Energy-based fault detection and isolation in an industrial steam turbine system

JH Smith

orcid.org/ 0000-0001-5184-4244

Dissertation accepted in fulfilment of the requirements for the degree *Master of Engineering in Electrical Engineering* at the North West University

Supervisor:	Prof G van Schoor
Co-supervisor:	Prof KR Uren
Co-supervisor:	Prof M van Eldik

Graduation:	June 2021
Student number:	24162159

Declaration

I, Jan-Hendrik Smith, hereby declare that the dissertation entitled "**Energy-based fault detection and isolation in an industrial steam turbine system.** " is my own original work. Whenever contributions of others are involved, every effort was made to indicate this clearly, with due reference to the literature. No part of this work has previously been submitted to any other university or institution for examination.

Jan-Hendrik Smith
Student number: 24162159
Signed on the 16th day of March, 2021 at Potchefstroom.

Abstract

System monitoring, especially with the Industry 4.0 push to automate the industry, is becoming more crucial for the successful operation of industrial systems. The effective and early detection and isolation of faults is an important part of both system monitoring and control. This can be a laborious process especially if the system consists of more than one domain of operation i.e. electrochemical or electromechanical etc. A method that can overcome this multi-domain complexity is the energy graph-based visualisation (EGBV) method that can be used as a hybrid fault detection and isolation (FDI) technique. This study aims to evaluate this energy-based approach to fault detection and fault isolation applied to a steam turbine system (STS).

A static thermodynamic model of the STS situated in Jeanschwalde Germany has been developed in the software package engineering equation solver (EES). The model is based on mass and energy balance equations pertaining to individual components of the system. The modelling knowledge has then been applied to simulate a fault type (FT) using an existing validated model of the STS developed in the software package Ebsilon by engineering staff working at the plant. Four faults have been considered, namely: Solid particle erosion (SPE), leakage of the turbine's overflow valve, overall ageing or wear of the turbines and pump cavitation. Physical data in the form of temperature, pressure and mass flow rates from the STS have also been obtained for a number of components in the system. The simulations and physical data have been combined, using statistical methods, in an attempt to create time series energy based data for the purposes of FDI.

An energy characterisation of the STS has been done and an attributed graph of the STS is composed pertaining to modern graph theory techniques. The attributed graph has been used to construct node signature matrices containing the energy data from the constructed time series. These node signature matrices are utilised in three analytic approaches to FDI. Each of the three approaches uses different aspects of the information encapsulated within the node signature matrices. Approach 1 and 2 are based on graph matching to compose cost matrices with the Heterogeneous Euclidean-Overlap Metric (HEOM) metric. Approach 1 uses a single distance parameter that is indicative of a fault being present in the system or not. Approach 2 utilises eigenvalues of the cost matrices as the mathematical analysis technique for FDI. Approach 3 uses a residual approach to determine whether a fault is present and to isolate one fault from another.

The FDI results obtained from the 3 approaches show that approach 2 had the best performance. The detection accuracy of approach 2 was 100% and an isolation accuracy ranging from 19%, for some variations of FT2, to 100% for for FT4. Approach 1 had the lowest accuracy in terms of fault isolation, attaining isolation accuracy of less than 20.4% for all FT's except FT1 which was isolated with a 100% accuracy. Approach 3 had the worst detection accuracy with moderate isolation, attaining only 72% accuracy in terms of fault detection and an isolation accuracy ranging between 5.2% and 70%.

Future research on this topic can entail the development of a general approach to energy-based FDI. The establishment of guidelines for composing and attributed graph of a system will be crucial for such future research. The application of alternative metrics for graph matching and the analysis of these matched graphs using machine learning techniques can also be considered. A comparative study of the effectiveness of these FDI methods compared to traditional methods like Principle Component Analysis (PCA) is also required.

Keywords: **Fault detection, Fault isolation, Energy visualization, Graph theory, Steam turbine system**

Acknowledgements

"For from Him and through Him and to Him are all things. To Him be the glory forever"

-Rom 11:36

I would like to thank the following persons and institutions, in no specific order, for the support, contributions and effort they made towards the completion of this dissertation:

- The North-West University, Potchefstroom Campus for the financial support and facilities that provided the opportunity to accomplish this research and allowed met to obtain this education.
- To the Institute for process technology, process automation and measurement technology in Zittau/Görlitz Germany for their funding, hospitality and efforts in providing the opportunity to acquire data form a physical system.
- To my study leaders, Prof. George van Schoor, Prof. Kenny Uren, and Prof. Martin van Eldik for their patience, valuable guidance and mentorship towards finishing this research. Without their support it would have been impossible to complete.
- A special thanks to Prof. Frank Worlitz for his effort in organising my trip to Germany. Thanks to him I will never forget the experience.
- To my lovely wife, Isabelle for her proofreading, grammar evaluation and moral support. Without her coffee my joy in finishing this dissertation would be dull.
- To my parents for their support during the years. Their confidence and prayers kept me going through the years since this academic journey started.
- To my fellow researchers colleagues from the McTronX research group for their feedback and interaction on my research.
- To all the unmentioned friends and family who supported me in numerous ways over the years.

Contents

List of Figures

List of Tables

“It is the glory of God to conceal a matter and the glory of kings to search it out.”

– King Solomon

Proverbs 25:2

CHAPTER 1

Introduction

This chapter gives a brief background of the research and basic context of the physical system. The research problem statement, objectives as well as the methodology are also described. Lastly the chapter gives an outline of the dissertation as well as the applicability and necessity of the study.

1.1 Background

Many electricity generating steam turbine powered generators in the 20th century are operating in service far beyond the designed lifetime of the machine [7]. The single largest cost of keeping a steam turbine running is the cost of dismantling parts of the system for maintenance. The monitoring of the STS, as a whole, on a regular or continuous basis allows for these costs to be lowered. Karlsson [4] states that when using fault diagnosis and identification methods the probability of having a breakdown that will result in a large dismantling of the system, can be greatly reduced if the fault diagnosis is acted on in time. According to [8], the dismantling of a STS in order to fix faults may also introduce larger problems since it increases the potential of, for example, vibrations and leakages. Due to the need for fault detection and isolation in modern systems the incentive for research in this regard is high. The industry push for automation, specifically with Industry 4.0, also increases the need for condition monitoring [9]. The potential loss in profits or equipment when faults are not detected and fixed in time can have unwanted and irreversible effects. Due to this fact an effective method of analysing a system and detecting faults is of utmost importance. It would be of immense value if a system can be monitored to such an extent that faults are detected before a large breakdown occurs. This could allow for predictive maintenance and can result in large cost reductions. The book by M. Mansouri [10] discusses numerous methods for fault detection. The application of data driven methods, among others, PCA and partial least square (PLS) is discussed. As a final chapter the book describes methods of state estimation for state space models and elaborate on the state estimation when using fault detection-based approaches. The book also describes the determination of upper and lower control limits for FDI applications and gives some examples of the application thereof in actual systems. The largest shortfalls of these methods seem to be either the lack of a holistic view of the system or the handicap imposed by the large amount of data required for the method to be effective. Both data driven and model-based methods attempt to address these shortfalls but seems ineffective if used as a standalone. In the article by Y. Ding et al. [11] the complexity and challenges with FDI in steam turbines using an artificial neural network (ANN) is discussed. A similar article by C. Karlsson [4] also utilizes an ANN for fault detection and interactive isolation in a STS. From both articles the common faults in these systems and the immense importance of its early detection, for supporting maintenance decisions,

is made clear. The shortfalls seen with data and modelled based methods, are typically addressed by hybrid systems.

Mainly two fault detection philosophies exist with some contention regarding the meaning of the terms. In order to clarify this, the terms are defined by [12] as follows:
The first being fault detection and diagnosis (FDD), referring to determination of a fault state in a system and obtaining information on what may be the cause, severity and location of the fault on a component level. The second is FDI, which refers to the determination of faults in a system or component, detecting in such the origin of the fault. The term isolating refers to the action of identifying unique faults. A fault can therefore be detected in a system but only isolated if the fault is unique to that specific case. By isolating faults, it allows for specific distinction between the different faults that can occur and results in a clear prognosis of the system's fault state. Based on the method of analysis used, fault isolation may also indicate the location of the fault in the system, however, the ability to identify the fault's location is normally referred to as fault identification rather than fault isolation. For the purpose of this research only FDI will be considered.

Besides the difficulties when detecting faults in general, the multi domain nature of most modern systems further increase the complexity of analysis. This implies that using energy and exergy as the main variables of analysis can be beneficial. This analysis approach is advantageous because the multi domain nature of these industrial plants does not impact the analysis as one moves from one domain (i.e. Mechanical, Electrical, Petrochemical etc.) to another. This can allow for a more effective and efficient approach to FDI. Further, since various aspects in terms of the information on the system's health is accumulated within the energy attributes of the system, the complexity of analysis is reduced. Lastly this approach may lead to effective energy and exergy-based methods of system control. In an article by Du Rand published in 2012 an approach for general fault detection (FD) in a 4^{th} generation nuclear high temperature gas-cooled reactor using enthalpy-entropy graphs were proposed [13]. The use of energy and energy attributes in industrial systems for thermal plant analysis is common [14]. The application of energy and specifically exergy for FD is however a relatively novel approach. The article publication by Du Rand was at the beginning of a branch of FD research at the North-West University (NWU) that later resulted in larger research focus, where the applicability of energy and energy derivatives as well as graph theory approaches for general FD are being investigated.

The resulting research had the aim of finding an effective hybrid method for general FD, specifically using energy and energy attributes as a global variable that will allow for simplification of the analysis, detection, isolation and diagnosis of faults in multi-domain systems such as electrochemical and electromechanical plants.

In an article [15], published in 2018, a vision for energy-based graph visualisation with the purpose of condition monitoring was proposed. The article described a graph approach using a simple Rankine cycle as an example for composing an energy based attributed graph. The term energy graph visualisation refers to the utilizing of energy attributes of a system or component in a mathematical and graphic sense. This means applying graph theory and the related analysis methods thereof to investigate and present the energy attributes in a useful way for FDI. This approach is also presented in [3]. Further a basic FDI scheme was set forth and illustrated with 3 case studies, 1), Heated two-tank system, 2), Gas to liquids process and 3), Closed loop Brayton cycle. All of these studies were done either on benchmark, small scale or modelled systems. There exists therefore a requirement for more research on the applicability of the method using data from an actual large-scale plant.

The opportunity presented itself in a town called Jaenschwalde, Germany at a lignite plant that is used to generate electrical power. The plant is a STS that consists of 6 sections or units each delivering around 500 MW of power. Due to the environmental damage that these systems cause, and the push from Germany's government to reduce CO_2 emissions, four of the six units of the plant were shut down and plans exist to shut the whole plant within a few years. The two remaining sections are being monitored for research application to this date. The research on magnetic bearings by the Institute for Process Technology, Process Automation and Measurement Technology (IPM) of the Hochscule is based on a part of this STS. Hence an ongoing relationship with the Hochschule Zittau/Gurlitz University of Applied Sciences and the plant is sustained. This arrangement enabled the obtainment of data and the use of a comprehensive software model for this research.

1.2 Problem statement

The applicability of the energy graph based visualisation (EGBV) technique for FDI purposes, as proposed by [15], should be evaluated when applied to a STS.

This requires:
1) Valid operational thermodynamic data from the STS that can be used for energy and exergy calculations.
2) Data for normal and fault operational conditions of the STS in order to be able to apply the proposed FDI technique.
3) Energy characterisation of the system of such a nature that mathematical analysis is simplistic.
4) The FDI method's performance should be tested in terms of sensitivity to system changes, robustness of the scheme and accuracy of the detection and isolation of faults.

The delimitations of the study are:
1) Only FDI will be evaluated as defined and not FDD.
2) The nature of the system under investigation in this research, is such that the system is considered on a system level, using the turbine only as the classifier for the system type, and does not imply a singular focus on faults within the turbine as a component. In this regard the system size is substantial and would not be practical for a "magister inżynier" (Ming) study to conduct fault detection on a subcomponent level. The FDI is thus restricted to a component and system level. This implies that the detection of faults is only effective to the extent that faults can be attributed to a specific component like a pump or turbine. It is, however, important to note that the turbine comprises of 3 units (high pressure turbine (HPT), medium pressure turbine (MPT) and low pressure turbine (LPT)), with various stages, which will be viewed as separate components since each of the individual stages (4 for the MPT and 3 for the LPT) can be modelled as individual turbines.
3) The nature of the EGBV only allows for systems with representative energy attributes to be monitored and thus the study will not conclude on systems other than STS for which energy attributes are available.

1.3 Objectives and methodology

1.3.1 Objectives

1) STS modelling

The FDI method can only be applied to a system from which comprehensive data can be obtained. This data can be from either or both a model or the physical system. For the study, both model and real system data are used. Thus, a representative thermodynamic model of the system must be developed. This representative model will be used for fault simulation purposes. Initially, a static model in the software EES was developed, however, since access to a comprehensive validated static model in the software Ebsilon was made available at a later stage in the research, and the focus of the study is on FDI, this EES model is only used to describe the fundamental working of the system. The aim is to obtain simulation results that describe the energy and exergy in the system. Verification of the EES model is also required to ensure the correctness of the fundamental mathematics used in the model and to ensure that the simulations of faults and operational states are coherent with the physical system. The definitions of verification and validation are as defined by the standard IEEE-STD-610.

The end result must be a detailed analysis of the complete system, referring to each of the components that are included in the graph, with data in the form of temperature, pressure, mass flow and work from which the exergy and energy values throughout the system can be obtained. This will be described in the next section.

2) Energy characterisation

The graphical visualisation of energy necessitates research regarding the specifics of energy characterisation. This refers to three objectives.
1) Establish the energy attributes that are useful for the specific system under investigation and determine the applicable mathematical methods to calculate it.
2) Determine the graph node and link connections and establish which components to include as nodes or not.
3) Compose a graph and an attributed graph based on objective 1 and 2, such that the graph complexity is minimal.

3) Experimental design.

To facilitate the comprehensive investigation of FDI the composition of representative data sets from the available practical data and model simulations needs to be carefully considered. With this in mind, an experimental design is developed to specify how the practical data and the simulation data are used to arrive at the appropriate data sets that will be used in further analysis. The data must be for specifically identified faults to enable a method for validation of the FDI scheme. Also, in order to evaluate robustness and sensitivity of the FDI method, the data needs to consist of the actual physical system fluctuations and noise. Data must therefore be obtained from both the model and the physical system.
The required end results are the following 4 datasets:
1) Normal operational state simulation data.
2) Fault state simulation data (for each fault under investigation).
3) Normal operational data from the physical system.
4) Fault data from the physical system if available.

Since the data obtained from sets (3) and (4) will have noise and fluctuations, specific statistical analysis as well as curve flattening and noise filtering methods are used to process the data.

4) Energy-based fault detection and isolation

Finally, the study will be concluded through testing the FDI method. Thus, each of the mentioned aspects need to come together using a logical and systematic approach. This is done through evaluation of the FDI by validating and verifying the FDI method and its performance. Both the robustness and sensitivity of the FDI needs to be taken into consideration. The following faults are considered.

