif(a,b)$

The probability distribution associated with $U(a,\ b)$ is defined as follows (Figure (C.2), (Kritzer *et al.*, 2014))

$$Pr(x \mid a,b) = \frac{1}{b-a} \ \ \forall \ \ a \leq x \leq b,$$

A special case of this distribution, when $b = \frac{1}{2\sigma\sqrt{3}}$ and $a = -b$ is given by

$$Pr(x \mid \sigma) = \frac{1}{2\sigma\sqrt{3}} \ \ \forall \ \ -\sigma\sqrt{3} \leq x \leq \sigma\sqrt{3}.$$

Appendix D

Beta Distribution

The Beta distribution is a continuous probability distribution that can be used to represent proportion or probability outcomes,

$$Pr(a \leq x \leq b) = \int_a^b f(x)dx \tag{D.1}$$

where $f(x)$, the density function, is

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{\beta(a,b)} \tag{D.2}$$

and a and b are the parameters of Beta distribution and Beta distribution is parameterized by $\beta(a,b)$. These parameters are the exponents of the random variable which control the shape of the distribution. Moreover, a and b describe the probability that Pr, the probability, gets on certain value. Beta distribution is known as conjugate prior for the binominal distribution (discrete nodes). It means that if the likelihood function is binominal, then a beta prior will give a beta posterior (Gupta & Nadarajah, 2004).

Appendix E

The Performance of Observer-based Residuals for Detecting Intermittent Faults: the Limitations

For each system by keeping as many factors as possible the same, such as input u, residual speed of response, residual design parameters and observer design parameters, try to simulate the residual performance as system complexity or n increases.

The effectiveness of the nonlinear observer-based residuals is limited by the system complexity. The evidence has been shown that both residual effectiveness and quality of residual performance decreases as n increases. Residuals effectiveness can change with fault position when n is fixed. However, residual effectiveness is not only dependent on these two factors.(Sedighi *et al.*, 2014).

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 22 (2014) 65 – 70

www.elsevier.com/locate/procedia

3rd International Conference on Through-life Engineering Services

The Performance of Observer-based Residuals for Detecting Intermittent Faults: the Limitations

T. Sedighi[a], P. D. Foote[b], S. Khan[a,b,*]

[a]*EPSRC centre in Through-life engineering services, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, UK, (e-mail: t.sedighi@cranfield.ac.uk.*
[b]*Enhanced Composites and Structures Centre, Manufacturing and Materials Department, School of Applied Sciences, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, UK (e-mail: p.d.foote@cranfield.ac.uk)*
[c] *EPSRC centre in Through-life engineering services, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, UK,(e-mail: samir.khan@cranfield.ac.uk)*

[*] Corresponding author. Tel.: +44-781-659-6270; fax: +0-000-000-0000. *E-mail address:* t.sedighi@cranfield.ac.uk

Abstract

In this paper a broad nonlinear system is considered. Attention is focused upon both performance of a high-gain observer-based residual and the investigation of residual effectiveness for detecting faults in actuators/components. Residual performances for different fault positions and various system complexities are compared. Both qualitative and quantitative evidence for selected fault positions indicated the performance and the effectiveness of the residuals decrease by ascending the system complexity. The poor performance of residuals in the more complex system may cause **No Fault Found (NFF)**. The methods may be extended to the more general class of nonlinear systems and different observers. Efficiency of the proposed approach is demonstrated through the intermittent failure case in a vehicle suspension system.

Peer-review under responsibility of the Programme Chair of the 3rd InternationalThrough-life Engineering Conference
Keywords: Nonlinear control systems; Fault detection; Observer-based residual, Mass-Spring-Damper system; Intermittent faults; No Fault Found; Vehicle suspension system.

1. Introduction

Faults are generally categorized according to whether they have developed slowly during the operation of a system usually characteristic of gradual component wear (incipient fault); arisen suddenly like a step change as a result of a sudden breakage (abrupt faults); or accrued in discrete intervals attributed to component degradation or unknown system interactions (intermittent faults). Intermittent faults can manifest in any system, mechanical or electronic, in an unpredictable manner, and if left unattended over time they may evolve into serious and persistent faults. The assumed unpredictability of an intermittent fault means that it cannot be easily predicted, detected nor is it necessarily repeatable during maintenance testing, thus faults of this nature raise many concerns in the realm of through-life engineering of products [1]. However, an intermittent fault, which is missed during standardised maintenance testing, by its very definition will reoccur at some time in the future. This therefore poses an ever increasing challenge in the maintenance of electronic, mechanical and hydraulic equipment. A substantial portion of malfunctions attributed to intermittent faults as tested healthy and may be categorized as No Fault Found (NFF) [2]. When the fault is not intermittent and the fault symptoms are consistent (hard fault), it is not difficult to isolate and repair. However, a fault that persists for a very short duration and manifests itself intermittently and only during a particular set of operational stresses can be extremely difficult to identify and isolate [3]. In general, intermittent faults typically tend to worsen with time, until eventually becoming substantial enough to be detected with conventional test equipment [4]. Hence, developing the capability for early detection and isolation of the intermittent fault can help to avoid major system breakdowns [5]. Faults can occur in actuators, process components or the sensors. Sensor faults are of particular importance, as they could affect the system performance, or result in a catastrophic mechanical failure. Model-based fault detection schemes can be powerful tools in determining sensor and actuator faults. The concept is to compare the behaviour of an actual process to that of a nominal fault-free model of the process driven by the same input signals. Model-based approaches are more powerful than data-driven signal-processing-based approaches [6] because they rely much more upon physical knowledge of the process and its interactions whereas signal processing techniques rely on large quantities of data to be recorded that may not be practical.

A model-based fault detection scheme consists of two main stages: residual generation and residual evaluation. The objective of designing residuals is to define a signal that can be

Peer-review under responsibility of the Programme Chair of the 3rd InternationalThrough-life Engineering Conference
doi:10.1016/j.procir.2014.07.012

compared to the appropriate measurements and estimations and then evaluated for possible presence of faults [7]. Early research on fault diagnosis, based on software and hardware, have been given. The robust observer-based method of generating residuals based on software is well-known. Such residuals have been designed based on adaptive observers [8], sliding-mode observers [8], bilinear observers [9], quasi-linear observers [10], neural-network-based adaptive observers [11], nonlinear high-gain observers [12], nonlinear canonical form observers [13] and nonlinear observer based on the existence of linearizing transformations [14]. Robust observer based-residuals, based on polynomial models, have been found effective, specially for hydraulic systems [15, 16] and the residuals generated by high-gain observers [18], have wide applicability. However, a wide study of the effectiveness of the residuals for higher dimensional nonlinear systems, with few output measurement has not been attempted. Limited evidence from a study of multi-tanks hydraulic benchmark system [19], food chain system and pipeline system [20] show that residual performances degrades significantly when system complexity increase. This issue and its relation to NFF with more details, will be addressed for the vehicle suspension systems in this paper.

The main object of the paper is to examine the effectiveness of a well established high-gain observer-based residual to detect the intermittent faults. In addition it is shown that the poor performance of residuals for more complex system may cause NFF events.

This paper is constructed as follows: Section 2 gives a system description, derives several models and maps and considers equilibrium points and control. Design of the observer-based residual for the considered system is addressed in Section 3 while a numerical example is provided in Section 4 to investigated the limits to fault detection as system complexity increases. Conclusions are presented in Section 5.

Nomenclature

C_n	restoring force of damper
e	error
f_i	intermittent fault
f_s	sensor fault
F_n	restoring force of spring
g	gravity
g_s	nonlinearity
c	damper constant
k	spring constant
L	length of spring
m	mass
n	number of masses
N	dimension of the system
n_c	choice of output
p_n	fault position
r_n	displacement
r_s	residual
u_n	applied control
y	output
η_y	additive offset

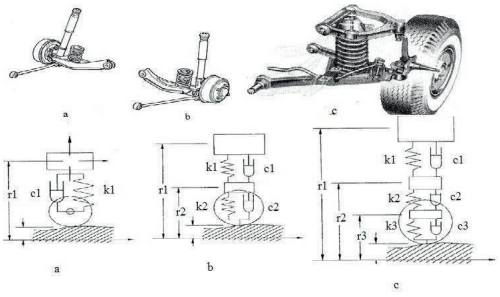

Fig. 1: The model vehicle suspension system.

2. System Description

Consider the class of nonlinear systems defined by the state-space form:

$$\begin{aligned} \dot{x}(t) &= h_x(x, u, g_s, f_i) \\ y(t) &= h_y(x, f_s) \end{aligned} \tag{1}$$

If the nonlinear function $h_x(x, u, g_s, f_i)$ is differentiable with respect to the state $x(t)$, then this class of the system may be expressed in terms of a linear unforced part, and nonlinear state dependent controlled part [21] and [22]:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + S g_s(x, u, t) + K_i f_i(t) \\ y(t) &= Cx(t) + K_{ss} f_s(t) \end{aligned} \tag{2}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ represent the state, input and output vectors respectively. $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $S \in \mathbb{R}^{n\times s}$, $K_i \in \mathbb{R}^{n\times r}$ and $K_{ss} \in \mathbb{R}^{p\times i}$ are known matrices, f_i and f_s present the intermittent and sensor faults respectively. The function $g_s(x, u, t) \in \mathbb{R}^s$ represents the known nonlinearity function.

To illustrate the application of the results obtained in sections 2 – 4, consider the dynamic characteristics of a model vehicle suspension system treating it as a Mass-Spring-Damper (M-S-D) system shown in Figure 1 and 2 where n masses, springs and dampers are connected together in series [23] and [24].

More thorough analysis of a full suspension system are quite complex involving all four tire/suspension systems acting independently. The quarter-car suspension model can be considered in the three levels of complexity shown in Figure 1. The one-degree of freedom model shown in Figure 1*a* considers displacement r_1 of the sprung mass m_1 of the vehicle and the primary suspension stiffness k_1 and damping c_1 only. Here the unsprung mass (mass of the wheels and other components such as lower control arms) and the mass of the tires are not considered. The two degree of freedom model shown in Figure 1*b* accounts for the dynamics of the unsprung mass as well and introduces a second equation of motion and degree of freedom for the displacement r_2 of the unsprung mass m_2, springs and dampers with k_2 and c_2. In this model, the tires are massless. A three-degree of freedom model is shown in Figure 1*c* where the dynamics of the tires are added to the analysis by treating them as a mass spring damper as well (see Figure 2), [25], and [26].