1) Solid particle erosion on different stages of the turbine.
2) Leakage of the overflow valve of different stages of the turbine.
3) Ageing and wear of the turbine overall.
4) Cavitation on the feeding pump.

1.3.2 Methodology

1) STS Modelling

The static model will be developed using the modelling approach as referred to in [16]. The model will be based on a simplified system using the software package EES. This model utilizes mass and energy balance equations to solve the thermodynamic parameters expected at various points in the system. The software allows for simultaneous solving of numerous equations. This enables one to define equations in terms of variables, like P for pressure or T for temperature, and solve it for specific parameter inputs throughout the system. Each component in the system is viewed as a controlled volume for which the mass and energy balances should hold. Connecting each of these components to one another, the system can be recomposed and mathematically modelled. The main use for this model is to grasp the inherent workings and modelling of the plant.
During a visit of 3 months to Germany, the validation of this model will be completed. This will be accomplished by comparing the simulation results to the physical system data and results obtained with an existing and more detailed dynamic model. The dynamic model was developed by a technical team in Germany using the software package "Ebsilon". It was created and validated to be used with the plant design in Jänschwalde. The verification of the static model will also be done by evaluating the model's ability to perform normal and faulty simulations. The dynamic model will also be used to simulate fault conditions. For the aforementioned modelling, changes will be made to the model in relation to each fault simulated. These changes include adjusting the model inputs or marginal values and adding specific fault conditions to the model set-up. Four main faults will be simulated as specified in the previous section 1.3.1 at 3 to 4 different intensities (e.g. for a leak, the intensities may be a high, medium and low mass flow rate).

2)Energy characterisation

The energy characterisation forms the connection between the system and the FDI scheme of the research. This is done by composing a node and link (also known as vertices and edges) graph and applying both the practical and modelled energy data to this graph. The graph is then referred to as an energy-based attributed graph as suggested by [3, 15] and [17]. The methodology includes determining which components to include as nodes or not. Also, the nodes which should be connected with links to ensure a simple graph that allow for fast and effective analysis, whilst still representing the structural composition of the system, are determined. This is done by analysing

the system according to five criteria, and then composing a decision matrix from which conclusions are made.
1) Nodes must be chosen for components for which faults will be detected.
2) Choose nodes to represent components with relatively high numbers of links connected to other components. For example see option 2 in Figure 1.1.
3) Ensure nodes represent components having large effects on the system's energy output.
4) The nodes must be representing components with high energy attributes relative to other components within the system. A typical measure of what constitutes large or small energy attributes will be to classify components with energy attributes above the average of the nominal energy attributes of all components as large energy attributes and include nodes for them. For example, see option 2 in Figure 1.2.
5) Use node placement in such a way that it minimises graph complexity.

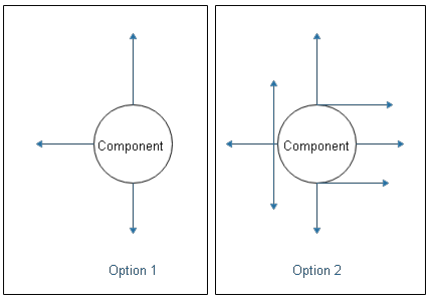


Figure 1.1: Example of component with high number of links

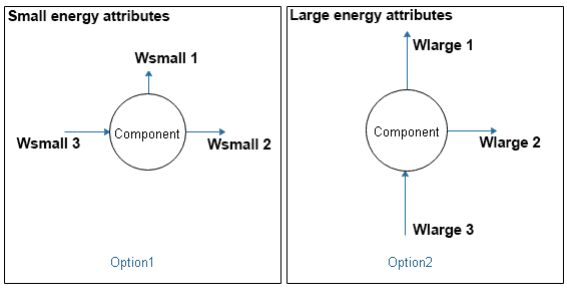


Figure 1.2: Example of component with large energy attributes

3) Experimental design.

In general, controlled fault operational data from a physical system can be obtained by using either existing fault data from the plant, or by emulating faults through knowingly inducing specific faults into the system. A third method would be to run a simulation of the faults, using a validated software model that is representative of the system.

With regards to the research, the major concern regarding the faults is that the physical system in Germany does not have fault data in the form of thermodynamic results. It is also not possible to induce faults directly into the system due to the fact that the plant is constantly operating, delivering power to the energy grid. Therefore, to obtain the fault state data simulations must be utilized. The only data obtained from the practical system, are under normal operational conditions while the load of the system would change. This results in a set of practical data that contains the normal noise of the system as well as the fluctuations in the data seen due to the control of the plant. A flow diagram of this approach is shown in Figure 1.3. This data is obtained through measuring temperature, pressure and mass-flow at various points in the system.

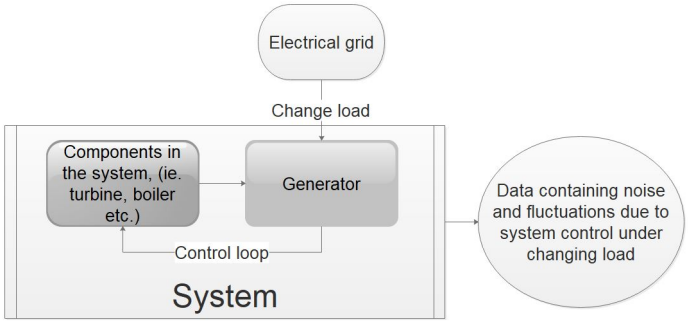


Figure 1.3: Flow diagram visualizing how the practical data are obtained

The simulation results give specific data that shows both the designed normal state results, as well as the results under fault conditions, at various points in the system. A flow diagram of this process is visible in Figure 1.4.

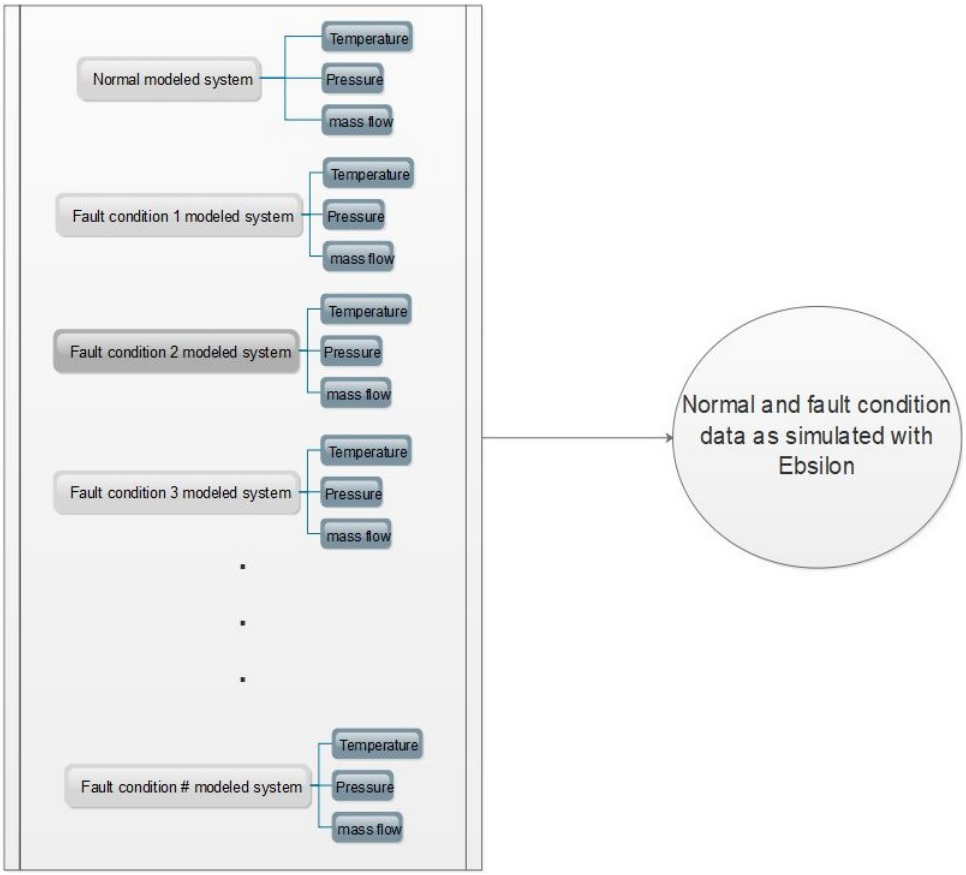


Figure 1.4: Flow diagram visualizing how the simulation data are obtained

Using a Simple Moving Average (SMA) filter the noise from the practical data are extracted. The average of the noise and a standard deviation above and below the mean can then be calculated. As an example, these results are then combined with the modelled results by adding the average noise to the simulation data and a number of fault simulation data sets with noise are obtained. The final data sets obtained will then consist of various parameters (temperature, pressure etc.) under specific conditions (normal or a fault condition), with as well as without noise.

4) Energy-based fault detection and isolation

For the research to be concluded the FDI application of the EGBV technique must be evaluated. The data of the system states, as mentioned before, are used as standalone data and in combination with the practical data. The FDI is done by composing a node signature matrix using the attributed energy-based graph. This matrix then consists of data for a specific fault state obtained from the model or practical system (the unhealthy state). This can be compared to a reference signature matrix (the healthy state) using a matching metric's like the HEOM or one of the metrics described in [18]. A resulting cost matrix is obtained. Repeating this for each fault, results in a set of cost matrices that can be analysed using eigenvalue decomposition. These results can then be qualitatively interpreted. In summary, by using graph matching techniques to compare the unhealthy signature to the healthy signature the faults are detected.

In order to evaluate the robustness and sensitivity of the FDI philosophy, the dataset with noise that occurs in the real system is used. By introducing a limit function in combination with some threshold value the sensitivity of the method is tested. Lastly the isolation of the fault signatures is done by simply subtracting the signature vectors from each other to obtain an isolability matrix. Since the faults are simulated, each fault is known and controlled for the implementation the FDI. To validate the FDI it can therefore simply be noted whether the resulting fault signatures detected and isolated are valid and unique or not. A quantifiable method for doing this would be to compile a confusion matrix.

1.4 Dissertation outline

1.4.1 Chapter 2: Literature

In this chapter a discussion of the relevant literature is given. It contains both a literature survey and detailed literature. The chapter focuses on elaborating specifically in regards to the four main aspects listed in the objectives. An explanation of the literature regarding thermodynamic modelling is given. Also, the various methods regarding FDI with an energy attributed approach are explained. Different methods of data analysis for the purposes of FDI are discussed and lastly the energy characterisation and graph theory aspects applicable to the study are elaborated on. The chapter motivates the relevance of the literature and gives a simple overview of each topic.

1.4.2 Chapter 3: STS Model

In chapter three the approach and step by step methods used to develop the thermodynamic model in EES are explained. This chapter serves as a description of the technical aspects regarding the model. A discussion of the validation process is also given as well as some elaboration regarding the deviations seen between the results of the model and the real system. Finally, a conclusion is made as well as a motivation on the use of the Ebsilon model that forms part of the validation process of the EES model.

1.4.3 Chapter 4: Energy characterisation

This chapter discusses the energy characterisation of the system. The methodology followed to determine which components to include as notes as well as which links to include between the nodes is explained. The physical analysis of the system to determine how the energy characterisation would be effective for the study outcomes is explained. Lastly a discussion on the effects of inclusion or neglecting of certain components as nodes are stated. This includes an explanation of

the approach followed if some of the components included in the energy characterisation do not allow for efficient data allocation.

1.4.4 Chapter 5: Experimental design and data collection

Chapter five gives an in-detail discussion on the methodology followed in Ebsilon to emulate faults. The obtaining of data from the physical system is also discussed. Finally, the analysis of the data from both the simulation as well as the physical system is explained. This includes an in-detail explanation of how the practical data and simulation data are combined in order to be used in the FDI philosophy. A short discussion on the different data analysis methods for thermodynamic data are discussed along with a defence of the methods used in the study.

1.4.5 Chapter 6: Energy-based fault detection and isolation

The second last chapter serves as an attempt to gather all aspects of the study with the application of the FDI philosophy as described by [3], [17] and [15]. The chapter gives an in-detail description of the mathematical approach as well as alternative methods in terms of the mathematical analysis. A discussion on the robustness and sensitivity of the FDI philosophy is also elaborated on. Lastly the chapter concludes by listing the fault conditions that were effectively detected and isolated or failed to do so.

1.4.6 Chapter 7: Conclusion

The last chapter concludes the dissertation with a quick summary of the research as well as some suggestions regarding possible topics of further research. Lastly some possible improvements that may be of good practice to study are given and the possible shortcomings of the study are listed and discussed.

CHAPTER 2

Literature

This chapter is a survey and review of all the literature that were studied. The chapter motivates the relevance of the literature and gives a simple overview of each topic as noted in the objectives. This includes a broad overview of the documentation regarding FDI and the use of energy attributes and the relative analysis methods. Previous research done on energy graph-based visualisation and exergy analysis is given. Lastly the literature on thermodynamic modelling and graph theory is noted.

2.1 Introduction

The focus of this chapter is to elaborate on the relevant literature regarding four main points. These points include a discussion of the literature on Rankine cycles and the thermodynamic modelling thereof. The literature and concepts regarding energy characterisation of a system are given and the relevant aspects of graph theory used as a tool for visualisation and data analysis is discussed. Lastly the literature regarding methods and philosophies for fault detection, fault detection and isolation and fault detection and diagnosis are given.

The chapter also contains an in-depth discussion of the energy based visualisation method for FDI and the modelling approach followed for development of thermodynamic models. The in-depth discussion includes some background and definitions on specific concepts in order to clarify the semantics for the rest of the dissertation.
The chapter ends with a critical review of the relevant literature as a motivation for the focus of the study.

2.2 Literature survey

The information regarding the EGBV for FDI purposes is based on the work done by G. Van Schoor, K. Uren, H. Marais and S. Greyling. [3, 15, 17, 19]. The literature builds on methods of FDI and concepts such as fault tolerance, fault diagnosis and performance assessment, modelled based vs. data driven approaches and sensitivity and robustness measures as mentioned in [20] as well as exergy analysis as defined from a fundamental approach described in [21], and lastly graph theory and graph matching methods as described in [18] and [22].

The book [20] is a thorough source on FDI methods and the standard analysis approaches that are

used such as the approach by the artificial intelligence community known as the DX approach. It provides a fundamental description of FD philosophies in general as well as the mathematics and techniques involved. The book introduces modern advanced FDI techniques and also describes the model, machine learning and hybrid-based methods that exist or are being researched. The definitions given on FDI topics by this source are used as the reference for the definitions used in the dissertation.

One motivation to pursue research on FDI for steam turbine systems is given by [4]. The motivation is based on the cost and time impact that these systems have when a breakdown occurs. The article gives a comprehensive explanation on FDI in a STS using an ANN. It discusses 7 typical fault types common in STS and gives a clear overview of the difficulties that arise with detecting these faults. Another motivation for FDI in STS is given by [8]. In the article, the difficulties and shortcomings of methods such as fuzzy logic and support vector machine (SVM) techniques are mentioned along with the cost implications that faults can have in these systems. The article focuses on a FD approach using flow graphs and a Naive Bayesian classifier. Lastly [7] motivates this research on STS. The article states that condition monitoring techniques on performance analysis (typically energy or exergy-based analysis) are less known than conventional techniques. It also claims that, in some cases, performance analysis is the only way of detecting and monitoring some modes of the system and is, therefore, necessary to investigate.

The book on modern advances in STS by T. Tanuma [23] describes the principles of steam turbine systems in detail, referring especially to the fundamental working of such systems as well as the operational procedures available to manage these systems, when used for electrical power generation applications. A detailed explanation on the Rankine cycle is given. This is further described in terms of the first principle mathematical modelling in and [24].

The first principle methods used for thermodynamic modelling with special focus on energy and exergy is described in [21] and [14]. The definitions and descriptions used in this study and the rest of this dissertation, when referring to energy, exergy and mass flow balance, are defined and described in these books as well as in [25]. The numerical analysis of thermodynamic components (compressor, boiler, turbine etc.) is described in detail in [26].

Robin J. Wilson [27] gives an in-depth description of graph theory in his book "Introduction to graph theory". All definitions and theorems related to graphs in this dissertation are obtained from this source. The application and description of graphs is also summarised in the notes by C. Griffin [28]. A detailed discussion of the visualizing abilities and data mining applications of graphs is given by [29]. This source also describes different algorithms for graph analysis and matching. Different graph matching metrics are explained and tested with a thorough comparison and analysis in [30]. A specific focus on improving the HEOM function is given, describing specific variations on the HEOM and their effects.

The discussion that follows, focuses on explaining the mentioned literature and concepts discussed, in more detail. More sources are mentioned in order to give a clear and holistic description of the relevant topics.

2.3 Fault detection and isolation

With the push to automatize the industry with the Industry 4.0 revolution, the need for condition monitoring is rapidly increasing. The article [9] provides a roadmap towards fully automated

factories. A step in this roadmap consists of data analysis especially emphasizing the monitoring of equipment effectiveness. This is done through diagnosis of a system resulting in a root-cause problem solving method. These diagnostics forms a part of common FD.
For the purposes of this dissertation the focus is on FDI, this however emphasizes the need to differentiate between FD, FDI and FDD. According to the IFAC technical committee SAFE-PROCESS as referenced in [12,31] the definition of the terms, detection, isolation, identification and diagnosis is as follows:

- Fault detection: Determination of faults present in a system and time of detection.
- Fault isolation: Determination of the kind, location and time of detection of a fault by evaluating its symptoms.
- Fault identification: Determination of the size and time-variant behaviour of a fault.
- Fault diagnosis: Determination of the kind, size, location and time of detection of a fault by evaluating its symptoms.

Therefore, FDI refers to the process of determining if a fault is present in the system at a certain time and also indicating the type and location of the fault detected, thereby distinguishing between the fault types and consistently isolating one fault from another.

As stated by [20, 32] and [10], the methods of FD can be grouped into categories of either model-based or data driven philosophies. The common difficulties with the modelled approach are that the FD is only as effective as the model's dynamic modelling capabilities. Developing such models are hard and complex especially in a dynamic multi domain system. Further, since the dynamic changes of a system results in certain changing parameters the model's accuracy can decay over time. This results in less effective modelled based FDI.

With data based FDI the complexity of implementation is less but the method is handicapped by the data it requires. Since the method relies on both pre-existing normal and faulty data in order to effectively execute FD, the method can only detect faults that previously occurred [20, 33]. Typical examples of this is ANN's, fuzzy logic (FL), SVM and combinations thereof. In [8] these methods are mentioned as background for a novel approach using flow graphs and a naive Bayesian classifier. A similar approach is followed in [4] using a Bayesian network (BN) in combination with an ANN. Another common method is called PCA. This technique is used in [34] as a method for FD in gas turbines. The advantage of PCA is that it is simple to implement and can be used as a dimensional reductive method since it forms part of the group of multidimensional descriptive methods called factorial methods. In various applications a combination of these techniques with graphs or some model-based approach are utilized. These approaches can be classified as hybrid methods. All of these modelled or data driven methods are generally used in combination with or singularly in STS for FDI. The general challenges of these methods in STS's is the complexity of the systems and the underlying requirements of the FDI methods [24].

Hybrid methods are a common compensation for the shortfalls of model and data driven methods. Unfortunately, many of these hybrid methods only combine some techniques whilst excluding the physical system information [19]. A hybrid method (the EGBV) that includes the physical structural information of the system is suggested by [15]. The article proposes a vision for condition monitoring using an energy approach. The multi-domain nature of modern industrial systems creates a unique opportunity for the use of energy and exergy analysis as a complexity reductive data abstraction method. This is due to the applicability of energy across domains. The universality of energy allows for additive treatment of all process inputs and outputs [19]. Three case studies are

given, proving the applicability of the method in large industrial system. The article by H. Marais et al. [19] describes the merits of specifically using exergy-based fault detection, building on the proposed EGBV method. Due to the first law of thermodynamics the conversion of energy in a real system can drive or detract a process and thus exergy as a measure of this is useful [21, 35]. H. Marais et al. states that the merits of hybrid exergy based FDI is observed in the reductive modelling effects since the structural information of the system acts as a form of model validation. Simultaneously the model complexity reduction of the method is beneficial.

With the automation drive instigated by Industry 4.0 in mind, the mentioned benefits of energy and especially exergy-based hybrid FDI methods are considered. This consolidates the FDI method used for the research. To validate the FDI method a quantitative analysis can be done. The performance metrics used for this analysis are categorised into two groups, defined by [36]. These categories are temporal and static metrics. The temporal metric quantifies how the FDI method respond to a time varying fault signal. The static metric quantifies the FDI performance with regards to a set of fault signals regardless of the time. Static performance metrics are therefore a time-independent performance metric. The static metric as described by [37] is used specifically for the verification of the FDI method in this research.

Since hybrid approaches combine model and data driven techniques, both a model and data are requirements for the research. The development and implementation of a representative model is therefore needed. This necessitates knowledge of the thermodynamics and mathematical modelling of the STS. Secondly comprehensive data acquiring techniques are necessary and will also be investigated. Lastly the EGBV presented by [15], utilizing graph theory, as part of the energy characterisation of the system under investigation, to accomplish FDI. In the next section the literature with regards to energy characterisation is elaborated on.

2.4 Energy characterisation

The concept of energy characterisation with regards to the EGBV implementation for FDI refers to three aspects. These are extrapolated from the articles [15], [3], [19], [31], [38], [39] and [6]. The process is shown in the flow diagram in Figure 2.1 and consists of three parts. The first is determining and calculating energy attributes (exergy, entropy, enthalpy, heat-flow) that will be used for the FDI analysis. Next analysing the system structure to determine the components of interest and find where in the system the energy attributes must be calculated or measured, thus reducing the size of the graph that will be composed. Finally composing a graph that represents the system structure and energy attributes.

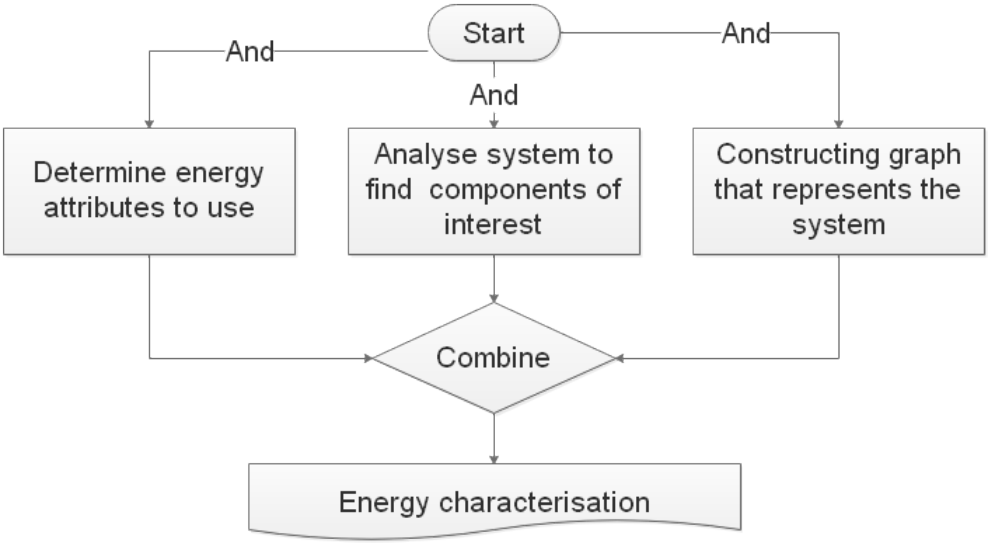


Figure 2.1: Flow diagram visualising the energy characterising process

The merits of exergy as one of the energy attributes is clear especially when considering the large-scale energy conversion that happens within a STS. Exergy can be loosely defined as a measure of the usable energy in a system or the universal measure of energy quality [14], and therefore it would be of value to use exergy as one of the energy attributes. In the absence of nuclear effect, magnetism, electricity and surface tension, exergy can be grouped into four types and can be mathematically written as shown in (2.1).

$$\dot{b} = \dot{b_k} + \dot{b_p} + \dot{b_{ph}} + \dot{b_{ch}}, \tag{2.1}$$

with: $\dot{b_k}$ = kinetic exergy, $\dot{b_p}$ = potential exergy, $\dot{b_{ph}}$ = physical exergy and $\dot{b_{ch}}$ = chemical exergy.

When the STS is analysed in terms of thermodynamic parameters, (2.1) reduces to only contain physical exergy. Fratzscher [14] defines physical exergy as "the maximum amount of work obtained when a working fluid stream is brought from its initial state to the environmental (or dead) state defined by P_0 and T_0 (environmental pressure and temperature) through physical processes only involving thermal interaction with the environment". This is mathematically expressed through (2.2).

$$\dot{b_{ph}} = \dot{m}[h_{initial} - h_0 - T_0(s_{initial} - s_0)], \tag{2.2}$$

with: T = Temperature, s = Entropy, h = Enthalpy and $\dot{m}$ = mass flow rate.

The subscript "0" and "$initial$" distinctively refer to the environmental and initial state value of the parameters. The exergy analysis of energy conversion systems such as STS is a common form of analysis. As stated by [40] exergy can be used as a measure of the thermal efficiency of a system. Utilizing this property of exergy is valuable when FDI is considered, since a fault can introduce some change in the thermal efficiency of a STS.

Since the STS system consists of various interconnected components, the exclusive use of exergy is not sufficient for applications of EGBV since the energy transfer between the nodes should also be accounted for. Considering the system as fundamentally a Rankine cycle [21], utilizing the energy attribute, heat-flow, between the components and work done by the components, a comprehensive view of the energy flow in the system can be obtained. R. Koningsveld [21] defines heat flow $\dot{Q}$ as the rate by which energy in the form of temperature difference is transferred between two systems. This is generally defined as the change in heat of a specific substance with mass flow

m. This is expressed by (2.3) as,

$$\dot{Q} = m\frac{dq}{dt}. \tag{2.3}$$

The specific heat (q) is given by (2.4).

$$q = h_{final} - h_{initial}, \tag{2.4}$$

with: h_{final} = the final enthalpy of the substance and $h_{initial}$ = the initial enthalpy of the substance.
Based on this the energy attributes, exergy and heat-flow, are used as the parameters of evaluations for the FDI analysis.

R. Wilson [27] defines a simple undirected graph in layman's terms as a diagrammatic representation of some system by means of points and lines. An example of this is shown in Figure 2.2

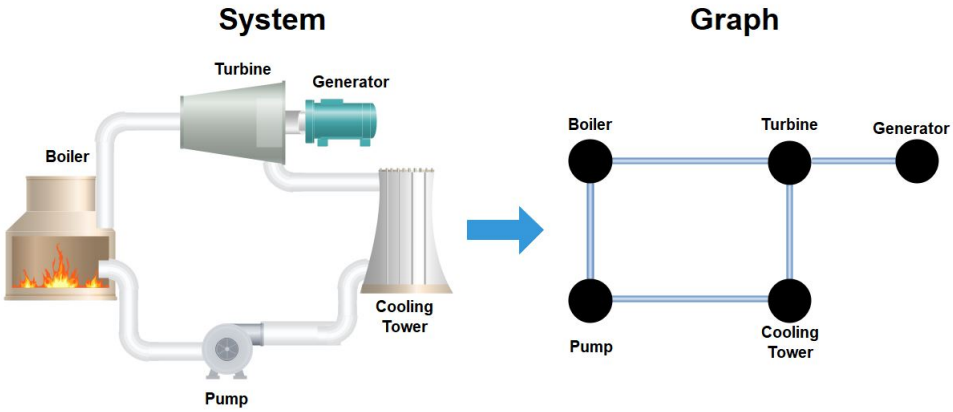


Figure 2.2: Example of a graph representation of a physical system

The points and lines are formally referred to as nodes and links or vertices and edges. For the purposes of this dissertation the terms nodes and links will be used. A formal mathematical definition for a simple undirected graph is given by [27] and [28] as: *A simple graph $G(N,L)$ consists of a non-empty finite set of elements called nodes $N(G)$, and a finite set of distinct unordered pairs of elements $L(G)$, called links. $N(G)$ is called the node set and $L(G)$ is referred to as the link set of G. A link a,b is said to join the nodes a and b, and is normally abbreviated to ab.*

The graph in Figure 2.2 can be written mathematically as $G(N,L)$ with $N(G) = \{b,t,g,c,p\}$ and $L(G) = \{bt,tg,tc,cp,pb\}$. The variables b,t,g,c,p distinctively refers to the Boiler, Turbine, Generator, Cooling Tower and Pump. Since heat flow is directional in a real system (moving from one component to another) a directional graph can be used to represent this. In such a case the links set $L(G)$, consists of an ordered finite set and is graphically indicated by an arrow rather than a line. The direction of heat flow is indicated mathematically by the node variables starting at the node where the heat flow originated and moving towards the node where the heat flow recedes. Typically heat flow from the turbine towards the cooling tower will be represented by a link $L(G) = \{tc\}$. The reverse heat flow will then be $L(G) = \{ct\}$. By attributing exergy and heat flow values to the nodes and links, an attributed graph is obtained that can be used as a representation of a specific state of operation of the system (e.g. faulty or normal state).

The EGBV allows for isolating a fault to the extent that the component, where the fault occurs, can be identified. This capability necessitates the inclusion of the component under evaluation

as a node when composing the graph. This is due to the nature of graphs, since a graph can only communicate information captured in its consisting nodes and links [27]. While executing this, the graph complexity should be kept to a minimum. A minimalistic graph results in complexity reduction in terms of the mathematical analysis. As stated by [29] the less nodes and links present in a graph and the more the graph is structured in an orderly manner, the less complex the mathematical methods for analysing the graph need to be.