A mass-spring-damper system is usually modeled by a set of differential equations. The system comprises of a finite number

of masses, springs and dampers on a line. In fact it is assumed that n masses, springs and dampers are connected together serially. The model that will be developed here could be extended so that the user is able to select any number of springs, dampers and masses to connect together to build the final system.
The dynamic of the n-th mass is given by

$$\begin{aligned}\ddot{r}_n &= (m_n)^{-1}\Big[-c_n(\dot{r}_n-\dot{r}_{n-1})-k_n(r_n-r_{n-1}-L_n)\\ &\quad +c_{n+1}(\dot{r}_{n+1}-\dot{r}_n)+k_{n+1}(r_{n+1}-r_n-L_{n+1})\\ &\quad +g+u_n+g_s+f_i\Big] \end{aligned} \tag{3}$$

where m_n represents the mass of n^{th} mass, r_n represents the displacement from a reference position of the n^{th} mass, c_n represents the restoring force of n^{th} damper, k_n represents the stiffness of n^{th} spring, L_n represents the length of n^{th} spring, g represents the gravity, u_n is the control applied on the n^{th} mass, g_s is the nonlinearity and f_i is the possible fault in the system. Hence the dynamic of the n^{th} mass with n degrees of freedom may be rewritten in the following form

$$m_n\ddot{r}_n = F_{n+1}-C_{n+1}+F_n-C_n+g+u_n+g_s+f_i \tag{4}$$

where F_n represents the restoring force of the n^{th} spring and C_n represents the restoring force of the n^{th} damper.
For relatively small displacements, restoring forces in (4), can be considered as linear function of displacements

$$\begin{aligned}F_n &= k_n(r_n-r_{n-1}-L_n)\\ C_n &= c_n(\dot{r}_n-\dot{r}_{n-1}).\end{aligned} \tag{5}$$

Also a situation in which the spring and damper restoring forces depend nonlinearly on displacement, hardening spring, where, beyond a certain displacement, large force increments are obtained for small displacement increments, case (5), can be rewritten as:

$$\begin{aligned}F_n &= k_{n_1}(x_n-x_{n-1}-L_n)+k_{n_2}(x_n-x_{n-1}-L_n)^3\\ C_n &= c_{n_1}(\dot{x}_n-\dot{x}_{n-1})+cn_2(\dot{x}_n-\dot{x}_{n-1})^3\end{aligned} \tag{6}$$

see [27].

To obtain the state-space equation of the M-S-D system define

$$X_n = r_n - L_n \tag{7}$$

from Figure (2b), where X_n is the amount of the stretch of the corresponding spring. Then the displacement r may be represented as

$$r = WX + L \tag{8}$$

where $r = \begin{bmatrix} r_1 & \dots & r_n \end{bmatrix}^T$, $X = \begin{bmatrix} X_1 & \dots & X_n \end{bmatrix}^T$, $L = \begin{bmatrix} L_1 & \dots & L_n \end{bmatrix}^T$ are extended vectors and

$$W = \begin{pmatrix} 1 & 0 & \dots & 0\\ 1 & 1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 1 & 1 & \dots & 1 \end{pmatrix}. \tag{9}$$

Hence, the the system equation in W-form is presented as

$$\begin{aligned}W\ddot{x} &= M^{-1}(-CW\dot{X})-M^{-1}(K(WX+L))+M^{-1}\bar{g}_s\\ &\quad +M^{-1}g+M^{-1}\bar{U}+M^{-1}\bar{f}_i\end{aligned} \tag{10}$$

with $\bar{g}_s = \begin{bmatrix} g_{s_1} & \dots & g_{s_n} \end{bmatrix}^T$, $\bar{U} = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix}^T$, $\bar{f}_i = \begin{bmatrix} f_{i_1} & \dots & f_{i_n} \end{bmatrix}^T$ and

$$M = \begin{pmatrix} m_1 & 0 & 0 & \dots & 0\\ 0 & m_2 & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ 0 & 0 & 0 & \dots & m_n \end{pmatrix}.$$

Now define $w = W(X-\bar{X})$, where $\bar{X}$ is the equilibrium point. Then the system (10) in w-form is

$$\begin{aligned}\ddot{w} &= M^{-1}(-C\dot{w})+M^{-1}(-K(w+W\bar{X}))+M^{-1}\bar{g}_s\\ &\quad +M^{-1}g+M^{-1}\bar{U}+M^{-1}\bar{f}_i.\end{aligned} \tag{11}$$

Note that at equilibrium ($U=0$, $w=W\bar{X}$, $\dot{w}=0$ and $\ddot{w}=0$), (11) will find the following form

$$M^{-1}KW\bar{X} = M^{-1}g. \tag{12}$$

Thus the system equation (11) can be rewritten as

$$\ddot{w} = -M^{-1}C\dot{w}-M^{-1}Kw+M^{-1}\bar{g}_s+M^{-1}\bar{U}+M^{-1}\bar{f}_i \tag{13}$$

or equivalently, in terms of its state space representation

$$\begin{aligned}\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix} &= \begin{pmatrix}0 & I_n\\ -M^{-1}K & -M^{-1}C\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\\ &\quad +\begin{pmatrix}0\\ M^{-1}\bar{g}_s\end{pmatrix}+\begin{pmatrix}0\\ M^{-1}\bar{f}_i\end{pmatrix}+\begin{pmatrix}0\\ M^{-1}\bar{U}\end{pmatrix}\end{aligned} \tag{14}$$

where $x_1 = w$ and $x_2 = \dot{w}$. The system output is $y = Cx + \eta_y$ where $C \in \Re^{n\times n}$ and η_y is an additive offset (output error/sensor fault) on each output.
When the system is consist of two messes, springs and dampers, then the system equation (14) without fault terms is presented as

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= \frac{1}{m_1}[-k_1x_1-c_1x_2+k_2x_3+c_2x_4+u_1\\ &\quad +g_{s_1}(x_1,x_2,x_3,x_4)]\\ \dot{x}_3 &= x_4\\ \dot{x}_4 &= \frac{1}{m_2}[k_1x_1+c_1x_2-2k_2x_3-2c_2x_4-u_1+u_2\\ &\quad -g_{s_1}(x_1,x_2,x_3,x_4)+g_{s_2}(x_1,x_2,x_3,x_4)].\end{aligned} \tag{15}$$

3. Fault Detection Filter

Not all the states $x(t)$ can be directly measured (as is commonly the case), therefore we can design an observer, $\hat{y}(t)$ to estimate them, while measuring only the output $y(t)$. The observer is basically a model of the plant; it has the same input and follows a similar differential equation. An extra term compares the actual measured output $y(t)$ to the estimated output of the observer $\hat{y}(t)$; minimising this error will cause the estimated states $\hat{x}(t)$ to tend towards the values of the actual real-system states $x(t)$. It is conventional to write the combined equations for the system plus observer using the original state x(t) plus the error state [19],

$$e(t) = x(t)-\hat{x}(t). \tag{16}$$

In general the fault detection system consists of two parts, 1) residual generation, 2) residual evaluation [28].

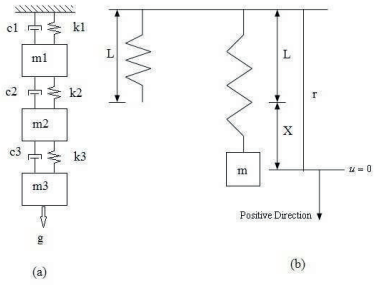

Fig. 2: The mass-spring-damper system.

3.1. Residual Generation

While a suitable observer is chosen for every case, if the error system stability is satisfied, then the following scalar observer-based residual can be generated for each output to detect the intermittent faults

$$r_s(t) = (y(t) - \hat{y}(t)) = \zeta Ce(t) + \zeta K_{ss} f_s(t) \tag{17}$$

where $\zeta \in \mathbb{R}^{n \times p}$, is a suitable weighting matrix to be designed. The problem can be stated as finding ζ, such that the following aims are achieved [29]:

- The effect of unknown input and disturbance signals on the residual signal are as small as possible while the effect of fault signal is as large as possible.
- The effect of parametric uncertainties on residual signal are as small as possible.
- The fault detection system is robust stable in the presence of exogenous signals and uncertainties.

The object is to show that the residuals are differing from zero when faults have occurred; however, the residual tends to zero in "no fault" situation.

3.2. Residual Evaluation

A common choice of evaluation signal is the 2-norm:

$$r_{s_{eval}} = \|r_s\|_2 \triangleq \sqrt{\int_0^\infty |r_s(\tau)|^2 d\tau}. \tag{18}$$

Since the evaluation function (18) can not be realised exactly, because the value of $\|r_s\|_2$ is not known until $t = \infty$, and it is reasonable to assume that faults could be detected, if they occur over finite time interval. Therefore equation (18) could be modified to

$$r_{s_{eval}} = \|r_s(t)\|_2 \triangleq \sqrt{\int_0^t |r_s(\tau)|^2 d\tau} \tag{19}$$

where τ is the time window and it is finite [27].

4. Simulation Outcomes

In this section the object is to investigate several M-S-D systems with the aim of showing a system dependent phenomenon which limits the effectiveness of the observer-based residuals.

4.1. Intermittent fault

Collapsing suspension due to coil spring failure seems to be a growing problem - caused by a combination of recent harsh winter conditions and weight-saving designs. A plastic coating is applied to coil springs when they are made to reduce the risk of corrosion. Over time, contact between coils as the spring is repeatedly compressed in service can cause damage to this coating. Most often coil spring failure seems to be caused by corrosion, accelerated by salt applied to the roads in winter. In other hand Electrolytic action between the salt solution, formed by road salting, and the iron in the spring generates free hydrogen atoms which enter the steel and can cause microscopic cracking. Cracks propagate and combine, ultimately leading to failure of the spring, (www.theaa.com).
Cracks and corrosion both can be classified as Intermittent faults, which will start with small failure in a short time but will get stronger and longer until it ends up with complete spring failure eventually.

Assume that at each failure, the length of spring will change unexpectedly. In other word a fault in position i is a change in the length of the i-th spring, so that $L = (L_0 + f_i L_0)$ in position i, and $L = L_0$ in all other sections of the system, where L_0 is the initial length of the spring.
Also, it is defined that $f_i(t)$ is a time varying of the form $f_i(t) = dd_i y_{n_c}(t)$, where dd_i, the maximum fault amplitudes, are constant and y_{n_c} is the designers's choice of output.
Hence the intermittent fault, $f_i(t)$, could be generated as combination of impulses at different amplitudes which will occurred in discrete intervals. We could model the fault as follows

$$f_i(t) = \begin{cases} 0 & for \quad 0 \le t < 55s \\ f_{i_1} & for \quad 55s \le t < 60s \\ 0 & for \quad 60s \le t < 120s \\ f_{i_2} & for \quad 120s \le t < 145s \\ 0 & for \quad 145s \le t < 190s \\ f_{i_3} & for \quad 190s \le t < 260s \\ 0 & for \quad 260s \le t < 270s \\ f_{i_4} & for \quad 270s \le t < 400s \end{cases} \tag{20}$$

where $dd_1 = 0.0025$, $dd_2 = 0.01$, $dd_3 = 0.15$ and $dd_4 = 0.25$ are constants, $n_c = 1, \cdots, n$ is the choice of output and t indicates the time.