The article by B. Gallagher [41] gives an overview of specific graph matching approaches. It notes that graph matching techniques can be either classified as exact or inexact matching techniques. These techniques can be based on graph structure or semantics. In a case where graphs have the same structure but the attributes to the nodes may vary, the matching is semantic based. This simplifies the matching problem since aspects like the NP-complete problem of sub-graph isomorphism [42] do not have to be considered. The resulting approaches used are described in [30]. The article focuses on the heterogeneous distance metric as a method of determining the distance between two vectors or matrices. It defines the Heterogeneous euclidean overlap metric (HEOM) as an improved metric for distance measuring when either categorical or continuous attributes are used. The mathematical expression can be seen in (2.5).

$$HEOM(\mathbf{x_i}, \mathbf{x_j}) = \sum_{n=r}^{R} d_r(x_{i,r}, x_{j,r}), \tag{2.5}$$

with: n = number of subset for which the distance is calculated and R = the number of measured predictors in the subsets. The metric d_r is given by,

$$d_r(x_{i,r}, x_{j,r}) = \begin{cases} \frac{|x_{i,r} - x_{j,r}|}{range_r} & \text{, if } r \text{ indexes a continuous attribute} \\ \delta_{i,j} & \text{, if } r \text{ indexes a categorical attribute} \end{cases}$$

with, $\delta_{i,j} = 1$ if $x_{i,r} \neq x_{j,r}$ or $\delta_{i,j} = 0$ if $x_{i,r} = x_{j,r}$.
The range is given by, $range_r = max\{x_{j,r}\} - min\{x_{j,r}\}$.

D. Randall Wilson [18] proposes further improvements for distance metrics describing specifically the methods used for different attribute types. Another approach can be to utilize eigenvector decomposition. In [43] a nearly optimal matching approach using eigenvalue decomposition for both directed and undirected graphs is proposed. The article describes the use of the attribute (called adjacency) matrices of the graphs. A detailed explanation of the general method is given as well as specific artificial examples. The utilizing of eigenvector decomposition, not for graph matching but as part of the FDI approach, was also proven to be useful for especially the EGBV technique in [3].

2.5 Thermodynamic model

A system model is defined by D. Karnop et al. [44] as simplified, abstract constructs used to predict a systems behaviour. According to [45] a model can either be dynamic or static with different levels of complexity. Models can be used for various simulations, where simulations are defined as the process of predicting functionality and performance of the system. The models can be generally classified under different levels of complexity. These classifications are, purely empirical, semi-empirical, semi-theoretical or fundamental principle models. Since the STS under investigation is analysed on a component level, a semi-theoretical model is developed. P. Rousseau [45] describes this type of model as a component based model that employs detailed mass-, momentum- and energy conservation analysis for the components in the system. Such a model can typically be used

to predict the performance of a system under different operational conditions such as faulty or normal conditions.

Rousseau [46] gives a general layout of the types of equations that are typically used to develop a model. These fundamental equations, component information, fluid parameters and properties are among others:
1) Conservation laws: Mass, momentum (and angular momentum) and energy.
2) Component characteristics: Heat transfer rates, pressure drops based on efficiencies, pipe diameters and heat transfer areas.
3) Fluid properties: Thermodynamic property tables and gas laws.
4) Boundary values: Temperatures, pressures and mass flow.

Effective modelling of a STS requires understanding of the working principles of the system. The book by H.B et al. [47] describes the design, application of and re-rating of steam turbines. M. Topel [48] defines a steam turbine as one of several kinds of turbo-machines. These are machines that are classified as devices in which energy is transferred from or to a continuously flowing fluid by means of dynamically rotating blades. It also states that the operation of this machine is directly associated with the Rankine cycle. The modern implementation of STS's commonly consist of multi stage arrangements of the turbine and can include a high, medium and low pressure turbine stage. To increase system efficiency or power output, the configuration of the STS can be adjusted. The book by Alexander S. Leyzerovich [24] explain the various types of steam turbines and system configuration in detail. Some of these system configurations are known as the reheat cycle and regenerative cycle [49].

A key concept in applying especially the conservation equations, is the control volume. The control volume is defined by A. Sonin [50] as an arbitrarily defined volume consisting of a closed bounding surface, called the control surface, that divides the "universe" into two sections. These sections or parts consist of the parts contained within the control volume and the rest of the universe outside of the control volume. The advantage given by a control volume is that it allows for a framework in which the integral laws in engineering analysis can be applied. An example of such a control volume is shown in Figure 2.3

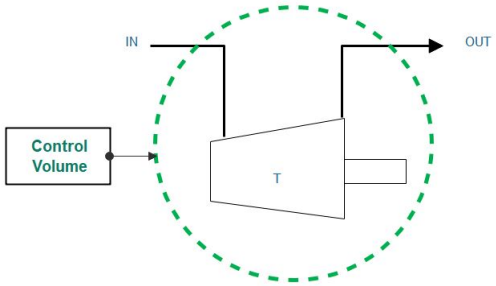


Figure 2.3: Example of a control volume

Using this definition of a control volume (CV), consider Figure 2.4 consisting of a CV with both steady incompressible mass flow rates and energy attributes moving in and out of it. The variables $\dot{W}_{in}$ and $\dot{W}_{out}$ represent the work entering and exiting the CV. Similarly, $\dot{m}_{in}$ and $\dot{m}_{out}$ represents the mass flow rate in and out of the CV. The variable $\dot{Q}_{out}$ is the heat flow out of the system and $\dot{E}_{CV}$ is the difference between the work in and out of the system and the heat flow, such that $\dot{E}_{CV} = (\dot{W}_{out} - \dot{W}_{in}) - \dot{Q}_{out}$.

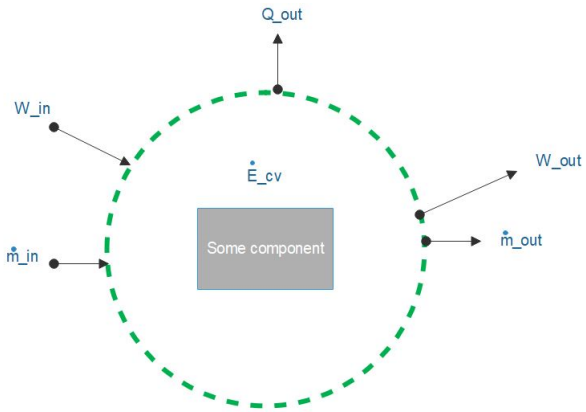


Figure 2.4: Control volume with mass and energy inputs

P. Rousseau [46] and W. Fratzscher [14] states that the differential form of the mass conservation equation for any finite CV can be written as shown in (2.6).

$$v\frac{\partial \rho}{\partial t} + \dot{m}_{out} - \dot{m}_{in} = 0, \tag{2.6}$$

with: v the volume and $\frac{\partial \rho}{\partial t}$ the partial derivative of the fluid density with regards to time. This implies that $v\frac{\partial \rho}{\partial t}$ indicates the mass inside the CV, $\dot{m}_{out}$ the mass flow rate out of the CV and $\dot{m}_{in}$ the mass flow rate into the CV.

Similarly the differential form of the energy conservation equation for a finite CV can be written as shown in (2.7).

$$\dot{Q} + \dot{W} = v\frac{\partial}{\partial t}\left(\rho h_o - p\right) + \dot{m}_{out}h_{out} + \dot{m}_{in}h_{in} - \dot{m}_{out}gz_{out} - \dot{m}_{in}gz_{in}, \tag{2.7}$$

with: g the gravitational acceleration and z the elevation of the fluid above the point of calculation. The term $\dot{Q}$ refers to the heat added to the CV and $\dot{W}$ refers to the work done on the CV. The term $v\frac{\partial}{\partial t}\left(\rho h_o - p\right)$ represent the work due to a change in density and pressure of the fluid inside the CV. The other four terms respectively represent the rate of change in internal and potential energy in the CV.

According to [25, 51] for steady state with no elevation of the fluid the conservation of mass and energy equations can be simplified as shown in (2.8) and (2.9) respectively.

$$\dot{m}_{in} = \dot{m}_{out} \tag{2.8}$$

$$\dot{E}_{cv} = \dot{W}_{in} - \dot{W}_{out} - \dot{Q}_{out} \tag{2.9}$$

Conveying this to a more general case with "n" inputs and "m" outputs to the CV, these equations can be rewritten as shown in (2.10) and (2.11) respectively.

$$\sum_{i=1}^{n} \dot{m}_i = \sum_{e=1}^{m} \dot{m}_e \tag{2.10}$$

$$\sum_{i=1}^{n} \dot{E}_i = \sum_{e=1}^{m} \dot{E}_e \tag{2.11}$$

To develop a semi-theoretical model at a component level model, it is necessary to establish which parts in the system are considered as components that should be modelled and which is a sub-part/component of the components that should not be modelled. For example, a pipe with various valves can be seen as one component with specific characteristics dependant on the valves. The STS in Jeanschwalde consists of turbines, compressors, pumps, heat exchangers and pipes. The components such as the pre-heaters, re-heaters, boilers and cooling towers are all categorised as forms of heat exchangers and will thus follow the same energy and mass conservation laws. The modelling of these five main components as well as some of their subcomponents is described in detail in both [26] and [45].

2.5.1 Turbine

Consider a generic drawing of a steam turbine as shown in Figure 2.5. The variables "h", "P" and "$\dot{m}$" represent enthalpy, pressure and mass flow rate respectively. The subscripts "oi" and "oe" represent the inlet and exit value of the variable in an ideal process. If the subscripts are changed from "oi" or "oe" to "$isen$" it represents the isentropic process.

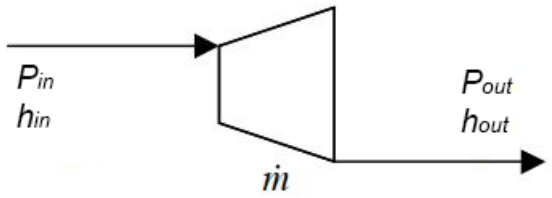


Figure 2.5: Simple generic turbine schematic

F. Alobaid [26] states that the pressure drop over the turbine is a function of the nominal mass flow rate "$\dot{m}_{nom}$", nominal turbine inlet pressure "$P_{in,nom}$" and the nominal outlet pressure $P_{out,nom}$" given that the current mass flow rate $\dot{m}$ is known as shown by (2.12).

$$\dot{m} = \dot{m}_{nom} \times \frac{P_{in}}{P_{in,nom}} \times \sqrt{\frac{1-\left(\frac{P_{out}}{P_{in}}\right)^{\frac{n+1}{n}}}{1-\left(\frac{P_{out,nom}}{P_{in,nom}}\right)^{\frac{n+1}{n}}}}, \tag{2.12}$$

with:

$$n = \frac{ln\left(\frac{P_{out}}{P_{out}}\right)}{ln\left(\frac{P_{out}}{P_{in}}\right) - ln\left(\frac{T_{out}}{T_{in}}\right)}. \tag{2.13}$$

The variable $\dot{m}$ represents the mass flow rate through the turbine.

The pressure ratio is then represented by (2.14),

$$PR = \frac{P_{out}}{P_{in}}. \tag{2.14}$$

According to P.G Rousseau [45], through application of the conservation of energy law, the work done by the turbine due to the isentropic fluid expansion process can be written as in (2.15).

$$\dot{Q}_T = \dot{m} \times \eta_T \left(h_{out} - h_{in}\right), \tag{2.15}$$

with: η_T the efficiency of the turbine.

This can be rewritten in terms of temperature as shown in (2.16).

$$\dot{Q}_T = \dot{m} \times \eta_T \times C_p \left(T_{isen} - T_{in}\right), \tag{2.16}$$

with: C_p the specific heat of the steam at constant pressure.

2.5.2 Compressor

According to [26] a compressor is simply a turbo machine that increases the pressure and enthalpy of a fluid. T. Gresh [52] classifies compressors into either axial or centrifugal compressors. For detailed analysis this source gives a good in-depth discussion of the design and performance optimisations of compressors. Regardless of the type of compressor, considering it as a single component (i.e. the whole system is within the CV under investigation) only a few equations are necessary to calculate the thermodynamic parameters. The relevant input parameters necessary may include design point values for pressure and temperature or any other thermodynamic parameter as long as at least two parameters are known. Also, the pressure ratio, rotational speed, efficiency and sometimes the characteristic curve can be necessary. Knowing the pressure ratio of a compressor, the pressure increase can be calculated similarly to a turbine. Consider the simple schematic of a generic compressor in Figure 2.6.

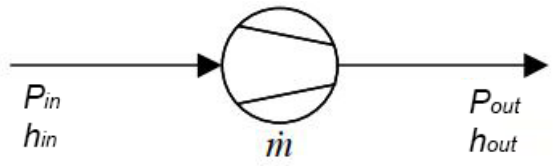


Figure 2.6: Simple generic compressor schematic

From Figure 2.6 it is clear that the mass balance for an incompressible fluid can be calculated using (2.17).

$$\dot{m}_{in} = \dot{m}_{out} \tag{2.17}$$

The ideal lossless energy balance is given by (2.18),

$$\dot{m}_{in} \times h_{in} = \dot{m}_{out} \times h_{out} + \dot{Q}, \tag{2.18}$$

with: $\dot{Q}$ is the work done by the compressor to compress the fluid.

It is stated by [45] that the compressors isentropic work can be calculated by (2.19).

$$\dot{Q}_c = \dot{m} c_p T_i \frac{1}{\eta_c} \times \left(PR^{\frac{\gamma-1}{\gamma}} - 1\right), \tag{2.19}$$

with: PR the pressure ratio of the compressor and η_c the compressors isentropic efficiency given by (2.20).

$$\eta_c = \frac{\dot{Q}_s}{\dot{Q}} \tag{2.20}$$

This can be rewritten in terms of enthalpy as shown in (2.21).

$$\eta_c = \frac{h_{isen,out} - h_{in}}{h_{out} - h_{out}} \tag{2.21}$$

The term γ refers to the ratio between the specific heat capacities at constant pressure and volume given by (2.22).

$$\gamma = \frac{C_p}{C_v} \tag{2.22}$$

2.5.3 Pump

According to [26] a compressor and pump differ in the sense that a pump is, unlike the compressor, designed for moving high density fluids. In order to derive a simplified model for a pump with an incompressible fluid like water, some of the thermodynamic fundamentals must be used. Consider the generic pump diagram in Figure 2.7.