4.2. Simulation Conditions

For numerical example a general $N = 2n$ dimension M-S-D system is considered. This system has a maximum of n inputs and n outputs. There is no disturbances, and the effect of a single fault $f_i(t)$ depends on the choice of K_i in (2). Here $\eta_y(t)$ is the term defining the sensor fault $f_s(t)$ of the form $\eta_0 sin(t)$ and the intermittent fault is of the form (20).
Note that the length of spring, L, is $1m$, the mass m is $1Kg$ and $g = 9.8\frac{N}{m^2}$ is the gravity.

4.3. Residual effectiveness investigation

For each system by keeping as many factors as possible the same, such as input u, residual speed of response, residual

design parameters and observer design parameters, try to simulate the residual performance as system complexity or n increases.

To investigate the residual effectiveness for increasing n the following steps are performed:

1. An intermittent fault of the form (20) is applied, in the one of the springs, p_i, $(1 \leq i \leq n)$.
2. An additive offset (output error), η_y, is made on each output y_i, so that $y_i(t) = x_i(t) + \eta_y$ where $\eta_y = \eta_0 sin(t)$.
3. The residual of the form (17) with corresponding observer is designed.
4. The number η_0, bounding the output error, is varied until the effect of the fault $f_i(t)$ on the residual, denoted by $R_{f_{max}}$ in Table 1, is approximately equal to the effect of the noise, η_y, on the residual denoted by R_{η_y} in Table 1. This condition is denoted by $\frac{\|R_f\|_{max}}{\|R_\eta\|_{max}} = 1$. It indicates a limit on the error. Increasing η_0 further means that, if η_y is present throughout the time frame, its effect on the residual (17) would mask the effect due to the fault $f_i(t)$. This condition may cause No Fault Found events, [19].
5. Finally the steps are repeated for different number of masses.

4.4. Results of investigation

The numerical results for the system mentioned above are summarised in Table 1 and Figures 3 – 7.

For M-s-D system, Figures 3 – 7 show, some results in graphical form when implementing the residual with a fault $f_i(t)$ is presented. For each case a nearly limited condition for η_0 is chosen. Table 1 is driven from the data and compares some important numbers for each n. for some cases the residual $R_{f_{max}}$ is so small and that it can not be distinguished from modeling error and control effects without filtering action.

Table 1: R_{fmax}, R_{η_y} and η_0 for fault $f_i(t)$, where $n_c = 1$ and $l = 1m$.

n	P_n	$R_{f_{max}}$	$R_{\eta_{ymax}} = \frac{\eta_0}{R_{f_{max}}}$	η_0
2	1	2.67×10^{-4}	0.149×10^{4}	0.04
2	2	2.10×10^{-5}	5.86×10^{5}	1.2
3	1	3.07×10^{-4}	0.048×10^{4}	0.15
3	2	5.82×10^{-5}	0.154×10^{5}	0.9
3	3	2.96×10^{-5}	0.405×10^{5}	1.2
4	1	3.64×10^{-4}	0.68×10^{4}	0.25
4	2	1.062×10^{-4}	0.68×10^{4}	0.73
4	3	7.14×10^{-5}	0.14×10^{5}	1
4	4	3.66×10^{-5}	0.34×10^{5}	1.25
5	1	4.60×10^{-4}	0.65×10^{4}	0.3
5	2	1.56×10^{-4}	0.38×10^{4}	0.6
5	3	1.13×10^{-4}	0.70×10^{4}	0.8
5	4	7.62×10^{-5}	0.14×10^{5}	1.1
5	5	3.93×10^{-5}	0.33×10^{5}	1.3
6	1	7.05×10^{-4}	0.049×10^{4}	0.35
6	2	2.48×10^{-4}	0.24×10^{4}	0.59
6	3	1.85×10^{-4}	0.43×10^{4}	0.8
6	4	1.43×10^{-4}	0.67×10^{4}	0.97
6	5	9.95×10^{-5}	0.11×10^{5}	1.1
6	6	5.19×10^{-5}	0.25×10^{5}	1.3

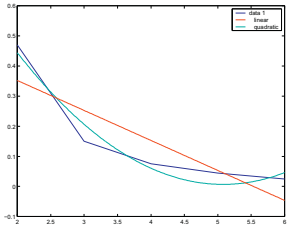

Fig. 3: η_0 against $n = 6$, $n_c = n$ and $P_n = 2$.

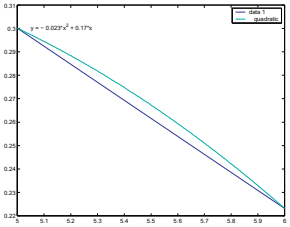

Fig. 4: η_0 against $n = 6$, $n_c = n$ and $P_n = 5$.

Note that a straight line can be fitted to the data as a limiting line for fault detection ($log(\eta_0)$). A quadratic curve is more accurate but is poor when extrapolations are made for more complex systems, [19]. Figures 2 and 3 display η_0, as the data pair $(log(\eta_0), n)$ is fitted for each system. These graphs clearly show that the effectiveness of the residuals decreases as n increases.

In Figure 2, consider that η_y is determined as a sensor fault, then if the magnitude of $\log \eta_0$ is grater than 0.3 or (equivalently $\eta_0 \geq 0.0003$), a fault in the second position is only detectable for a system with 2 masses or less. But in Figure 3, a fault in position 5 is detectable for a system with 5 masses or less. For the system with more masses, residual cannot detect the fault $f_i(t)$, and is masked by the effect of the sensor fault $f_s(t)$.

Figures 5 – 7 show that the intermittent faults detection may be delayed due to the effect of the sensor faults causing No Fault Found (NFF). Figures 5 and 7 show that the complete masking due to the effect of the sensor fault has been occurred and made the intermittent fault detection impossible. Figure 6 also show that a sensor fault has masked part of the intermittent fault.

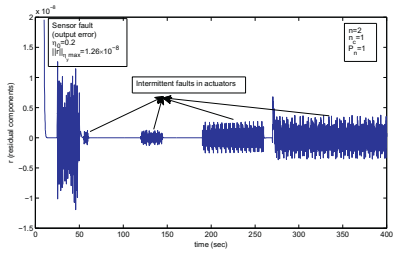

Fig. 5: Sensor faults masked the intermittent fault, $n = 2$, $n_c = 1$ and $P_n = 1$ for $\eta_0 = 0.1$.

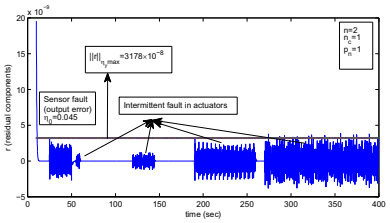

Fig. 6: Sensor faults masked part of the intermittent fault, $n = 2$, $n_c = 1$ and $P_n = 1$ for $\eta_0 = 0.045$.

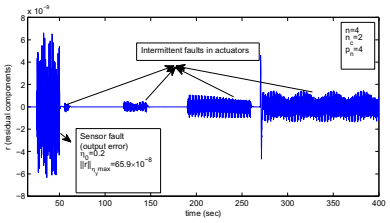

Fig. 7: Sensor faults masked the intermittent fault, $n = 4$, $n_c = 2$ and $P_n = 4$ for $\eta_0 = 0.2$.

5. Conclusions

The development of state-space models and several transformations for use in satisfying the objectives concerning fault detection for a nonlinear M-S-D systems have been discussed. In particular, an observer-based residual has been proposed for the state space model.

An extensive investigation has been made into the effectiveness and performance of the residuals based on an observer design. For each system and a fixed controller, a specific form of sensor fault and a specific input structure, residual performance has been investigated for detecting the intermittent faults of different positions.

The effectiveness of a nonlinear observer-based residuals has been shown to be limited by the system complexity. The evidence has been shown that both residual effectiveness and quality of residual performance decreases as n increases. Residuals effectiveness can change with fault position, when n is fixed. However, the residual effectiveness is not only dependent on these two factors.

The simulations also show that the sensor fault may be able to mask the effect of the intermittent faults in the actuator/components, resulting a very late detection of the intermittent faults and NFF.

Future investigation is needed to compare the performance of different observers to detect the intermittent faults as one of the main root causes of NFF in the presence of the sensor faults and unknown inputs/disturbances.

References

[1] Sedighi T., Phillips P., and P. Foote. Model-based intermittency fault detection. *2nd International Through-Life-Engineering Services Conf.*, volume 11, 68–73, 2013 .

[2] Hockley C., and P. Phillips. The impact of No-Fault-Found (NFF)on through-life engineer services. *Int. J. of Quality in Maintenance Engineering*, volume 18, 141–153, 2012.

[3] Sedighi T., Phillips P., and P. Foote. Unknown Input Observer-Based Intermittent Fault Detection. *Submitted to int. J. of Control*, 2013

[4] Steadman B., Berghout F., and N. Olsen. Intermittent fault detection and isolation system. *IEEE Autotes Conf.*, 2008.

[5] Qi H., Ganesan S., and F. Pecht. No fault found and intermittent failure in electronic products, 2008.

[6] Frank P.M. Analytical and qualitativemodel-based fault diagnosis: a survey and some new results. *Europe J. Control*, volume 2, 6–28, 1996.

[7] Abid M. Fault detection in nonlinear systems: an observer-based approach. *PhD thesis*, 2012.

[8] Marino R., and P. Tomie. Adaptive observer with arbitrary exponential rate of converges for nonlinear systems. *IEEE, Trans. Autom. Control*, volume 40, 1300–1304, 1995.

[9] Edwards C., Spurgeon, S.K. and R.J. Patton. Sliding mode observer for fault detection and isolation. *Automatica*, volume 36, 541–553, 2005.

[10] Yu D., and D.N. Shields. A bilinear fault detection observer. *Automatica*, volume 32, 1597–1602, 1997.

[11] Edelmajer A., Boker J., Szigeti F., and L. Kevicsky. Robust detection filter design in the precence of time-varing system pertubations. *Automatica*, volume 33, 471–475, 1997.

[12] Young H.K., Frank L.L., and T.A. Chaoki. A dynamic recent neural-network-based adaptive observer for a class of nonlinear systems. *Automatica*, volume 33, 1539–1543, 1997.

[13] Nijmeijer H., and A. Van Der Schaft. *Nonlinear Dynamic Control systems*. spring-Verlag, New York, 1990.

[14] Xia X.H., and W.B. Gao. Nonlinear observer design by observer error linearization. *SIAM J. of control and optimization*, volume 27, 199–216, 1989.

[15] Krener A.J., and A. Isodori. Linearization by output injection and nonlinrear observers. *System and control letters*, volume 3, 47–52, 1983.

[16] Shields D.N., and S. Du. Fault detection observer for continuous nonlinear systems of general degree. *Int. Jnl. of control*, volume 76, 953–962, 2003.

[17] Shields D.N. and S. Daley. A quantitative fault detection method for a class of nonlinear systems. *Trans. Inst. of Measurement and Control (MC)*, volume 20, 125–133, 1998.

[18] Hammouri H., Kinneart M., and E. Yaagoubi. Observer-based approach to fault detection and isolation for state affine systems. *European Journal of Control*, volume 4, 2–16, 1999.

[19] Shields D.N. Models, residual design and limits to fault detection for a complex multi-tank hydraulic control system *Int. Jnl. of Control*, volume 76, 781–793, 2003.

[20] Shields D.N., Ashton S.A., and S. Daley. Design of observer for detecting faults in hydraulic sub-sea pipelines, *Control Engineering Practice*, volume 9, 297–311, 2001.

[21] Aldeen M., and R. Sharma. Estimation of states, faults and unknown disturbances in nonlinear systems. *Int. Jnl. of Control*, 2006.

[22] Rajmani. Observer for Lipschitz nonlinear systems. *IEEE Trans. on Automatica Control*, volume 43, 397–400, 1998.

[23] Patton R., Frank P., and R. Clark. *Fault Diagnosis in Dynamic Systems, Theory and Applications*. Prentice Hall, London, 1989.

[24] Figliola R.S., and D.E. Beasley. Theory and design for mechanical measurments, *4th Ed. Wiley*, 199–201, 474–477 and 482–489, 2006.

[25] Gillespie, T.D. Fundemental of vehicle dynamics, *Society of Automative Engineer, Inc.*, 1992.

[26] Genta G. Motor vehicle dynamics: model and simulation, *World Scientific publishing*, 1997.

[27] Sedighi T., D.N. Shields. The performance of the high-gain observer-based residual for detecting faults in a mass-spring-damper system. *UKACC Conf.*, Glasgow, 2006.

[28] Puig, V., Oca S.M., and J. Blesa. Adaptive threshold generation in robust fault detection using interval models: time-domain and frequency-domain approaches. *I*nternational Journal of Adaptive Control and Signal Processing, 2012.

[29] Ahmadizadeh S., Zarei J., and H.R. Karimi. A robust fault detection design for uncertain Takagi-Sugeno models with unknown inputs and time-varying delays, *Science direct*, volume 11, 98–117, 2014.

References

Abdelghani, Maher, & Friswell, Michael I. 2007. Sensor validation for structural systems with multiplicative sensor faults. *Mechanical systems and signal processing*, **21**(1), 270–279.

Abdi, Hervé, & Williams, Lynne J. 2010. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, **2**(4), 433–459.

Abdo, Ali, Ding, Steven X, Damlakhi, Waseem, & Saijai, Jedsada. 2011a. Robust fault detection filter design for uncertain switched systems with adaptive threshold setting. *Pages 5467–5472 of: 2011 50th ieee conference on decision and control and european control conference*. IEEE.

Abdo, Ali, Ding, Steven, Damlakhi, Waseem, & Saijai, Jedsada. 2011b. Robust fault detection filter design for uncertain switched systems with adaptive threshold setting. *Pages 5467–5472 of: Cdc-ece*. IEEE.

Abid, Muhammad. 2010. *Fault detection in nonlinear systems: an observer-based approach*. Ph.D. thesis, Uni Duisburg-Essen.

Adhyaru, Dipak M. 2012. State observer design for nonlinear systems using neural network. *Appl. soft comput.*, **12**(8), 2530–2537.

Ahmad, Syed Wakil. 2017. Intermittent fault diagnosis and health monitoring for electronic interconnects.

Ahmadizadeh, Saeed, Zarei, Jafar, & Karimi, Hamid Reza. 2013. Robust fault detection of linear uncertain time-delay systems using unknown input observers. *J. applied mathematics*, **2013**.

Ahmadizadeh, Saeed, Karimi, Hamid Reza, & Zarei, Jafar. 2014. Robust fault detection design for unknown inputs takagi-sugeno models with parametric uncertainties and time-varying delays. *Pages 756–761 of: Ecc*. IEEE.

Al-Salami, Ibrahim M., Chabir, Karim, Sauter, Dominique, & Aubrun, Christophe. 2010. Adaptive thresholding for fault detection in networked control systems. *Pages 446–451 of: Cca*. IEEE.

Albert, Jim. 2011. *Learnbayes: Functions for learning bayesian inference*. R package version 2.12,.

Aldeen, M., & Sharma, R. 2008. Estimation of states, fault and unknown disturbances in non-linear systems. *International journal of control*, **81**, 1195–1201.

Alwi, H., Edwards C., & Marcos, A. 2012. Fault reconstruction using a lpv sliding mode observer for a class of lpv systems. *Journal of the franklin institute*, **349**, 510–530.

Alwi, H., Edwards, C., & Tan, C.P. 2009. Sliding mode estimation scheme for incipient sensor faults. *Automatica*, **45**, 1679–1685.

Analysis, Fluid Systems. 2004. *Lumped fluid systems*. `http://www.dartmouth.edu/~sullivan/22files/Fluid%20sys%20anal%20w%20chart.pdf`.

Anderson, Ken, & Synaptics, Universal. 2015. Reducing nff and improving operational availability through intermittent fault detection. *In: Proc. 2015 nff symposium*.

Andrade, Luciano CM, Oleskovicz, Mário, & Fernandes, Ricardo AS. 2016. Adaptive threshold based on wavelet transform applied to the segmentation of single and combined power quality disturbances. *Applied soft computing*, **38**, 967–977.

Andrieu, Christophe, & Atchade, Yves F. 2007. On the efficiency of adaptive MCMC algorithms. *Electronic communications in probability*, **12**, 336–349.

Anilkumar, P. 1994. On estimating the mean function of a gaussian process. *Statistics & probability letters*, **19**(1), 77–84.

Armaou, Antonios, & Demetriou, Michael A. 2008. Robust detection and accommodation of incipient component and actuator faults in nonlinear distributed processes. *Aiche journal*, **54**(10), 2651–2662.

Azarian, Armin. 2009. *A new modular framework for automatic diagnosis of fault, symptoms and causes applied to the automotive industry*. Ph.D. thesis, Uni Karlsruhe.

Baghernezhad, F., & Khorasani, K. 2016. Computationally intelligent strategies for robust fault detection, isolation, and identification of mobile robots. *Neurocomputing*, **171**, 335–346.

Balan, Anoop Korattikara. 2014. *Approximate markov chain monte carlo algorithms for large scale bayesian inference.* Ph.D. thesis, University of California, Irvine, USA. base-search.net.

Baligar, Vishwanath P, Patnaik, Lalit M, & Nagabhushana, GR. 2006. Low complexity, and high fidelity image compression using fixed threshold method. *Information sciences*, **176**(6), 664–675.

Barker, Matthew, & Rayens, William. 2003. Partial least squares for discrimination. *Journal of chemometrics: A journal of the chemometrics society*, **17**(3), 166–173.

Barrau, Axel, & Bonnabel, Silvère. 2016. The invariant extended kalman filter as a stable observer. *Ieee transactions on automatic control*, **62**(4), 1797–1812.

Basseville, M., Abdelghani, M., & Benveniste, A. 2000. Subspace-based fault detection algorithms for vibration monitoring. *Automatica*, **36**, 101–109.

Basseville, Michèle. 1997. Information criteria for residual generation and fault detection and isolation. *Automatica*, **33**(5), 783–803.

Bejarano, F. J., Pisano, A., & Usai, A . 2011. Finite-time converning jump observer for switched linear systems with unknown inputs. *Journal of nonlinear analysys: Hybird systems*, **5**, 174–188.

Boulkroune, Boulaid, Galvez-Carrillo, Manuel Ricardo, & Kinnaert, Michel. 2011. Additive and multiplicative fault diagnosis for a doubly-fed induction generator. *Pages 1302–1308 of: Cca*. IEEE.

Brandherm, Boris, & Jameson, Anthony. 2004. An extension of the differential approach for Bayesian network inference to dynamic Bayesian networks. *International journal of intelligent systems*, **19**(8), 727–748.

Brik, J., & Zeitz, M. 1988. Extended luenberger observer for non-linear multivariable systems. *International journal of control*, **47**, 1823–1836.

Butts, C. T. 2003. Network inference, error, and informant inaccuracy: a bayesian approach. *Social networks*, **25**(2), 103–140.

Butz, C. J., dos Santos, Andre E., de S. Oliveira, Jhonatan, & Gonzales, Christophe. 2016. A simple method for testing independencies in bayesian networks. Lecture Notes in Computer Science, vol. 9673. Springer.

Butz, C.J. 2001. On axiomatizing probabilistic conditional independencies in bayesian networks. *Lecture notes in computer science*, **2198**, 131–140.

Cai, Baoping, Liu, Yu, & Xie, Min. 2016. A dynamic-bayesian-network-based fault diagnosis methodology considering transient and intermittent faults. *Ieee transactions on automation science and engineering*, **14**(1), 276–285.

Cai, Baoping, Huang, Lei, & Xie, Min. 2017. Bayesian networks in fault diagnosis. *Ieee transactions on industrial informatics*, **13**(5), 2227–2240.

Cai, Baoping, Liu, Yonghong, Hu, Jinqiu, Liu, Zengkai, Wu, Shengnan, & Ji, Renjie. 2018. *Bayesian networks in fault diagnosis: Practice and application*. World Scientific.

Castillo, I., Edgar, T. F., & Fernandez, B. R. 2012. Robust model-based fault detection and isolation for nonloinear processes using sliding modes. *International journal of robust and nonlinear control*, **22**, 89–104.

Cecati, Carlo. 2015. A survey of fault diagnosis and fault-tolerant techniques—part ii: Fault diagnosis with knowledge-based and hybrid/active approaches.

Chen, J., & Patton, R. J. 1999. *Robust model-based fault diagnosis for dynamic systems.* Kluwer Academic Publishers. London, Great Britania.

Chen, J., Cao, Y., & Zhang, W. 2015. A fault detection observer design for lpv systems in finite frequence domain. **88**.

Chen, Jie, & Patton, Ron J. 2012. *Robust model-based fault diagnosis for dynamic systems*. Vol. 3. Springer Science & Business Media.

Chen, Jie, Patton, Ron J, & Zhang, Hong-Yue. 1996. Design of unknown input observers and robust fault detection filters. *International journal of control*, **63**(1), 85–105.

Chen, W., & Chowdhury, F. N. 2010. A synthesized design of sliding-mode and luenburger observer for early detection for incipation faults. *International journal of adaptive control and signal processing*, **24**, 1021–1035.

Chen, W., & Saig, M . 2006a. Fault detection and isolation based on novel unknown input observer design. *In: Proceedings of the 2006 american conference*.

Chen, W., & Saig, M . 2006b. Unknown input observer design for a class of nonlinear systems: an lmi approach. *In: Proceedings of the 2006 american conference*.

Chen, Wei. 2011. *Fault detection and isolation in nonlinear systems: observer and energy-balance based approaches*. Ph.D. thesis, Uni Duisburg-Essen.

Cheng, Jie, & Greiner, Russell. 2001. Learning bayesian belief network classifiers: Algorithms and system. *Pages 141–151 of: Conference of the canadian society for computational studies of intelligence*. Springer.

Chickering, David Maxwell. 2013. Learning equivalence classes of bayesian networks structures. *Corr*, **abs/1302.3566**.