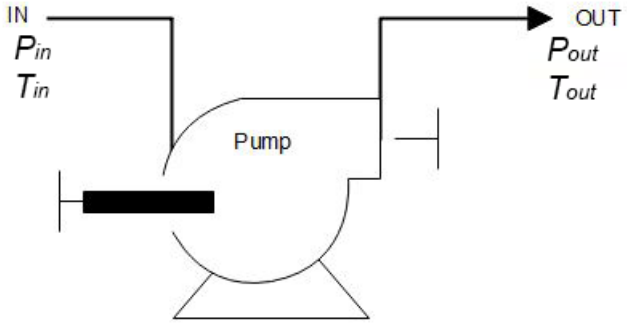


Figure 2.7: Simple generic pump schematic

The conservation of energy is shown in (2.23).

$$\delta u = \delta q + \delta w, \tag{2.23}$$

with: δu the change in specific internal energy, δq the specific heat transferred to the fluid while pumping and δw the work done by the pump on the fluid.
Noting that for ideal reversible processes, the change in specific heat transfer and the work can be written as shown in (2.24) and (2.25).

$$\delta q = T\delta s, \tag{2.24}$$

$$\delta w = -p\delta v, \tag{2.25}$$

with: δv the change in specific volume and δs the change in entropy.
Since enthalpy can be written as $h = u + pv$ (u= specific internal energy of the fluid) and for an isentropic process $\delta s = 0$ it follows that, for an incompressible fluid being pumped, the change in entropy is given by (2.26).

$$h_{out,s} - h_{in} = v\,(P_{out} - P_{in})\,. \tag{2.26}$$

The variable $h_{out,s}$ refers to the total enthalpy for an isentropic process measured at the outlet of the pump. From (2.26) it follows that the rate of isentropic heat transfer to the fluid as work, is obtained by (2.27).

$$\dot{Q}_s = \dot{m}\left(h_{out,s} - h_{in}\right) \tag{2.27}$$

This can be rewritten in terms of the pressure and density of the fluid as shown in (2.28).

$$\dot{Q}_s = \frac{\dot{m}}{\rho}\,(P_{out} - P_{in}) \tag{2.28}$$

The heat transfer rate of the pump can then be obtained by (2.29).

$$\dot{Q}_{po} = \frac{\dot{m}}{\rho\eta_p}\,(P_{out} - P_{in})\,, \tag{2.29}$$

with: $\eta_{po} = \frac{\dot{Q}_s}{\dot{Q}} > 0$.

F. Alobaid [26] states that if the head (H) of the pump, under steady state operation, must be calculated and the max and nominal pump head (H_{max} and H_{nom}) is known along with the nominal volumetric flow rate of the liquid ($\dot{V}_{liq,nom}$) and the nominal pump rotational speed (ω_{nom}), then (2.30) can be used.

$$H = \left(H_{max} - (H_{max} - H_{nom})\left(\frac{\dot{V}_{liq}}{\dot{V}_{liq,nom}}\right)^2\right)\left(\frac{\omega}{\omega_{nom}}\right)^2, \tag{2.30}$$

with: $\dot{V}_{liq}$ the actual volumetric flow rate of the fluid and ω the rotational speed of the pump.

2.5.4 Heat Exchanger

A heat exchanger is defined by both O. Khayal [53] and R. Koningsveld [21] as a device used to transfer heat between two or more sources. These devices can be used to either transfer heat between two fluid streams or to-or-from a source to fluid and vice versa. O. Khayal classifies heat exchangers into five points as:
1) Recuperators and regenerators.
2) Direct or indirect contact transfer process.
3) Construction geometry.
4) Single phase or two-phase heat transfer mechanism.
5) Parallel, counter, or cross flow arrangements.

As seen in Figure 2.8 the heat exchanger can be visualized as having a primary and secondary side. The subscripts $_P$ and $_S$ indicates the primary and secondary sides respectively.

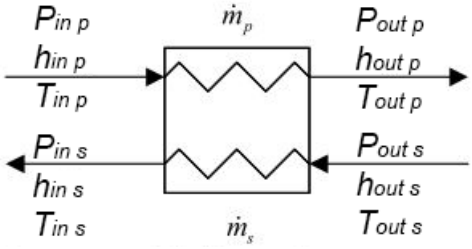


Figure 2.8: Simple generic heat exchanger schematic

The pressure drop over each side of the heat exchanger can be modelled similarly to a pipe or duct as stated by [45]. However if the classification of the heat exchanger is known a more accurate calculation can be done as shown by [53]. The pressure drop can be categorised into three main stages. These are the entrance, core and exit losses as the fluid enters the heat exchanger, moves through the core and then exits the heat exchanger again. Mathematically it can be expressed as shown in (2.31).

$$\Delta P_{HX} = \Delta P_{in} + \Delta P_{core} + \Delta P_{ac} + \Delta P_{exit}, \tag{2.31}$$

with ΔP_{ac} the pressure drop in the core due to a change in acceleration of the fluid as it moves from one phase to another.
Each of these pressure drops can be seen as a function of the average input and output pressure, $\Delta P = f(P_{average})$. To simplify (2.31) it can be stated that the pressure drop is caused by the product of some loss constant and the input pressure of the heat exchanger. This is expressed in (2.32).

$$\Delta P_{HX} = \alpha P_{in} \tag{2.32}$$

Where α represents some pressure drop constant.

Similarly the output pressure can then be calculated by (2.33).

$$P_{out} = \gamma P_{in}, \tag{2.33}$$

with: $\gamma = (\alpha - 1)$

The maximum theoretical heat transfer can be calculated by (2.34).

$$\dot{Q}_{max} = \left[\dot{m} C_p\right]_{min} \Delta T_{max} \tag{2.34}$$

with: $\dot{m}$ the mass flow rate of the fluid stream under consideration and $\Delta T = T_{in,p} - T_{in,s}$.

A simple method to calculate an estimate heat transfer rate is simply to write the heat transfer rate as some fraction of the maximum heat transfer rate. This fraction is known as the heat exchanger effectiveness (ϵ). This can be mathematically expressed as shown in (2.35).

$$\dot{Q} = \epsilon \dot{Q}_{max} \tag{2.35}$$

For more information on the design and internal working of heat exchangers and boilers the book by R. Shah et. Al [54] was found to be of value.

2.5.5 Pipe

Since a pipe (liquid flow) or a duct (gas flow) can simply be visualized as a stationary or, when moving, directional fluid stream the modelling can be greatly simplified. Considering a well insulated pipe with no heat losses it can be stated that the heat transfer rate is simply zero ($\dot{Q} = 0$). In such a case the change in entropy between the input and output of the pipe is also zero ($\Delta S = 0$). Using (2.32) the pressure drop can be calculated in the exact manner as for heat exchangers. Rousseau [45] gives an alternative to (2.32), using the density (ρ) and fluid velocity (V) as shown in (2.36).

$$\Delta P_{pipe} = K\frac{1}{2}\rho V^2, \tag{2.36}$$

with: K simply the loss factor due to friction similar to α in (2.31).
The change in mass flow for an incompressible fluid like water will also be zero for the pipe resulting in $\Delta\dot{m} = 0$.

Additional details to calculate specific aspect of these components can be found in [21] and [26].

2.5.6 System modelling approach

The modelling approach followed by [26] and [45] consists of a component break down of the system which implies that each component in the system is modelled separately and then all of the models is combined to form the final model. Each component is modelled in terms of its thermodynamic parameters and then systematically connected coherently with the system layout. Simplifying the system by dividing it into main components and then systematically adding more components helps with managing large system model development. The approach by P.G Rousseau [45] can be summarised in five steps:
1) Compose a simplified system diagram.
2) Define the known boundary conditions of the various components.
3) Determine the component's characteristics.
4) Solve each component's thermodynamic parameters mathematically.
5) Connect each component by starting at some point in the system and calculating the mass and energy balances transferred from each component to the next.
Once any two thermodynamic parameters (e.g. T, P), for some point is known, any other parameter (e.g. s, h) can be obtained from the thermodynamic tables. Using this fact the necessary enthalpy and entropy parameters can be calculated and used to determine the exergy and heat flow in the system.

The article by T. Koroglu et al. [55] describes an exergy analysis for a steam power plant on a marine vessel. The details regarding both a conventional and advanced exergy analysis are given. The article explains and illustrates the advantage of exergy analysis and especially shows how it can be used to improve system efficiency. The method also establishes to what extent the system has endogenous and exogenous exergy destruction and concludes that the avoidable exergy destruction can have a potential system efficiency improvement of 10%. The software Ebsilon was used as the thermodynamic modelling tool.

A similar study on exergy analysis was done by M. Elhelw [**?**]. In this article the effects on exergy destruction, by changing some input parameters of a component in a steam turbine plant, were evaluated. The article gives a clear explanation of the calculations done to determine exergy destruction for each type of component. A detailed mathematical flow diagram for solving a system for exergy analysis purposes is also available. The article concludes that the turbine for both half and full load operation is the second largest source of exergy destruction with the boiler being first.

The research done by G. Ahmadi et al. [56] results in similar conclusions as [**?**]. The article notes however that only by analysing the system using energy analysis will the condenser result in being the component with the highest potential for optimisation. Including exergy however, clearly indicates that the condenser operates at almost maximal efficiency. The software used as the thermodynamic tool was EES. Each of these three sources confirms the effectiveness of using exergy analysis for thermal plant optimisation. They also agree on the mathematical equations

used in general for exergy calculation as is shown in (2.37), (2.38) and (2.39) with b_Q the exergy transferred by heat at temperature T and $\dot{b}$ the total exergy of a fluid flow.

$$b_Q = Q\left(1 - \frac{T_o}{T}\right) \tag{2.37}$$

$$b = (h - h_o) - T_o\,(s - s_o) \tag{2.38}$$

$$\dot{b} = \dot{m}b = \dot{m}\,((h - h_o) - T_o\,(s - s_o)) \tag{2.39}$$

2.6 Data analysis methods

The data from the physical system is commonly measured in the form of temperature, pressure and mass flow values. To analyse this data it is necessary to remove unwanted noise, fluctuation and outliers in the data set. These fluctuations and noise can be due to normal fluctuation in the system or it can be due to imperfections in measuring instruments [57]. The removed data can be used as an indication of the amount of variance that can be expected under normal operations, if neglecting sensor noise and outliers. S. Athichanagorn [58] defines outliers as data points that lie away from the main trend of the data. S. Moosavi et al. [57] states that these outliers demonstrate a problem with either the measuring equipment, the data recording process or human error. The article focused on de-noising techniques for pressure data. Techniques like SMA, Savitzky-Golay filter (SG), Autoregressive moving average (AMA) and locally weighted scatter-plot smooth (LOWESS & LOESS), among others, are evaluated and compared using the least square error method. Each of these filters can be mathematically described as follows:

1) SG: For a dataset of n pairs (x_i, y_i) where x_i is the independent variable and y_i is the dependent variable, the data is fitted using the polynomial expression (2.40). The SG then requires that the sum of least squares be minimised by (2.41) where e_t is the total error and $N+1$ is the total number of data points.

$$p_n(x) = \sum_{i=0}^{n} a_i x^i \tag{2.40}$$

$$e_t = \sum_{i=0}^{N} [p_n(x_i) - y_i]^2 \tag{2.41}$$

2) AMA: For a time series of data x_i the AMA combines a auto regressive model with the moving average calculation as expressed in (2.42). The terms $\theta_1, \theta_2, ..., \theta_p$ are the model parameters and η_t is the noise for the auto regressive part. The terms $\sigma_1, \sigma_2, ..., \sigma_q$ are the model parameters and η_t is the noise for the moving average part.

$$x_t = \eta_t + \sum_{i=0}^{q} \sigma_i \eta_{t-1} + \sum_{i=0}^{p} \theta_i x_{t-i} \tag{2.42}$$

3) LOWESS&LOESS: For this filter the least square criterion is modified by (2.43) using the non-negative weight factor W_i after which the square error is minimised. For each point in the span of the dataset the weight factor is calculated using (2.44).

$$e_t = \sum_{i=0}^{N} W_i e_i^2 \tag{2.43}$$

$$W_i = \left(1 - \left|\frac{x - x_i}{d(x)}\right|^3\right)^3 \tag{2.44}$$

The article by A. Raudys [59] specifically focuses on the optimising of moving average methods for stock price applications. The article describes the application and effects of simple, weighted, sinus weighted, double exponential and more moving average methods. The simplicity and effectiveness of the SMA for data smoothing is emphasised. For a set of data points x, both [57] and [59] give the mathematical expression for the SMA as shown in (2.45).

$$\bar{x}(i) = \frac{1}{2N+1}\left(x(i+N) + x(i+N-1) + ... + x(i-N)\right). \tag{2.45}$$

With: $\bar{x}(i)$ the smoothed value, N the number of neighbouring data points on each side and $2N+1$ the span or window for which the average is calculated.
S. Moosavi et al. states that this technique is in principle a low pass filter.

An article by N. Gallagher [60] focused on the SG smoothing technique as a digital filter. It describes the technique as a method to increase the signal-to-noise ratio without distorting the signal itself. The method uses convolution by fitting a low-degree polynomial to successive subsets of adjacent data points using the linear least square method. The article concludes that the filter is most popular in signal processing due to its ability to reduce high and low frequency noise when combined with a differentiation method. According to [57] one of the advantages of this technique is that it keeps features of the noise or fluctuations such as extrema that are often removed by techniques such as SMA. Another article that give more detail on the application of this method is [61].

The AMA process is described by R. Katz et al. [62]. The article evaluates different AMA processes in order to fit time series data. According to [57] this method has long been used for the filtering and smoothing of data. The advantage of this method is that it implements both the autoregressive and moving average where the auto regressive part describe the time series as a linear function of previous observed data points. The article by R. Katz et al. concludes that for meteorological time series a lower order autoregressive process fits time series data adequately.

The article by S. Moosavi et al. concluded that the LOWESS & LOESS method gives the best results for the application that was tested, but it also states that there are advantages to the other methods and combinations thereof. Using the SMA for data containing low noise or fluctuations is accurate and simple to implement but the accuracy significantly decreases for high amounts of noise. A last method that is commonly used for noise and fluctuation removal is the Kalman filter. This filter is not considered since the moving average filter that was used is a less complicated method that is sufficient for this study. The article by Y. Zhang et al. [63] can be seen for more detail on this method as applied to temperature sensor data.

2.7 Critical literature review

The literature regarding the FDI for STS clearly indicates a need for research on energy based FDI in these systems. The research on traditional methods of FDI in STS like, among others, PCA and SVM is thoroughly described in literature with specific focus on the use of the aforementioned techniques. The application of energy and more specifically, exergy is commonly used as a performance analysis for STS, but not for fault detection [16], [40]. The energy approach utilizing graphs by [15] will be implemented with data from the lignite plant in Jeanschwalde. The application of the method will therefore be with reference to the article [15] as well as the articles [6] and [51] which further implements the method.

As a method of validation, the FDI will be evaluated using a static metric as a performance metric. The metric will be used with reference to the approach given by [37]. The energy characterisation as applied to the EGBV for FDI is a relatively novel concept and is not yet fully established. The approach used for this was extracted from analysing the articles by G. Van Schoor et al. [15] as well as the lignite plant in Jeanschwalde and basic graph theory as described by [27]. The formal establishing of a method for energy characterising will therefore be a topic of interest for further studies.

The methods for graph composition and mathematical analysis thereof is fully described in [27], [22] and the notes of [28]. The structural properties of interest in a system are connected with the design of FDI using three markers of interest [22]. In terms of graph matching and data mining [29] may be considered as good sources. The article, describe the theoretical methods to use graphs for web-based data mining. This research is of value in terms of the methods available for graph matching that can be extended to graphs in this research. Another article [64], that is a useful source to note, describes a balanced graph matching technique that can be used to improve standard matching approaches. The method from this source is however not used in this research but forms a good reference for an alternative method to evaluate matching robustness.

In terms of the thermodynamic model the books by [21] and [26] describe the fundamental theory. The specific modelling approach used in EES is extracted from the notes in and [46].