Choi, Han Ho, & Chung, Myung Jin. 1996. Observer-based h controller design for state delayed linear systems. *Automatica*, **32**(7), 1073–1075.

Chookah, Mohamed, Nuhi, Mohammad, & Modarres, Mohammad. 2011. A probabilistic physics-of-failure model for prognostic health management of structures subject to pitting and corrosion-fatigue. *Reliability engineering and system safety*, **96**(12), 1601–1610.

Chung, Kai Lai. 2001. *A course in probability theory*. Second edn. San Diego: Academic Press.

Civelek, Cem. 2006. Mathematical modelling of rotational mechanical elements of higher order and their characteristics. *Mathematical and computer modelling*, **43**(7-8), 957–964.

Cobb, Barry R., & Shenoy, Prakash P. 2012. Hybrid bayesian networks with linear deterministic variables. *Corr*, **abs/1207.1369**.

Connor, Nick. 2020 (May). *What is reynolds number for pipe flow – definition*. `https://www.thermal-engineering.org/what-is-reynolds-number-for-pipe-flow-definition/`.

Darouach, Mohamed, Zasadzinski, Michel, & Xu, Shi Jie. 1994. Full-order observers for linear systems with unknown inputs. *Ieee transactions on automatic control*, **39**(3), 606–609.

Das, Sourish, Roy, Sasanka, & Sambasivan, Rajiv. 2018. Fast gaussian process regression for big data. *Big data research*, **14**, 12–26.

De Kleer, Johan. 2009. Diagnosing multiple persistent and intermittent faults. *Pages 733–738 of: Ijcai*.

Dempster, A.P., Laird, N.M., & Rubin, D.B. 1977. Maximum liklihood from incomplete data via the EM algorithm. *Journal of the royal statistical society b*, **39**, 1–38. with discussion.

Deng, Jie. 2002. A bayesian network-based interactive and iterative reasoning for decision support system under uncertainty. *Pages 464–470 of:* Arabnia, Hamid R., & Mun, Youngsong (eds), *Ic-ai*. CSREA Press.

Dennis, J. E., & Woods, D. J. 1987. Optimization on Microcomputers: The Nelder–Mead Simplex Algorithm. *Pages 116–122 of:* Wouk, A. (ed), *New computing enviroments: Microcomputers in large-scale computing*. Philadelphia: Society for Industrial and Applied Mathematics.

Ding, S. X., Frank P. M, & Ding, E. L. 2000. A unified approach to the optimisation of fault detection systems. *International journal of adaptive control and signal processing , ieee transaction*, **14**, 725–745.

Ding, Steven. 2005. Model-based fault diagnosis in dynamic systems using identification techniques, silvio simani, cesare fantuzzi and ron j. patton, springer: London, 2003, 282pp. isbn 1-85233-685-4. *International journal of robust and nonlinear control*, **15**(11), 509–512.

Djeffal, Abdelhamid, Babahenini, Mohamed Chaouki, & Taleb-Ahmed, Abdelmalik. 2017. Fast binary support vector machine learning method by samples reduction. *Ijdmmm*, **9**(1), 1–16.

Doshi-Velez, Finale. 2012. *Bayesian nonparametric methods for reinforcement learning in partially observable domains.* Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.

Ebden, Mark, *et al.* 2008. Gaussian processes for regression: A quick introduction. *The website of robotics research group in department on engineering science, university of oxford*, **91**, 424–436.

Eddine, D., & Belkhiat, C. 2015. Fault tolerant control for a class of switched linear systems using generalized switched observer scheme. *Journal of control engineering and applied informatics*, **17**.

Edwards, C., & Tan, C.P. 2006. Sensor faults tolerant control using sliding mode observers. *Control engineering practice*, **14**, 897–908.

Elliott, Robert J, Aggoun, Lakhdar, & Moore, John B. 2008. *Hidden markov models: estimation and control*. Vol. 29. Springer Science & Business Media.

Emami-Naeini, A., Akhter M., & Rock, S. 1988. Effect of model uncertainty on failure detection: the threshold selector. *Automatic control, ieee transaction*, **33**, 1106–1115.

Ferdowsi, H, & Jagannathan, S. 2013. A unified model-based fault diagnosis scheme for non-linear discrete-time systems with additive and multiplicative faults. *Transactions of the institute of measurement and control*, **35**(6), 742–752.

Ferdowsi, Hasan, & Jagannathan, Sarangapani. 2011. A unified model-based fault diagnosis scheme for nonlinear discrete-time systems with additive and multiplicative faults. *Pages 1570–1575 of: Cdc-ece*. IEEE.

Figliola, Richard S, & Beasley, Donald E. 2014. *Theory and design for mechanical measurements*. John Wiley & Sons.

Figueroa, Gustavo A., & Sucar, Enrique L. 1999. A temporal bayesian network for diagnostic and prediction. *Pages 13–20 of: Proceedings of the fifteenth annual conference on uncertainty in artificial intelligence (uai–99)*. Morgan Kaufmann Publishers.

Fowler, Michael. 2007. Viscosity introduction: Friction at the molecular level. *Lectures on fluids, university of virginia publ*.

Frank, Paul M. 1995. Residual evaluation for fault diagnosis based on adaptive fuzzy thresholds.

Friedman, Nir, & Koller, Daphne. 2003. Being bayesian about network structure. a bayesian approach to structure discovery in bayesian networks. *Machine learning*, **50**(1–2), 95–125.

Gao, Z., Jing B. Shi P. Qian M., & Lin, J. 2012. Active fault tolerant control design for reusable launch vehicle using adaptive sliding mode technique. *Journal of the franklin institute*, **349**, 1543–1560.

Garcia-Alvarez, Diego, Fuente, Maria J., & Sainz, Gregorio I. 2011. Design of residuals in a model-based fault detection and isolation system using statistical process control techniques. *Pages 1–7 of:* Mammeri, Zoubir (ed), *Etfa*. IEEE.

Gasse, Maxime, Aussem, Alex, & Elghazel, Haytham. 2012. An experimental comparison of hybrid algorithms for bayesian network structure learning. *Pages 58–73 of:* Flach, Peter A., Bie, Tijl De, & Cristianini, Nello (eds), *Ecml/pkdd*. Lecture Notes in Computer Science, vol. 7523. Springer.

Gasse, Maxime, Aussem, Alex, & Elghazel, Haytham. 2014. A hybrid algorithm for bayesian network structure learning with application to multi-label learning. *Expert systems with applications*, **41**(15), 6755–6772.

Genta, Giancarlo. 1997. *Motor vehicle dynamics: modeling and simulation*. Vol. 43. World Scientific.

Ghanmi, Nabil, Mahjoub, Mohamed Ali, & Amara, Najoua Essoukri Ben. 2011. Characterization of dynamic bayesian network. *Ijacsa editorial*.

Gillespie, Thomas D. 1992. *Fundamentals of vehicle dynamics*. Vol. 400. Society of automotive engineers Warrendale, PA.

González, Alejandro Rodríguez, Alor-Hernández, Giner, Mayer, Miguel Angel, Robles, Guillermo Cortes, & Pérez-Gallardo, Yuliana. 2013. Application of probabilistic techniques for the development of a prognosis model of stroke using epidemiological studies. *Ijdsst*, **5**(4), 34–58.

Guang-Ren, D., & Patton, R.J. 2001. Robust fault detection using luenberger-type unknown input observers-a parametric approach. *Int. j. systems science*, **32**(4), 533–540.

Guo, Hongwei. 2011. A simple algorithm for fitting a gaussian function [dsp tips and tricks]. *Ieee signal process. mag.*, **28**(5), 134–137.

Guo, S., Zhu F., & Xu, L. 2015. Unknown input observer design for takagi-sugeno fuzzu stochastic system. *International journal of control, automation and systems*, **13**, 1–7.

Gupta, Arjun K, & Nadarajah, Saralees. 2004. *Handbook of beta distribution and its applications*. CRC press.

Hahne, J. M., Bießmann, F., Jiang, N., Rehbaum, H., Farina, D., Meinecke, F. C., Muller, K. R., & Parra, L. C. 2014. Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control. *Ieee transactions on neural systems and rehabilitation engineering*, **22**(2), 269–279.

Hair Jr, Joseph F, Hult, G Tomas M, Ringle, Christian, & Sarstedt, Marko. 2016. *A primer on partial least squares structural equation modeling (pls-sem)*. Sage publications.

Haranadh, G., & Sekhar, C. Chandra. 2008. Hyperparameters of gaussian process as features for trajectory classification. *Pages 2195–2199 of: Ijcnn*. IEEE.

Haykin, Simon, & Network, Neural. 2004. A comprehensive foundation. *Neural networks*, **2**(2004), 41.

Hoshino, Takahiro. 2008. A bayesian propensity score adjustment for latent variable modeling and mcmc algorithm. *Computational statistics & data analysis*, **52**(3), 1413–1429.

Hosseini, Seyedmohsen, Ivanov, Dmitry, & Dolgui, Alexandre. 2019. Ripple effect modelling of supplier disruption: integrated markov chain and dynamic bayesian network approach. *International journal of production research*, 1–20.

Huby, G. 2012. "no fault found: Aerospace survey results. *Copernicus technology ltd*.

Iamsumang, C. 2015. *Computational algorithm for dynamic hybrid bayesian network in on-line system health management applications.* Ph.D. thesis, University of Maryland, College Park, MD, USA.

Imsland, L., Johnsen T. A. Grip H. F., & Fossen, T. I. 2007. On nonlinear unknown input observers applied to lateral vehicle velocity estimation on banked roads. *International journal of control*, **80**, 1741–1750.

Incarbone, Luca, Auzanneau, Fabrice, Hassen, Wafa Ben, & Bonhomme, Yannick. 2014. Embedded wire diagnosis sensor for intermittent fault location. *Pages 562–565 of: Sensors, 2014 ieee*. IEEE.

Jablonski, Aleksander. 1985. Statistical model of an atom in electron scattering calculations. *Physica a: Statistical and theoretical physics*, **129**(3), 591–600.

Jackman, Simon. 2000. Estimation and inference are missing data problems. unifying social science statistics via bayesian simulation. *Political analysis*, **8**, 307–332.

Jensen, F.V., & Nielsen, T.D. 2007. *Bayesian networks and decision graphs*. 2nd edn. New York, NY: Springer.

Jiang, Huaiguang, Zhang, Jun Jason, Gao, David Wenzhong, & Wu, Ziping. 2014. Fault detection, identification, and location in smart grid based on data-driven computational methods. *Ieee trans. smart grid*, **5**(6), 2947–2956.

Jiang, Lian Lian, & Maskell, Douglas L. 2015. Automatic fault detection and diagnosis for photovoltaic systems using combined artificial neural network and analytical based methods. *Pages 1–8 of: Ijcnn*. IEEE.

Joachims, T. 2005. A support vector method for multivariate performance measures. *In: Proceedings of the international conference on machine learning*.

Jung, Sungmin, Moon, Gyubok, Kim, Yongjun, & Oh, Kyungwhan. 2009. Planning based on dynamic bayesian network algorithm using dynamic programming and variable elimination. *Pages 109–114 of:* Gupta, Gourab Sen, & Mukhopadhyay, Subhas Chandra (eds), *Icara*. IEEE.

Kalchbrenner, Nal, Grefenstette, Edward, & Blunsom, Phil. 2014. A convolutional neural network for modelling sentences. *arxiv preprint arxiv:1404.2188*.

Khan, AQ, Ding, SX, & Abid, M. 2008. Residual generation and evaluation scheme for detection of faults in nonlinear systems using convex optimization. *Pages 84–89 of: Proc. of 6th workshop on advanced control and diagnosis*.

Khan, Samir, Phillips, Paul, Jennions, Ian, & Hockley, Chris. 2014a. No fault found events in maintenance engineering part 1: Current trends, implications and organizational practices. *Reliability engineering & system safety*, **123**, 183–195.

Khan, Samir, Phillips, Paul, Hockley, Chris, & Jennions, Ian. 2014b. No fault found events in maintenance engineering part 2: Root causes, technical developments and future research. *Reliability engineering & system safety*, **123**, 196–208.

Khodakarami, Vahid. 2009. *Applying bayesian networks to model uncertainty in project scheduling*. Ph.D. thesis, Queen Mary University of London, UK. British Library, EThOS.

Kim, Charles J. 2014 (Nov. 25). *System and method of detecting and locating intermittent and other faults*. US Patent 8,897,635.

Kincaid, C. 2005 (May). *Guidelines for selecting the covariance structure in mixed model analysis, paper 198-30 in proceedings of the thirtieth annual sas users group conference*. `https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/198-30.pdf`.

Klimenko, Dimitri, Kurniawati, Hanna, & Gallagher, Marcus. 2015. A stochastic process model of classical search. *Corr*, **abs/1511.08574**.

Koenig, Damien. 2005. Unknown input proportional multiple-integral observer design for linear descriptor systems: application to state and fault estimation. *Ieee trans. automat. contr.*, **50**(2), 212–217.

Kong, Young-Bae, Chang, Kyung-Bae, & Park, Gwi-Tae. 2006. Clustering algorithm using bayes' rule in mobile wireless sensor networks. *Pages 1306–1310 of:* Huang, De-Shuang, Li, Kang, & Irwin, George W. (eds), *Icic (2)*. Lecture Notes in Computer Science, vol. 4114. Springer.

Kritzer, Peter, Niederreiter, Harald, Pillichshammer, Friedrich, & Winterhof, Arne (eds). 2014. *Uniform distribution and quasi-monte carlo methods - discrepancy, integration and applications*. Radon Series on Computational and Applied Mathematics, vol. 15. De Gruyter.

Kwoh, C.-K., & Gillies, D. F. 1996. Using hidden nodes in Bayesian networks. *Artificial intelligence journal*, **88**, 1–38.

La, Viet-Phuong, & Vuong, Quan-Hoang. 2019. bayesvl: Visually learning the graphical structure of bayesian networks and performing mcmc with 'stan'. *The comprehensive r archive network (cran)*.

Lawrence, CE. 2005. Gibbs sampling and bayesian inference. *Encyclopedia of life sciences*.

Lee, Colin, & van Beek, Peter. 2017. Metaheuristics for score-and-search bayesian network structure learning. *Pages 129–141 of:* Mouhoub, Malek, & Langlais, Philippe (eds), *Canadian conference on ai*. Lecture Notes in Computer Science, vol. 10233.

Lee, In Soo, Kim, J. T., Lee, J. W., Lee, D. Y., & Kim, Kyung Youn. 2003. Model-based fault detection and isolation method using art2 neural network. *Int. j. intell. syst.*, **18**(10), 1087–1100.

Lefebvre, Dimitri. 2014. Fault diagnosis and prognosis with partially observed petri nets. *Ieee trans. systems, man, and cybernetics: Systems*, **44**(10), 1413–1424.

Lerner, Boaz, & Malka, Roy. 2011. Investigation of the k2 algorithm in learning bayesian network classifiers. *Applied artificial intelligence*, **25**(1), 74–96.

Lewis, D., & Catlett, J. 1994. Heterogeneous uncertainty sampling for supervised learning. *Pages 148–156 of: In proceedings of the eleventh international conference on machine learning*. Morgan Kaufmann.

Li, Jian, & Yang, Guang-Hong. 2016. Simultaneous fault detection and control for switched systems with actuator faults. *Int. j. systems science*, **47**(10), 2411–2427.

Li, Junyi, & Chen, Jingyu. 2014. A hybrid optimization algorithm for bayesian network structure learning based on database. *Jcp*, **9**(12), 2787–2791.

Li, Rong. 2017. *Control theory for classes of nonlinear systems with application in insurance premium-reserve model*. Ph.D. thesis, University of Liverpool.

Li, Shun, Shi, Da, & Tan, Shaohua. 2012. Financial data modeling using a hybrid bayesian network structured learning algorithm. *Ijcini*, **6**(1), 48–71.

Li, Weihua, & Xu, Yabing. 2010. Gearbox incipient fault diagnosis using feature sample selection and principal component analysis. *Ijmic*, **10**(3/4), 246–254.

Li, Xiaobin, Qian, Jiansheng, & Wang, Gai-ge. 2013. Fault prognostic based on hybrid method of state judgment and regression. *Advances in mechanical engineering*, **5**, 149562.

Li, Zhenhai, & Jaimoukha, Imad M. 2009. Observer-based fault detection and isolation filter design for linear time-invariant systems. *Int. j. control*, **82**(1), 171–182.

Liberatore, Sauro, Speyer, Jason L., & Hsu, Andy Chunliang. 2006. Application of a fault detection filter to structural health monitoring. *Automatica*, **42**(7), 1199–1209.

Lin, Paul P, & Li, Xiaolong. 2006. Fault diagnosis, prognosis and self-reconfiguration for nonlinear dynamic systems using soft computing techniques. *Pages 2234–2239 of: International conference on systems, man and cybernetics*, vol. 3. IEEE.

Liu, C. S., & Peng, H. 2002. Inverse-dynamics based state and disturbance observer for linear time-invariant systems. *Journal of dynamic systems, measurment and control*, **124**.

Liu, C., Jiang B., & Zhang, K. 2014. Sliding mode observer-based actuator fault detection for a class linear uncertain systems. *Pages 230–234 of: Guidance, navigation and control conference (cgncc)*.

Liu, J., Wang J., & Yang, G. H. 2005. An lmi approach to minimum sensitivity analysis with application to fault detection. *Automatica*, **41**, 1995–2004.

Lo, Ndeye Gueye, Flaus, Jean-Marie, & Adrot, Olivier. 2019. Review of machine learning approaches in fault diagnosis applied to iot systems. *Pages 1–6 of: 2019 international conference on control, automation and diagnosis (iccad)*. IEEE.

Lorton, A., Fouladirad, Mitra, & Grall, Antoine. 2013. A methodology for probabilistic model-based prognosis. *European journal of operational research*, **225**(3), 443–454.

Lou, Chuyue, Li, Xiangshun, & Atoui, M Amine. 2020. Bayesian network based on adaptive threshold scheme for fault detection and classification. *Industrial & engineering chemistry research*.

Macci, Claudio. 1996. On the lebesgue decomposition of the posterior distribution with respect to the prior in regular bayesian experiments. *Statistics & probability letters*, **26**(2), 147–152.

MacKay, David JC. 1992. Bayesian model comparison and backprop nets. *Pages 839–846 of: Advances in neural information processing systems*.

Magoun, A. Dale. 2000. Data driven statistical methods. *Technometrics*, **42**(2), 218.

Maher, Patrick. 2010. Bayesian probability. *Synthese*, **172**(1), 119–127.

Malandain, Mathias, Maheu, Nicolas, & Moureau, Vincent. 2013. Optimization of the deflated conjugate gradient algorithm for the solving of elliptic equations on massively parallel machines. *J. comput. physics*, **238**, 32–47.

Malhotra, Ruchika, & Bansal, Ankita Jain. 2012. Fault prediction using statistical and machine learning methods for improving software quality. *Jips*, **8**(2), 241–262.

Marcot, Bruce G. 2012. Metrics for evaluating performance and uncertainty of bayesian network models. *Ecological modelling*, **230**, 50–62.

Marom, Gil, Haj-Ali, Rami, Raanani, Ehud, Schäfers, Hans-Joachim, & Rosenfeld, Moshe. 2012. A fluid-structure interaction model of the aortic valve with coaptation and compliant aortic root. *Med. biol. engineering and computing*, **50**(2), 173–182.

Marrison, N. A. 1992. *Real time fault detection and diagnosis in dynamic engineering systems using constraint analysis.* Ph.D. thesis, University of Glasgow, UK. British Library, EThOS.

Marshall, Samuel A. 1978. *Introduction to control theory*. Macmillan International Higher Education.

Martinez-Gardea, Marcela, Guzmán, Iordan J Mares, Lua, Cuauhtemoc Acosta, Di Gennaro, Stefano, & Alvarez, Ivan Vázquez. 2015. Design of a nonlinear observer for a laboratory antilock braking system. *Journal of control engineering and applied informatics*, **17**(3), 105–112.

Martinussen, Torben, & Scheike, Thomas H. 2006. *Dynamic regression models for survival data*. Springer.

Mashkov, Viktor, Fiser, Jirí, Lytvynenko, Volodymyr, & Voronenko, Maria. 2019. Diagnosis of intermittently faulty units at system level. *Data*, **4**(1), 44.

Mattei, Massimiliano. 2001. An lmi approach to the design of a robust observer with application to a temperature control problem for space vehicle testing. *Automatica*, **37**(12), 1979–1987.

Mau, B. 1996. *Bayesian phylogenetic inference via markov chain Monte Carlo methods*. Ph.D. thesis, Wiscontin University.

Meenatchisundaram, S. 2015 (Nov.). *mathematical modeling of liquidlevel systems*. https://www.slideshare.net/meenasundar/class-7-mathematical-modeling-of-liquidlevel-systems.

Mekki, H, Mellit, Adel, & Salhi, H. 2016. Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules. *Simulation modelling practice and theory*, **67**, 1–13.

Melkumyan, Arman, & Ramos, Fabio. 2009. A sparse covariance function for exact gaussian process inference in large datasets. *Pages 1936–1942 of:* Boutilier, Craig (ed), *Ijcai*.

Miljkovic, Dubravko. 2011. Fault detection methods: A literature survey. *Pages 750–755 of: Mipro*. IEEE.

Ming, Yue. 2014. Rigid-area orthogonal spectral regression for efficient 3d face recognition. *Neurocomputing*, **129**, 445–457.

Mondal, Sharifuddin, Chakraborty, G, & Bhattacharyya, K. 2009. Unknown input nonlinear observer for component fault detection and isolation of dynamic systems. *International journal of automation and control*, **3**(2-3), 154–170.