The data analysis methods described in [59] regarding the simple moving average as a filter in stock prices is only noted for basic fundamental understanding of the SMA. Due to the simplicity and effectiveness of the moving average and more specifically the SMA for data smoothing, this method will be used. The article by [57] is used as the basis for applying the filtering or smoothing method. It is important to note that a number of alternative methods is available but most require more complex data preprocessing or more complex code than simply using a moving average.

CHAPTER 3

Steam turbine system model

This chapter discusses the development of the static model in EES. A thorough theoretical overview of the physical system is given along with context of the actual set-up of the plant. All the assumptions for the static model are declared in the modelling methodology section and a short description of the modelling software EES is given. The modelling of each component based on the literature is described and finally a system model in EES is developed. The model validation is done by using the data obtained from the dynamic Ebsilon model. Lastly the chapter concludes with a brief note regarding the use of the dynamic model in Ebsilon for fault simulation.

3.1 Introduction

For energy based FDI there are three high level objectives that must be completed before the FDI can be accomplished. These three objectives can be describes as: Obtaining a model to simulate the system, composing an attributed graph of the system and acquisition of normal-and-fault state data from the physical system and through simulations. In chapter 2 section 2.5, the literature and theory regarding the development of a static thermodynamic model was described. The focus was placed on the mathematics needed to model a component in terms of energy and mass balance, and how different characteristics can be calculated for certain components. Each of the fundamental components in a typical STS were discussed and a final description of the system modelling approach were given.

In this chapter the system in Jeanschwalde will be described in terms of the thermodynamic cycle. This is done to show that the theoretical system cycle is a Rankine cycle and allows for the development of a static thermodynamic model based on the literature and equations of chapter 2.

In order to accomplish this the different components and their working principles are clearly stated with a theoretical description of the system. The modelling methodology and the assumptions for the purpose of mathematical modelling are stipulated as well as a short description of the modelling environment. The energy and mass balance equations for each of the components in the physical system are then composed and finally the model is validated. The chapter ends with a short conclusion of the thermodynamic model's validation and its accuracy and hence, completes the first high level objective.

3.2 Theoretical system description

A steam turbine system is commonly known to work on a simple Rankine cycle. The simplest of these systems consists of mainly four components. These components being the water pump, boiler with super heater, steam turbine and cooling tower [23] and [24]. A simple system schematic can be seen in Figure 3.1. The common thermal fluid is water. The water will move through the system while undergoing various phase changes from liquid to two-phase mixture, steam and superheated steam. The process can be described by systematically moving through the system from point 1 to 8 as indicated in Figure 3.1.

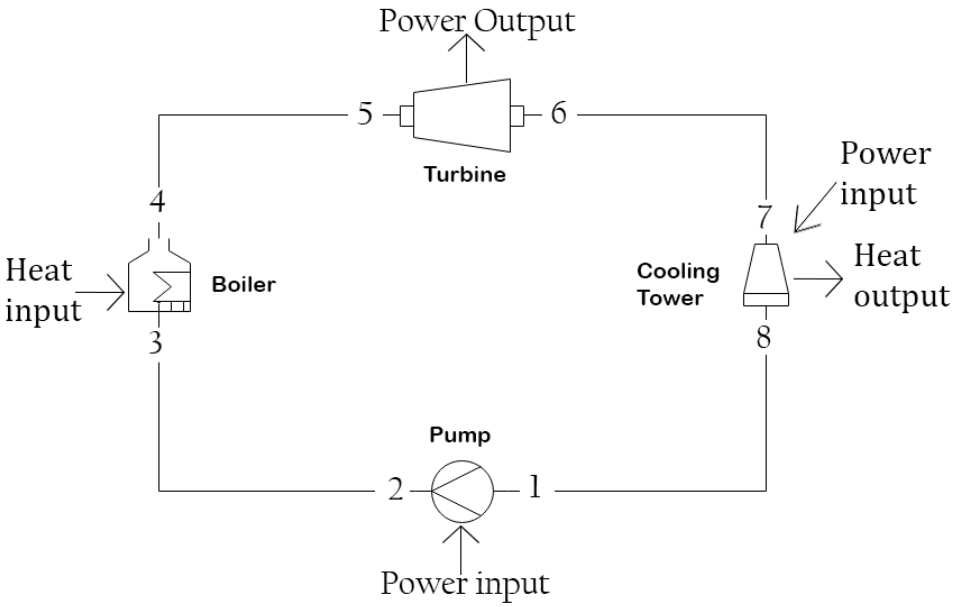


Figure 3.1: Simple Rankine cycle

Starting at the pump inlet (point 1 in the schematic), water at atmospheric pressure and temperature is compressed as it moves through the water pump. The pump will typically be either driven by an electrical motor or a turbine in the system. To illustrate this power addition some power input to the pump is shown on the schematics. Due to the higher pressure at the pump outlet (point 2) the cold-water stream is transported to the boiler (point 3). Once again, some energy is added to the boiler, typically this would be coal in a lignite plant. The water is then heated to superheated steam (point 4) and then moves to the turbine (point 5). The super-heated steam is adiabatically expanded through the turbine, resulting in some power generated in the form of mechanical torque. The fluid will exit the turbine (point 6) as a mixture of water and water vapour [24]. The temperature at this point is still higher than atmospheric temperature and must be cooled down. In order to accomplish this the fluid is moved through the cooling tower (point 7 to 8) at which point the extra energy is extracted from the fluid causing it to cool down. Finally, the cycle can repeat in the close looped system.

A common method of visualizing the process is to draw a T-s (temperature-enthalpy) or h-s (entropy-enthalpy) graph. The latter is also known as a Mollier diagram. These graphs can be seen in Figures 3.2 A and B.

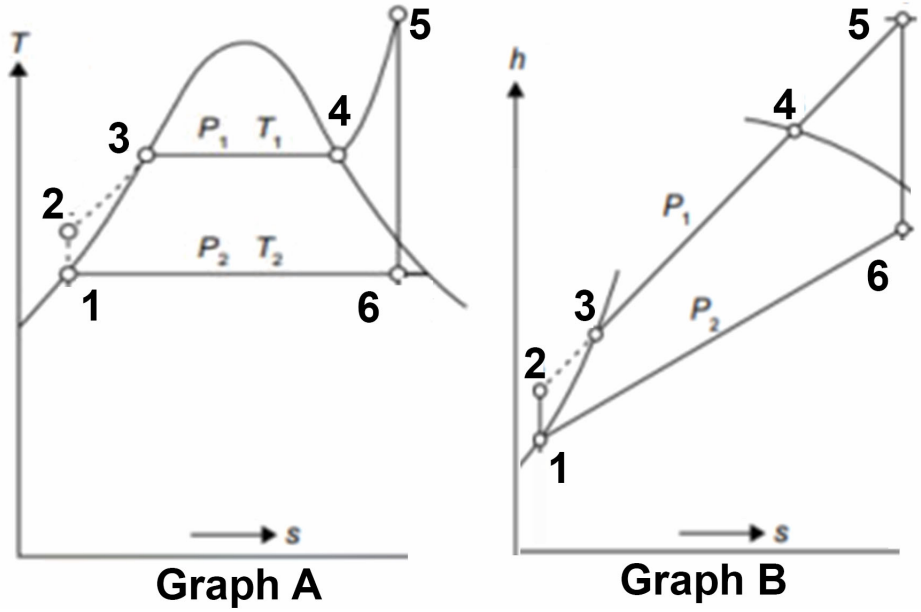


Figure 3.2: T-s (graph A) and P-s (graph B) graphs of the Rankine cycle

The ideal conditions at each of the points is described by [23] as:

1) The fluid, in a condensed water state, enters the pump and an adiabatic pressure rise ensues (points 1 to 2).
2) The pressurised water enters the boiler and is heated under constant pressure to saturation at point 3.
3) The saturated water is evaporated under constant pressure (points 3 to 4).
4) The evaporated water vapour is superheated by the final part of the boiler (points 4 to 5).
5) The fluid, in a superheated state, enters the turbine and expands adiabatically as it moves to the turbine exit (points 5 to 6).
6) Wet steam exits the turbine and then condenses as it moves through the cooling tower (point 6 to 1).

Since most industrial systems require high efficiencies and performance the simple Rankine cycle is normally expanded into some complex cycle with multiple turbine stages and a variety of pre-heaters throughout the cycle. The boiler is also commonly equipped with multiple stages. Such a cycle is known as a reheat regenerative Rankine cycle [49].

The physical system in Germany is one of these typical systems. As mentioned in the background section 1.1 the physical system consists of 6 units/blocks of which only 2 units is in operation. A detailed schematic of one of these units are shown in A, Figure A.1. For analysis purposes the schematic can be simplified as shown in Figure 3.3.

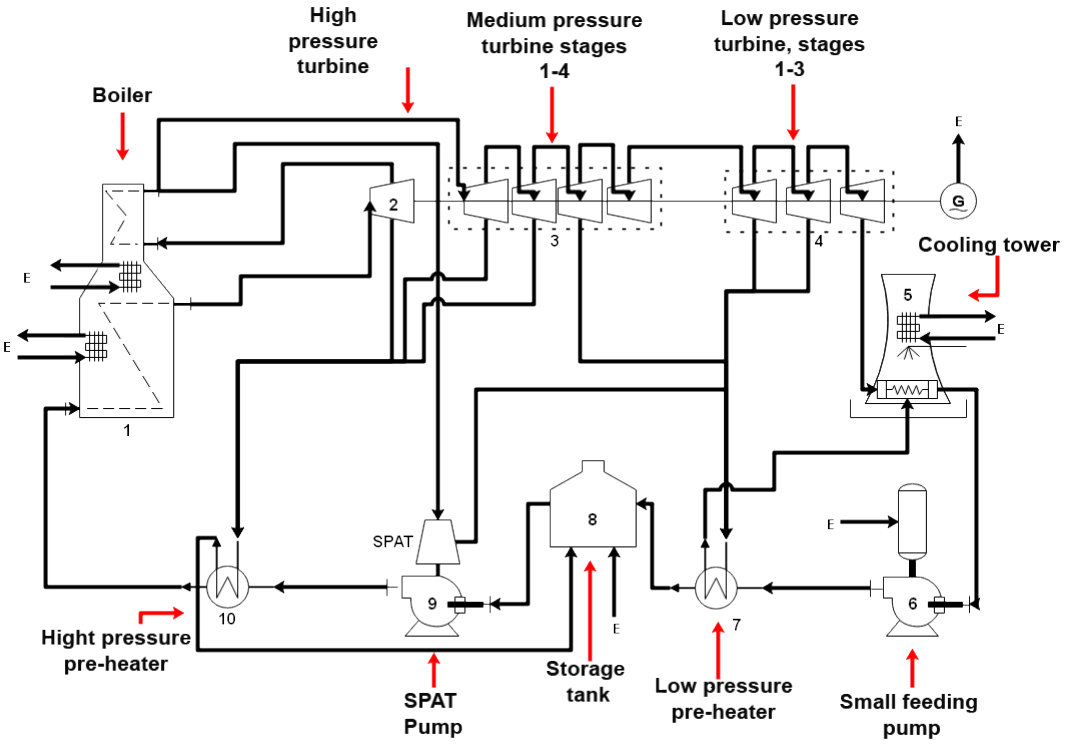


Figure 3.3: Reduced system diagram of the physical system

Each component is labelled from 1 to 10 and the input or output of energy from the environment in the form of water, coal or electrical energy is indicated with an "E" (Environment). The thermodynamic cycle of the system can be illustrated by starting at the inlet of the small water pump (point 6). Water at atmospheric pressure and temperature is pumped towards the low-pressure pre-heaters (point 7). The pre-heaters use energy from partial fluid streams of the 3^{rd} stage of the medium pressure turbine and the low-pressure turbine to heat the water. This preheated water then enters a central storage tank (point 8) at +- 150^{o}C. The storage tank serves as a central water storage to which external water can be added if required. The main water pump in the system (point 9) is driven by a turbine ("Speise pumpen antriebs turbine" (SPAT)) and pumps water from the storage tank to a second set of high-pressure pre-heaters (point 10). The high-pressure pre-heater uses energy from the partial fluid streams of the 1^{st} and 2^{nd} stages of the medium pressure turbine and the high-pressure turbine to heat the fluid to approximately 230 ^{o}C. This fluid then finally enters the boiler (point 1) where it is heated to superheated steam. The boiler consists of a primary and secondary side. The superheated water vapour is adiabatically expanded in the high-pressure turbine (point 2) that generates mechanical torque. According to [24] the high-pressure turbine will cause a pressure and temperature drop that will result in the fluid becoming steam with a quality of 100%, or at the lowest 90%. In order to increase efficiency and power output the fluid is reheated by the boiler before it enters the medium pressure turbine (point 3). A partial fluid flow from this reheated fluid-stream is also used to drive the SPAT. The medium and low-pressure turbines consist of 4 and 3 turbine stages respectively. As the fluid is adiabatically expanded through each of the stages, the temperature drops to a final value of approximately 35 ^{o}C at the outlet of the low-pressure turbine (point 4). A shaft connecting each of the turbines allows for each stage to add mechanical torque to the shaft used to drive the electrical generator. The cooler fluid exiting the low-pressure turbine is still at a temperature higher than atmospheric temperature and is condensed by the cooling tower (point 5). This cycle repeats to produce consistent electrical energy.

3.3 Modelling methodology

The modelling methodology followed is similar to the approach described by [45]. The process consists of four basic aspects, 1) Conservation laws, 2) Component characteristics, 3) Fluid properties and 4) Boundary values. Through awareness of the aforementioned process the system can be modelled using the component analysis methods described in the literature. The variables such as temperature, pressure, entropy, enthalpy, exergy and work at various points in the system are required as the output of the model. The model is built by composing a EES program starting with defining the component characteristics and boundary values. A component based solving approach is followed as illustrated with the flow diagram in Figure 3.4. This approach is followed in order to systematically solve the system parameters for each component.

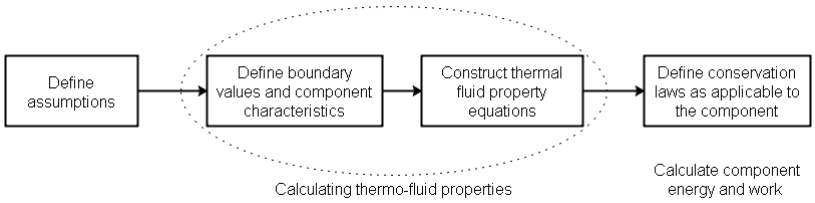


Figure 3.4: Process flow of the component based modelling methodology

The component characteristics such as the isentropic efficiency, mechanical efficiency and pressure ratios are obtained from either the STS specification document or the Ebsilon model set-up values.

3.3.1 Assumptions

To ensure simplicity of the model and to increase the accuracy, some assumptions are made. It is assumed that the system is evaluated as a steady flow system due to the scale of the system, that would be impractical for this study, to model otherwise. This includes one-dimensional flow, resulting in each component being analysed as a complete system, as defined in [65]. This allows for the energy and mass balance equations to be simplified from the differential form to a steady state form.

Regarding the component geometry, the inlets and outlets of the components are assumed to be in the same xy-plane with no elevation in the z-axis direction. This implies that no potential energy components need to be included in the energy balance equation. The system is also stationary with regards to the environment, allowing the exclusion of the kinetic energy component in the energy balance equations.