Moraes, R, Barbosa, Ricardo, Duraes, João, Mendes, Naaliel, Martins, Eliane, & Madeira, Henrique. 2006. Injection of faults at component interfaces and inside the component code: are they equivalent? *Pages 53–64 of: 2006 sixth european dependable computing conference*. IEEE.

Mori, Junichi, & Mahalec, Vladimir. 2016. Inference in hybrid bayesian networks with large discrete and continuous domains. *Expert syst. appl.*, **49**, 1–19.

Murphy, Kevin P. 2012. *Machine learning: A probabilistic perspective*. Adaptive Computation and Machine Learning. Cambridge, MA: MIT Press.

Murphy, Kevin Patrick, & Russell, Stuart. 2002. Dynamic bayesian networks: representation, inference and learning.

Mutoh, Yasuhiko. 2009. Simple design of the state observer for linear time-varying systems. *Pages 225–229 of:* Filipe, Joaquim, Andrade-Cetto, Juan, & Ferrier, Jean-Louis (eds), *Icinco-spsmc*. INSTICC Press.

Naik, Amol Subodh. 2010. *Subspace based data-driven designs of fault detection systems*. Ph.D. thesis, Uni Duisburg-Essen.

Nannapaneni, Saideep, Mahadevan, Sankaran, & Rachuri, Sudarsan. 2016. Performance evaluation of a manufacturing process under uncertainty using bayesian networks. *Journal of cleaner production*, **113**, 947–959.

Ng, K. Y., Tan, C. P., & Oetomo, D. 2012. Disturbance decoupled fault reconstruction using cascaded sliding mode observers. *Automatica*, **48**, 794–799.

Nguang, Sing Kiong, & Lin, Chih-Min. 1999. Robust filtering: a model matching approach. *International journal of systems science*, **30**(10), 1143–1151.

Nguyen-Trang, Thao, & Vovan, Tai. 2017. A new approach for determining the prior probabilities in the classification problem by bayesian method. *Adv. data analysis and classification*, **11**(3), 629–643.

Niculita, Octavian, Irving, Phil, & Jennions, Ian K. 2012. Use of cots functional analysis software as an ivhm design tool for detection and isolation of uav fuel system faults. *In: Proceedings of the prognostic and health management society conference*, vol. 3.

Niculita, Octavian, Jennions, Ian K, & Irving, Phil. 2013. Design for diagnostics and prognostics: A physical-functional approach. *Pages 1–15 of: Aerospace conference, 2013 ieee*. IEEE.

Nikoukhah, R . 1995. A new methodology for observer design and implementation. *Inria report*.

Oblak, Simon, Skrjanc, Igor, & Blazic, Saso. 2007. Fault detection for nonlinear systems with uncertain parameters based on the interval fuzzy model. *Eng. appl. of ai*, **20**(4), 503–510.

Ostrom, C.W. 2010. *Time series analysis: Regression techniques*. Thousand Oaks/ Newbury Park, CA: Sage Publications Inc.

Pampel, Fred C. 2000. *Logistic regression: A primer*. Sage Publications, Inc.

Parlangeli, Gianfranco, Pacella, D., & Corradini, Maria Letizia. 2007. Fault identification and accommodation for incipient and abrupt faults. *Pages 1003–1008 of: Cdc*. IEEE.

Patton, Ron J, Frank, Paul M, & Clarke, Robert N. 1989. *Fault diagnosis in dynamic systems: theory and application*. Prentice-Hall, Inc.

Peñarrocha, I., Sanchis, R., & Albertos, P. 2009. Observer design for a class of nonlinear discrete systems. *European journal of control*, **15**(2), 157–165.

Prasov, Alexis A, & Khalil, Hassan K. 2012. A nonlinear high-gain observer for systems with measurement noise in a feedback control framework. *Ieee transactions on automatic control*, **58**(3), 569–580.

Puig, V., Oca, S. M., & Blesa, J. 2012. Adaptive threshold generation in robust fault detection using interval models: time-domain and frequency-domain approaches. *International journal of adaptive control and signal processing*.

Qi, Haiyu, Ganesan, Sanka, & Pecht, Michael G. 2008. No-fault-found and intermittent failures in electronic products. *Microelectronics reliability*, **48**(5), 663–674.

Qian, Xiaoning, & Dougherty, Edward R. 2016. Bayesian regression with network prior: Optimal bayesian filtering perspective. *Ieee trans. signal processing*, **64**(23), 6243–6253.

Qning, X., Hau, Z., Feng, Y., Qiao, W. X., & Yong, Y. H . 2014. Effective model based fault detection scheme for rudder servo system. *Journal of cent. south university*, **21**, 4172–4183.

Rajmani, R. 1998. Observer for lipschitz nonlinear systems. *Ieee transaction on automatic control*, **43**, 397–401.

Raoufi, R., Marquez H. J., & Zinober, A. S. I. 2010. Sliding mode observers for uncertain lipschitz system with fault estimation synthesis. *International journal of robust and nonlinear control*, **20**, 1785–1801.

Reggia, J., Nau, D., & Wang, P. 1985. A formal model of diagnostic inference. *Information sciences*, 227–285.

Ríos, Héctor, Punta, Elisabetta, & Fridman, L. 2017. Fault detection and isolation for nonlinear non-affine uncertain systems via sliding-mode techniques. *International journal of control*, **90**(2), 218–230.

Roberts, Gareth O, & Rosenthal, Jeffrey S. 2009. Examples of adaptive mcmc. *Journal of computational and graphical statistics*, **18**(2), 349–367.

Rohmer, Jeremy. 2020. Uncertainties in conditional probability tables of discrete bayesian belief networks: A comprehensive review. *Engineering applications of artificial intelligence*, **88**, 103384.

Ross, S.M. 2003. *Introduction to probability models*. Vol. 8. Academic press New York.

Rotondo, Damiano, López-Estrada, Francisco-Ronay, Nejjari, Fatiha, Ponsart, Jean-Christophe, Theilliol, Didier, & Puig, Vicenç. 2016. Actuator multiplicative fault estimation in discrete-time lpv systems using switched observers. *Journal of the franklin institute*, **353**(13), 3176–3191.

Ruderman, Michael. 2013. Tracking control of motor drives using feedforward friction observer. *Ieee transactions on industrial electronics*, **61**(7), 3727–3735.

Sadler, D. Royce. 1975. *Numerical Methods for Nonlinear Regression*. University of Queensland Press.

Said, A., Djamel, B., & Boudjema, I. 2013. Structural analysis for fault detection and isolation using the matching rank algorithm for residual generation: application on an isdustrial water heating system. *Journal of control engineering and applied informatica*, **15**.

Salmerón, Antonio, Rumí, Rafael, Langseth, Helge, Nielsen, Thomas D, & Madsen, Anders L. 2018. A review of inference algorithms for hybrid bayesian networks. *Journal of artificial intelligence research*, **62**, 799–828.

Samadi, M. Foad, & Saif, Mehrdad. 2017. State-space modeling and observer design of li-ion batteries using takagi-sugeno fuzzy system. *Ieee trans. contr. sys. techn.*, **25**(1), 301–308.

Samantaray, SR. 2009. Decision tree-based fault zone identification and fault classification in flexible ac transmissions-based transmission line. *Iet generation, transmission & distribution*, **3**(5), 425–436.

Sang, Huiyan, & Huang, Jianhua Z. 2012. A full scale approximation of covariance functions for large spatial data sets. *Journal of the royal statistical society: Series b (statistical methodology)*, **74**(1), 111–132.

Schneider, C. 2015. *Using unsupervised machine learning for fault identification in virtual machines.* Ph.D. thesis, University of St Andrews, UK. British Library, EThOS.

Schneider, Henning, & Frank, Paul M. 1996. Observer-based supervision and fault detection in robots using nonlinear and fuzzy logic residual evaluation. *Ieee trans. contr. sys. techn.*, **4**(3), 274–282.

Sedighi, T, Foote, PD, & Khan, Samir. 2014. The performance of observer-based residuals for detecting intermittent faults: the limitations. *Procedia cirp*, **22**, 65–70.

Sedighi, Tabassom. 2019a. Using dynamic and hybrid bayesian network for policy decision making. *International journal of strategic engineering (ijose)*, **2**(2), 22–34.

Sedighi, Tabassom. 2019b. *Using Dynamic Bayesian Network (DBN) for Evaluation*. `https://uk.mathworks.com/matlabcentral/fileexchange/69698-using-dynamic-bayesian-network-dbn-for-evaluation`. [Online; accessed 9-January-2019].

Sedighi, Tabassom, & Varga, Liz. 2019. Cecan evaluation and policy practice note (eppn) for policy analysts and evaluators-a bayesian network for policy evaluation.

Sedighi, Tabassom, Phillips, P., & Foote, P.D. 2013. Model-based intermittent fault detection. *In: 2nd international through-life engineering services conference*.

Sedighi, Tabassom, Foote, Peter D, & Khan, Samir. 2015. Intermittent fault detection on an experimental aircraft fuel rig: Reduce the no fault found rate. *Pages 110–115 of: 2015 4th international conference on systems and control (icsc)*. IEEE.

Sharma, Abhishek B, Golubchik, Leana, & Govindan, Ramesh. 2010. Sensor faults: Detection methods and prevalence in real-world datasets. *Acm transactions on sensor networks (tosn)*, **6**(3), 1–39.

Sharma, R., & Aldeen, M. 2011. Fault and disturbance reconstruction in non-linear systems using a network of interconected sliding mode observers. *Iet control theory and applications*, **5**, 751–763.

Shen, Qiang, & Chouchoulas, Alexios. 2000. Knowledge-based fault detection in industrial plants supported by rough-fuzzy learning. *Ifac proceedings volumes*, **33**(11), 657–662.

Shen, Yirong, Seeger, Matthias, & Ng, Andrew Y. 2006. Fast gaussian process regression using kd-trees. *Pages 1225–1232 of: Advances in neural information processing systems*.

Skima, Haithem, Medjaher, Kamal, Varnier, Christophe, Dedu, Eugen, & Bourgeois, Julien. 2016. A hybrid prognostics approach for mems: From real measurements to remaining useful life estimation. *Microelectronics reliability*, **65**, 79–88.

Slezak, Dominik. 2009. Degrees of conditional (in)dependence: A framework for approximate bayesian networks and examples related to the rough set-based feature selection. *Information sciences*, **179**(3), 197 – 209.

Suresh, Yeresime, Kumar, Lov, & Rath, Santanu Ku. 2014. Statistical and machine learning methods for software fault prediction using ck metric suite: a comparative analysis. *Isrn software engineering*, **2014**.

Syed, Wakil Ahmad, Perinpanayagam, Suresh, Samie, Mohammad, & Jennions, Ian. 2016. A novel intermittent fault detection algorithm and health monitoring for electronic interconnections. *Ieee transactions on components, packaging and manufacturing technology*, **6**(3), 400–406.