Regarding the pipes and its connections, all pipes are assumed to be ideal, thus adiabatic (heat losses is zero), and with no friction losses. To account for actual pressure drops in the system, a constant loss factor is included that accounts for the fraction of pressure drop seen in the physical system, due to valves and other losses that are not accurately accounted for in the model.

It is assumed that the compression and expansion of the fluid that happens within the turbines, compressors and pumps, is adiabatic. This allows for composing an energy balance for a component without having any heat loss to the environment. It is also assumed that the water is an incompressible fluid while in liquid state from the outlet of the cooling tower to the outlet of the SPAT pump.

These assumptions can be summarised as:

1) The system operates under steady state one-dimensional flow.
2) The system geometry is static and in one plane with regards to the z-axis.
3) There is no friction within the pipes in the system and all pipes are adiabatic.
4) Adiabatic expansion and compression occur within the turbines, compressors and pumps.
5) The water is an incompressible fluid while in the liquid phase.

3.4 Modelling environment

The modelling software used for the program is known as Engineering Equation Solver or EES. The software allows for the simultaneous solving of multiple equations. According to the software manual [66] there are multiple algorithms used depending on what needs to be solved. If there is one degree of freedom, EES will either use a Golden Section search or a recursive quadratic approximation method. The recursive Quadratic The approximation method is usually the faster method. The Golden Section method is, however, more reliable. For cases with Multi-dimensional optimisation, the software will implement either Direct Search or a Variable Metric algorithm. For differential equations, a fourth-order Runge-Kutta algorithm is utilized as a function with four inputs. The inputs include the initial value and final value of the intendent variable, the final value of the dependent variable, and the step size. In this sense, the software does not solve the equations by following a linear solving algorithm but rather solves each equation if the number of variables defined is equal to the number of equations [67] and [68]. Built-in functions allow for the calculation of fluid properties through the use of steam and thermodynamic-tables. In this regard, two-parameter values as input are required for the functions to be able to obtain any other thermodynamic parameter from the steam tables.

The composition of a parametric table allows for repetitive calculations using changing variables. This tool is useful when the model is evaluated for a number of different input conditions to simulate various operating points. It can also be used to model changes in component parameters. For example, different loss values for a valve, depending on how much the valve is opened.

3.5 Component modelling

Since a component based modelling approach is followed, as stated in the modelling methodology section 3.3, the mathematics related to each component is discussed in detail. Once all mathematics for each component is clear the final model can be composed by combining each of the components following the system layout as shown in Figure 3.3. The calculation of thermodynamic parameters using the built-in function with the steam tables in EES is simply calculated by (3.1).

$$Z = f\{X, Y\} = Z(Steam\_IAPWS; X = X_{value}; Y = Y_{value}), \tag{3.1}$$

with: Z the thermodynamic parameter that is unknown,
and X and Y the known thermodynamic parameters at which Z is calculated. The $Steam\_IAPWS$ refers to the build in steam tables in EES.

3.5.1 Turbine

There exist two basic types of turbines (Impulse and reaction turbines). The turbines installed at the plant are all Russian made reaction, Leningradsky Metallichesky Zavod (LZM) turbines. These

machines consist of stationary blades through which the steam passes to cause a reaction of the rotor blades that then turns the axis of the machine [47]. The machines can be installed in various mechanical configurations that include a horizontal or vertical installation. The turbine installation in the plant consists of only horizontal configurations. This allows for assumption 2 to be valid. From the system diagram in Figure 3.3 it can be seen that the high, medium and low-pressure turbines are all connected in series causing one to feed the other with mechanical torque. Since the model is component based each of these turbines can be evaluated as a whole. The medium and low-pressure turbine are evaluated as 4 and 3 separate turbines respectively, with each stage of the medium and low-pressure turbines taken as a separate unit.

Considering the turbine as a unit within a control volume as described in the literature section 2.5.1, (3.1) can be used to calculate the thermodynamic parameters at the input and output of the turbine, given that at least two parameters are known.

For a specific input temperature (T_{in}) and pressure (P_{out}) of the turbine and with the power generated ($\dot{W}_T$), the isentropic efficiency (η_s), mechanical efficiency (η_m) and the nominal mass flow rate ($\dot{m}_{nom}$) all known, the adiabatic expansion of the steam turbine work can be calculated using (3.2).

$$\dot{W}_T = \dot{m}_T \times \eta_m \times (h_{out} - h_{in}) \tag{3.2}$$

A reference for the variables is given in the schematic shown in Figure 3.5.

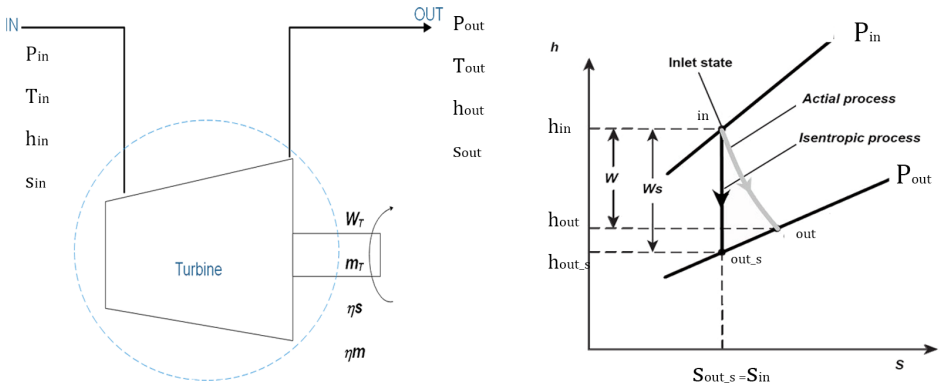


Figure 3.5: Schematic of a turbine (left) and the actual vs isentropic enthalpy-entropy graph (right)

Since all turbines in reality have an isentropic efficiency this should be accounted for. As seen from the h-s graph the actual energy produced by the turbine is less than the isentropic energy input. This implies that the isentropic thermodynamic parameters must be used to calculate the real amount of energy that must be supplied to the turbine. Using (3.3) as defined by [21], the work done by the turbine in the isentropic case can be calculated.

$$\dot{W}_s = \dot{m}(h_{in} - h_{out\_s}), \tag{3.3}$$

with, h_{oes} the isentropic enthalpy.
The relation between the actual work and the isentropic work is given in terms of the isentropic efficiency as shown in (3.4).

$$\eta_s = \frac{\dot{W}_T}{\dot{W}_s} \tag{3.4}$$

Through the use of (2.14) with a pressure ratio determined by the turbine manufacturer the output pressure (P_{oe}) can be calculated. This results in all necessary equations to solve any of the thermodynamic parameters. Once each of the parameters is obtained the exergy change over the turbine can be calculated using (3.5.

$$\Delta \dot{b} = \dot{m}_T([(h_{out\_s} - h_o) - T_o(s_{out\_s} - s_o)] - [(h_{in} - h_o) - T_o(s_{in} - s_o)]) \tag{3.5}$$

The amount of energy entering and exiting the turbine can simply be calculated by substituting the inlet and outlet enthalpy (h_{in} and h_{out}) into h_x in (3.6).

$$\dot{q} = \dot{m} \times h_x \tag{3.6}$$

3.5.2 Pump

Pumps are some of the most frequently used machinery in the industry and can be found in single or multiple stages. Since all pumps are either used to increase flow or pressure a general approach to the modelling of such machines can be taken. Considering that in this regard pumps convert mechanical energy into kinetic energy, the outlet pressure of a pump will correspond to the kinetic energy [69]. The book by B.El Hefni [69] gives a detailed explanation of the modelling of centrifugal pumps and is used along with [26] and [21] to develop the pump model. Considering a real pump with isentropic efficiency (η_s) and mechanical efficiency (η_m) as stipulated by the manufacturers, a similar approach as with the turbine can be followed.

Figure 3.6 gives a schematic of the pump with the variables that are used or calculated. Assuming that at least 2 input thermodynamic parameters (P_{in} (pressure) and T_{in} (temperature)) are known, (3.1) can be used to calculate the other thermodynamic parameters at the inlet. The outlet pressure can simply be calculated with (3.7) using the pressure ratio (PR_{po}) as specified by the manufacturer.

$$PR_{po} = \frac{P_{out}}{P_{in}} \tag{3.7}$$

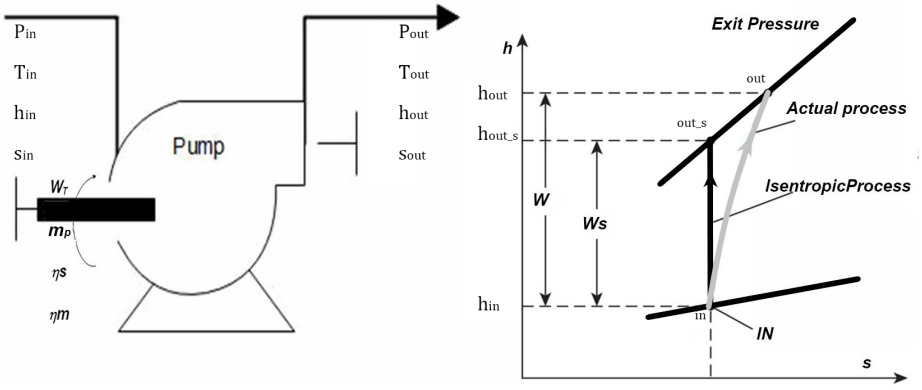


Figure 3.6: Schematic of a pump (left) and the actual vs isentropic enthalpy-entropy graph (right)

The isentropic process requires less energy input than the actual process as shown in Figure 3.6. This implies that the thermodynamic parameters for the actual process must be used to calculate the actual required energy. The actual work done by the pump can be calculated using (3.8).

$$\dot{W}_p = \dot{m}_p \eta_m \times (h_{out} - h_{in}) \tag{3.8}$$

To calculate the isentropic work the isentropic efficiency can be used as shown in (3.9).

$$\dot{W}_s = \eta_s \dot{W} = \dot{m}(h_{out} - h_{in}) \tag{3.9}$$

To determine the mass flow rate of the fluid moving through the pump the average mass flow rate ($\dot{m}$) can be used as shown in (3.10).

$$\dot{m} = \frac{\dot{W}\rho}{P_{out} - P_{in}}, \tag{3.10}$$

where: ρ is the average density of the fluid between the inlet and output of the pump.

3.5.3 Boiler

According to [69] the boiler forms the most complex subsystem of the power plant. The boiler can be described as a component consisting of a combination of 2 circuits, 1) the water/steam circuit and 2) the flue-gas circuit. According to [70] the boiler is also known as the steam generator because, as the name suggests, the main purpose is to generate steam through the heating of water. In this sense the boiler differs from a preheating heat exchanger especially when considering the differences between the energy sources of the two components. The book by B. El Hefni [69] describes a simplified boiler model that models the combustion inside the boiler furnace and the heat transfer, to the cold fluid flowing within the tubes of the boiler. When considering the boiler only as two sections consisting of the furnace (the hot section) and the fluid tubes (the cold section), the boiler model can be developed as if it is a heat exchanger with the exception that the boiler as a unit will consist of heat exchange sections representing the boiling, super heating and reheating part of the unit. The schematic in Figure 3.7 represents a simplified visualisation of the boiler consisting of both the boiling (primary chamber), super heating and reheating section of the physical system in the plant. When considering each of these sections the hot side of the heat exchanger will be referred to as the primary side and the cold side will be referred to as the secondary side.

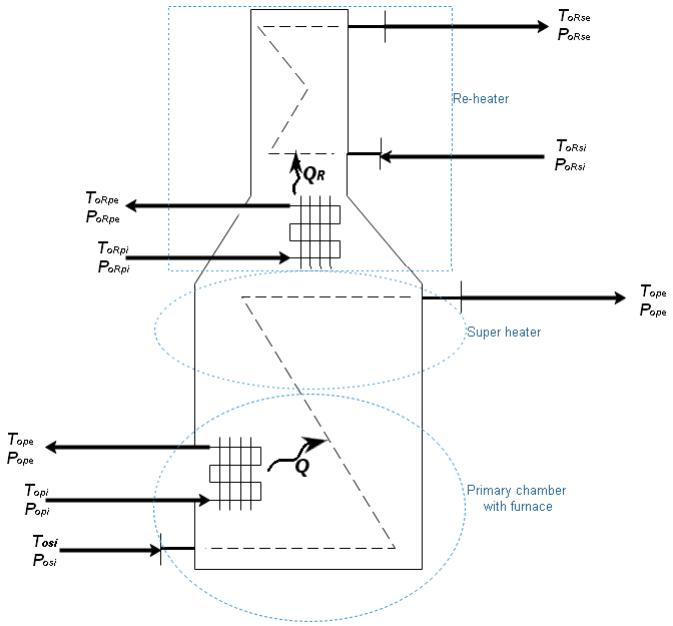


Figure 3.7: Simplified schematic of the boiler

The subscripts R, P and S in the drawing refers to the reheat section, the primary side and the secondary side respectively.

With the aforementioned facts in mind each section of the boiler is thus modelled as individual, but connected, heat exchangers, with the fluid entering the primary heat exchanger which is connected to the super heater. After the steam is cooled through adiabatic expansion in the turbine, it is reheated in the reheat heat exchanger. Assuming that the pressure (P_{osi}), (P_{opi}) and temperature (T_{osi}), (T_{opi}) is known the other thermodynamic parameters at these points are calculated using (3.1).

The pressure (P_{ope}), can be calculated using (2.33) with $0 < \alpha > 1$, determined by the boilers geometry, and ΔP_{XH} determined by the boiler specifications. The outlet pressure and temperature on the secondary side (P_{ope} and T_{ope}) are indirectly determined by the energy requirement of the boiler and limited by the maximum pressure and temperature specifications of the boiler. Using (2.35) the maximum heat transfer can be calculated. The immediate heat transfer is given by (3.11).

$$\dot{Q} = \dot{m}\Delta h. \tag{3.11}$$

The mass flow, $\dot{m}$ for this case is determined by the pump supplying the fluid to the boiler.

3.5.4 Condensing tower

The cooling tower, as the name suggest, is a component used to cool a fluid from some temperature to another. In its application within a power plant, the cooled temperature is normally equal or close to the atmospheric temperature. The book by John C. Hensley [1] describes various types of these cooling towers, among which is the mechanical draft tower, Figure 3.8. This tower consists of multiple fans at the the bottom of the tower to provide an increased airflow into the tower. As stipulated by D. Kroger [71] an air-cooled heat exchanger is added inside the airflow path as shown in Figure 3.8. D. Kroger also notes that the use of approximate methods for design and evaluation of performance characteristic are adequate as shown by a number of authors. When considering the cooling tower to be a simple heat exchanger the approach by [45] can be followed to develop a model.