Tao, Yuan, Shen, Dong, Wang, Youqing, & Ye, Yinzhong. 2015. Reliable control for uncertain nonlinear discrete-time systems subject to multiple intermittent faults in sensors and/or actuators. *Journal of the franklin institute*, **352**(11), 4721–4740.

Theilliol, Didier, Noura, Hassan, & Sauter, Dominique. 1998. Fault-tolerant control method for actuator and component faults. *Pages 604–609 of: Proceedings of the 37th ieee conference on decision and control (cat. no. 98ch36171)*, vol. 1. IEEE.

Tinga, Tiedo. 2010. Application of physical failure models to enable usage and load based maintenance. *Reliability engineering and system safety*, **95**(10), 1061–1075.

Ukil, Abhisek, & Zivanovic, Rastko. 2015. Abrupt change detection in power system fault analysis using adaptive whitening filter and wavelet transform. *Corr*, **abs/1503.05275**.

Vanhatalo, Jarno, Jylänki, Pasi, & Vehtari, Aki. 2009. Gaussian process regression with student-t likelihood. *Pages 1910–1918 of:* Bengio, Yoshua, Schuurmans, Dale, Lafferty, John D., Williams, Christopher K. I., & Culotta, Aron (eds), *Nips*. Curran Associates, Inc.

Vapnik, Vladimir Naumovich. 1998. *Statistical learning theory*. New York, NY, USA: Wiley.

Venkatasubramanian, V., Rengaswamy, R., & Kavuri, S.N. 2003. A review of process fault detection and diagnosis. part i: Quantitative model-based methods. *Computers and chemical engineering*, **27**(3), 293–311.

Vlasselaer, Jonas, & Meert, Wannes. 2012. Statistical relational learning for prognostics. *Pages 45–50 of: Proceedings of the 21st belgian-dutch conference on machine learning*.

Wan, Cen, & Freitas, Alex Alves. 2015. Two methods for constructing a gene ontology-based feature network for a bayesian network classifier and applications to datasets of aging-related genes. *Pages 27–36 of: Bcb*. ACM.

Wan, Yiming, Wang, Wei, & Ye, Hao. 2016. Integrated design of residual generation and evaluation for fault detection of networked control systems. *International journal of robust and nonlinear control*, **26**(3), 519–541.

Wang, Chunyi, & Neal, Radford M. 2012. Gaussian process regression with heteroscedastic or non-gaussian residuals. *Corr*, **abs/1212.6246**.

Wang, Dan, & Huang, Jie. 2001. A neural network-based approximation method for discrete-time nonlinear servomechanism problem. *Ieee trans. neural networks*, **12**(3), 591–597.

Wang, Dejun, & Song, Shiyao. 2014. Robust model-based sensor fault monitoring system for nonlinear systems in sensor networks. *Sensors*, **14**(10), 19138–19161.

Wang, Haiqin, & Druzdzel, Marek J. 2013. User interface tools for navigation in conditional probability tables and elicitation of probabilities in bayesian networks. *Corr*, **abs/1301.4430**.

Wang, J. M., Fleet, David J., & Hertzmann, Aaron. 2005. Gaussian process dynamical models. *Pages 1441–1448 of: Nips*.

Wang, Limin, & Zhao, Haoyu. 2015. Learning a flexible k-dependence bayesian classifier from the chain rule of joint probability distribution. *Entropy*, **17**(6), 3766–3786.

Wang, Shengwei, & Xiao, Fu. 2004. Detection and diagnosis of ahu sensor faults using principal component analysis method. *Energy conversion and management*, **45**(17), 2667–2686.

Wang, Shibin, Zhu, Zhongkui, He, Yingping, & Huang, Weiguo. 2010. Adaptive parameter identification based on morlet wavelet and application in gearbox fault feature detection. *Eurasip j. adv. sig. proc.*, **2010**.

Wang, Y. Q., Ye H., & Wang, G. Z. 2007. Fault detection of ncs based on eigendecomposition, adaptive evaluation and adaptive threshold. *International journal of control*, **80**, 1903—1911.

Wang, Yongqiang, Ye, Hao, Ding, Steven X, Wang, Guizeng, & Zhou, Donghua. 2009. Residual generation and evaluation of networked control systems subject to random packet dropout. *Automatica*, **45**(10), 2427–2434.

Wang, YQ, Ye, Hao, & Wang, GZ. 2007. Fault detection of ncs based on eigendecomposition, adaptive evaluation and adaptive threshold. *International journal of control*, **80**(12), 1903–1911.

Wende, TIAN, Minggang, HU, & Chuankun, LI. 2014. Fault prediction based on dynamic model and grey time series model in chemical processes. *Chinese journal of chemical engineering*, **22**(6), 643–650.

Widodo, Achmad, & Yang, Bo-Suk. 2007. Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical systems and signal processing*, **21**(6), 2560–2574.

Williams, Christopher KI, & Rasmussen, Carl Edward. 2006. *Gaussian processes for machine learning*. Vol. 2. MIT press Cambridge, MA.

Wilson, A.G., & Adams, R.P. 2013. Gaussian process covariance kernels for pattern discovery and extrapolation. *Corr*, **abs/1302.4245**.

Witczak, Marcin. 2003. *Identification and fault detection of non-linear dynamic systems*. Ph.D. thesis, Computer Science and Telecommunications, University of Zielona Gora.

Wünnenberg, Jürgen. 1990. *Observer based fault detection in dynamic systems: = beobachtergestützte fehlerdetektion in dynamischen systemen*. Ph.D. thesis, Uni Duisburg-Essen.

Woodberry, Owen, Nicholson, Ann E, Korb, Kevin B, & Pollino, Carmel. 2004. Parameterising bayesian networks. *Pages 1101–1107 of: Australasian joint conference on artificial intelligence*. Springer.

Wu, Lifeng, Yao, Beibei, Peng, Zhen, & Guan, Yong. 2017. An adaptive threshold algorithm for sensor fault based on the grey theory. *Advances in mechanical engineering*, **9**(2), 1687814017693193.

Xia, Yinglong, & Prasanna, Viktor K. 2008. Junction tree decomposition for parallel exact inference. *Pages 1–12 of: Ipdps*. IEEE.

Xiang, Y., Beddoes, M. P., & Poole, D. 1990 (May). Sequential updating of conditional probability in Bayesian networks by posterior probability. *Pages 21–27 of: Proc. 8th canadian artificial intelligence conf.*

Xuhui, B., Zhongsheng H. Fashan F., & Ziyi, F. 2012. Model free adaptive control with disturbance observer. *Journal of control engineerinf and applied informatic*, **14**.

Yamaguchi, Takuma, Inagaki, Shinkichi, & Suzuki, Tatsuya. 2012. Data based construction of bayesian network for fault diagnosis of event-driven systems. *Pages 508–514 of: Case*. IEEE.

Yan, Ke, Ji, Zhiwei, & Shen, Wen. 2017. Online fault detection methods for chillers combining extended kalman filter and recursive one-class svm. *Neurocomputing*, **228**, 205–212.

Yan, Rongyi, He, Xiao, & Zhou, Donghua. 2016. Detecting intermittent sensor faults for linear stochastic systems subject to unknown disturbance. *Journal of the franklin institute*, **353**(17), 4734–4753.

Yan, Xing-Gang, & Edwards, Christopher. 2008. Adaptive sliding-mode-observer-based fault reconstruction for nonlinear systems with parametric uncertainties. *Ieee transactions on industrial electronics*, **55**(11), 4029–4036.

Yang, Zhenyu, & Stoustrup, Jakob. 2000. Robust reconfigurable control for parametric and additive faults with fdi uncertainties. *Pages 4132–4137 of: Proceedings of the 39th ieee conference on decision and control (cat. no. 00ch37187)*, vol. 4. IEEE.

Ye, Hao, Wang, Guizeng, & Ding, Steven X. 2004. A new parity space approach for fault detection based on stationary wavelet transform. *Ieee trans. automat. contr.*, **49**(2), 281–287.

Young, Rosaline. 2018 (Nov.). *Lesson 8: Modeling physical systems with linear differential equations*. `https://slideplayer.com/slide/12325346/`.

Zhang, Junfeng, Christofides, Panagiotis D, He, Xiao, Wu, Zhe, Zhang, Zhihao, & Zhou, Donghua. 2018a. Event-triggered filtering and intermittent fault detection for time-varying systems with stochastic parameter uncertainty and sensor saturation. *International journal of robust and nonlinear control*, **28**(16), 4666–4680.

Zhang, Junfeng, Christofides, Panagiotis D, He, Xiao, Albalawi, Fahad, Zhao, Yinghong, & Zhou, Donghua. 2018b. Intermittent sensor fault detection for stochastic ltv systems with parameter uncertainty and limited resolution. *International journal of control*, 1–9.

Zhang, Qinghua, & Basseville, Michèle. 2014. Statistical detection and isolation of additive faults in linear time-varying systems. *Automatica*, **50**(10), 2527–2538.

Zhang, Wei, Su, Housheng, Zhu, Fanglai, & Azar, Ghassan M. 2015. Unknown input observer design for one-sided lipschitz nonlinear systems. *Nonlinear dynamics*, **79**(2), 1469–1479.

Zhang, X., Polycarpou, M. M., & Parisini, T. 2010. Fault diagnosis of a class of nonlinear uncertain systems with lipschitz nonlinearities using adaptive estimation. *Automatica*, **46**, 290–299.

Zhang, Yongquan, Cao, Feilong, & Xu, Zongben. 2011. Estimation of learning rate of least square algorithm via jackson operator. *Neurocomputing*, **74**(4), 516–521.

Zhang, Yu, Deng, Zhidong, Jiang, Hongshan, & Jia, Peifa. 2006. Dynamic bayesian network (dbn) with structure expectation maximization (sem) for modeling of gene network from time series gene expression data. *Pages 41–47 of:* Arabnia, Hamid R., & Valafar, Homayoun (eds), *Biocomp*. CSREA Press.

Zhao, Yan, Lam, James, & Gao, Huijun. 2009. Fault detection for fuzzy systems with intermittent measurements. *Ieee trans. fuzzy systems*, **17**(2), 398–410.

Zhao, Yinghong, He, Xiao, & Zhou, Donghua. 2018. Intermittent fault detection with t2 control chart. *Ifac-papersonline*, **51**(24), 1298–1304.

Zhou, Donghua, Zhao, Yinghong, Wang, Zidong, He, Xiao, & Gao, Ming. 2020. Review on diagnosis techniques for intermittent faults in dynamic systems. *Ieee transactions on industrial electronics*, **67**(3), 2337–2347.

Zhu, F., & Yang, J. 2013. Fault detection and isolation design for uncertain nonlinear systems based on full-order, reduced-order and high-order high-gain sliding-mode observers. *International journal of control*.

Zhu, Yaguang, & Jin, Bo. 2016. Analysis and modeling of a proportional directional valve with nonlinear solenoid. *Journal of the brazilian society of mechanical sciences and engineering*, **38**(2), 507–514.