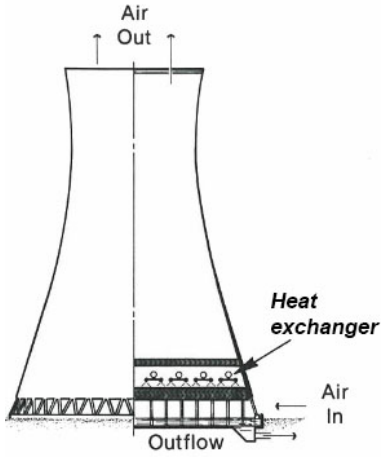


Figure 3.8: Schematic of a mechanical draft cooling tower as obtained from [1]

In Figure 3.9 the schematic for the cooling tower, on which the model is based, can be seen. The hot water entering the cooling tower is referred to as the primary side and the cool fluid entering from the environment is referred to as the secondary side. On the schematic a third input stream is shown on the primary side. This stream is mixed at the inlet with the main fluid stream on the

primary side. The stream is a partial fluid stream that was used in the low-pressure pre-heaters in the physical system.

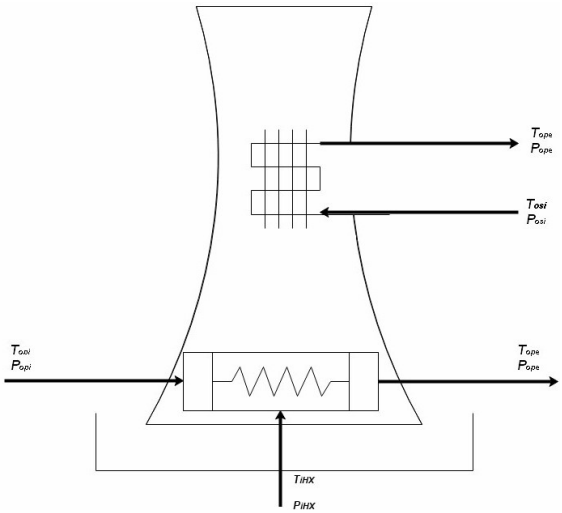


Figure 3.9: Schematic of the cooling tower as used for the model

Given that the inlet pressure and temperature (P_{in_p} and T_{in_p}) on the primary side is known the enthalpy, entropy and other thermodynamic parameters can be calculated using (3.1). The pressure drop between the input and output of the primary or secondary side can simply be evaluated using (2.32). Following the same approach as for the boiler, the output pressure can then be calculated using (3.12).

$$P_{out} = P_{in} - \Delta P. \tag{3.12}$$

The fluid stream that is supplied by the heat exchanger, and is entering the primary side of the cooling tower, has the same pressure as P_{in_p} and will therefore have no effect on the pressure of the output stream P_{out_p}. The aforementioned approach can be repeated for the pressure on the secondary side of the cooling tower assuming that the input pressure is known and the temperature is equal to atmospheric temperature.

The temperature on the primary exit of the cooling tower is taken to be equal to the atmospheric temperature. This allows for the calculation of the heat transferred from the primary to the secondary stream using (3.13).

$$\dot{Q} = \epsilon(\dot{m}_{in\_p} h_{in\_p} + \dot{m}_{HX} h_{HX} - \dot{m}_{out\_p} h_{out\_p}), \tag{3.13}$$

with: ϵ the cooling tower effectiveness as given by the plants specification sheet, and $\dot{m}_{in\_p}$ determined by the low pressure turbine exit stream with $\dot{m}_{HX}$ given by the fraction of the mass flows of the particular turbines used for the low pressure pre heater. The mass balance is given by equation (3.14).

$$\dot{m}_{out\_p} = \dot{m}_{in\_p} + \dot{m}_{HX} \tag{3.14}$$

3.5.5 Pipes and splits

All pipes are treated as well insulated and therefore the entropy change over the pipe is zero since the heat transfer is zero. A simple set-up of a pipe schematic is shown in Figure 3.10. The pipes

play a large role in the model since it forms the connecting part between the components. For most cases it is assumed that the pipes have no pressure drop or friction losses. In the cases where large pressure drops occur in the physical system, equation (2.32) is used with $0 < \alpha < 1$ a factor accounting for the pressure drop. The mass flow is taken to be constant between the input and output in order to adhere to the law of conservation of mass.

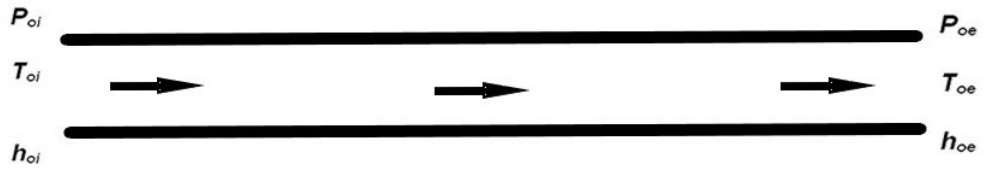


Figure 3.10: Schematic of a pipe as used for the model

In terms of junctions the only changes made to the approach used for the pipes is that the mass balance equation is adapted. The fractional mass flow is determined simply by a scaling constant. This is shown in equation (3.15).

$$\dot{m}_{Split} = \kappa(\dot{m}_{oi}), \tag{3.15}$$

with: $0 < \kappa > 1$ the fraction of mass flow from the inlet that splits. The value of κ is typically determined by the pipe geometry but is seen as a manipulated variable base on the required mass flow rates of the split. The mass balance can therefore be written as shown in equation (3.16).

$$\dot{m}_{oi} = \kappa \dot{m}_{oi} + \dot{m}_{oe} \tag{3.16}$$

A simple schematic of a split is shown in Figure 3.11. The split can either be a fluid flow in or out of a pipe section (Split A) or a component (Split B).

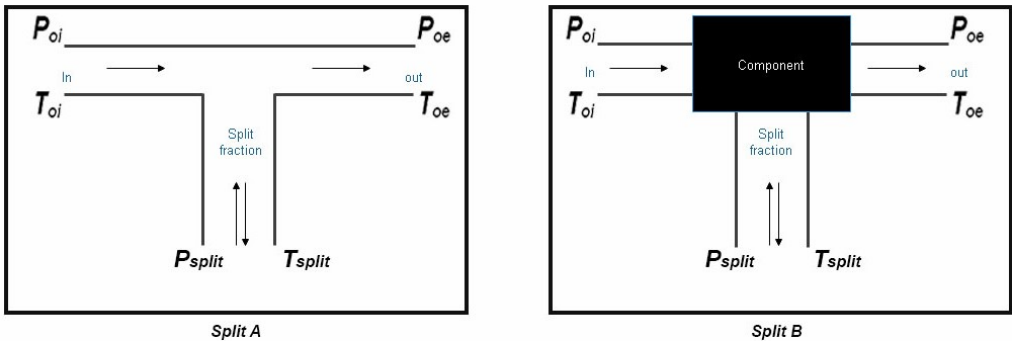


Figure 3.11: Schematic of the two basic split or junction set-ups as used in the model

3.6 System model

The system model is developed by following the schematics of the physical system (Figure 3.12) and applying the aforementioned equations to each component under evaluation.

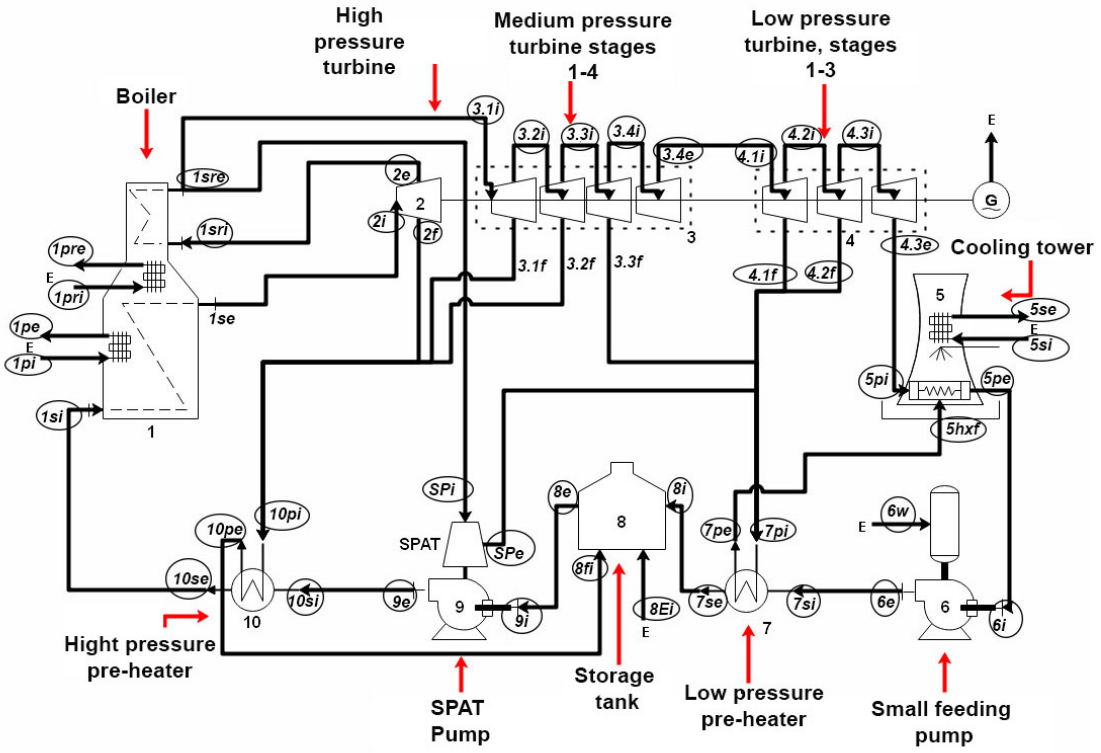


Figure 3.12: Numbered and simplified schematic of the physical system used for the model

As shown in the schematic each component is numbered, starting at the boiler (no 1) and ending at the high-pressure pre-heater (no 10). The identifier at each component's input and output is given by numbering it with the letters "i" and "e" with the fractional fluid stream numbered with a "f". The heat exchanging components like the boiler and cooling tower have some extra letters "p" and "s" indicating the primary and secondary sides. An external input to a component of either work, heat flow or any energy attribute is indicated using a capital "E". Each of the identifiers are included in a circle that encapsulate the fluid stream to which it refers.

In terms of solving the equations of each component, the manipulated inputs to each component are the temperature and pressure. The manipulated variable throughout the whole system is the mass flow rate. In order to explain the process of solving the system equations, consider the cooling tower starting at point $5pe$. The temperature and pressure at this point are taken to be equal to the atmospheric state. Using the equations mentioned for the pipes, the thermodynamic parameters, pressure, temperature, enthalpy, and entropy can be calculated. The mass flow at the exit of the pipe (also inlet of the pump (point $6i$)) is simply taken to be equal to the mass flow at point $5pe$ in order to adhere to the law of conservation of mass. Using the aforementioned equations for the pump, the input and exit thermodynamic parameters can be calculated. Once again, the mass flow is simply taken to be equal to the mass flow at point $6i$. This approach is followed for each component in the system until the loop is closed. This results in a series of equations for which most are balanced in terms of the number of variables and equations. The mass flow equations however, will have one more variable than the number of equations. In order to solve this, an equation is added by composing an energy balance for the turbines and the electrical generator shown in equation (3.17).

$$W_G \times \eta_G = \eta_m(\dot{W}_{HPT} + \sum_{i=1}^{i=4} \dot{W}_{MPT(i)} + \sum_{j=1}^{j=3} \dot{W}_{LPT(j)}), \tag{3.17}$$

with: η_g and η_m the generators efficiency and the turbines mechanical efficiency respectively. The subscripts i and j are indications of each stage of the medium and low-pressure turbines.

The power required from the electrical generator is then determined by the electrical grid. Choosing a value for W_G that represents this grid requirement allows for solving all the equations.

Since all of the equations to solve each component and the thermodynamic cycle is now defined, the next section will focus on the validation of the model.

3.7 Model validation

The model validation is defined as the process of determining how well or accurately the model represents the physical system. This implies that data are required for some point in the system where parameters are calculated. In order to obtain comprehensive and representative data in this regard, the existing validated model in Ebsilon is used. By comparing the model results from the Ebsilon model with the EES model a conclusion on the accuracy of the EES model can be made. According to the system specification document [72], the nominal system electrical power output is 500 MW per block. The system is, however, commonly operating at 525.750 MW as concluded from the staff and physical measurements of the plant. The physical measurements also showed 2 distinctive operating conditions, one at 475 MW and the other at 540 MW. Due to these values the 100% operational condition is taken to be 525.750 MW with the other significant operational points at 90% (475 MW), 95% (500 MW) and 103% (540 MW). The validation is done by running a simulation in both models for these different operational power requirements from the electrical generator and also by changing some component parameters to see if the models still predicts the same system response. One more simulation was ran at 85% to evaluate the model at non-common operational conditions. Two sets of simulations from each model are compared with each other. The first set consists of simulations that are done with the generator power requirement at 100% and 103% respectively with the 100% requirement equal to 525.750 MW. This results in two simulations from both models as shown in Table 3.1. The second set consists of a simulation done at a different efficiency for the HPT. For the first set of simulations, the HPT efficiency was set equal to the physical system specifications. The efficiency was then changed to be 3% less ($\eta_{hpt} = 0,96\%$) than specified.

The results from the two models are then compared in order to validate the EES model. In Table 3.1 only mass flow is compared since the variation in the results of the models is clear enough from only evaluating the mass flow. In Table 3.2 the power results of the HPT, MPT and LPT are compared while also showing the generator power (W_G). The turbine power results for the 100% generator load case (W_G=525.750 MW) is the power corresponding with the thermodynamic results shown in Table 3.1.

A summary of the compared power results, for both sets of simulations, is given in Table 3.2. From the results it is clear that the two models give reasonably similar results with a percentage deviation smaller than 0.7%. Simulations for generator power requirements of 95%, 90% and 85% were also ran. Only results for the 100% and 103% cases are shown since for this power requirement the EES model gave accurate results in comparison with the Ebsilon model.

The deviation of 0.7% increased for the smaller load requirements ($W\_G <$ 95% of 525750 kW) and the larger changes in the efficiency of the high pressure turbine. Beside this fact, even deviations of 0.7% are substantial when considering that the power difference between the models in such a

case is approximately 1 MW. It is noted that the load distribution between the turbines differs for each model run, but are similar for the two models. For the 100% load requirement case 27% of the total load is generated by the HPT while for the 103% load requirement only 25% of the total load is generated by the HPT.

Table 3.1: Mass flow comparison between Ebsilon model and EES model

Generator load at 100%				
	Ebsilon	**EES**	**Difference**	
ID	$\dot{m}$	$\dot{m}$	$\dot{m}$	%
	kg/s	kg/s	kg/s	-
1si	434.5	434.5	0	0.0000%
1se	433.3	433.3	0.04	0.0092%
2i	430.3	430.3	-0.01	-0.0023%
2e	430.3	430.3	-0.01	-0.0023%
1sri	387	387	0.01	0.0026%
1sre	388.1	388.2	0.06	0.0155%
3.1i	365	365.1	0.05	0.0137%
3.2i	352.7	352.7	0.01	0.0028%
3.3i	337.5	337.5	-0.03	-0.0089%
3.4i	324.9	324.9	0.03	0.0092%
3.4e	324.9	324.9	0.03	0.0092%
4.1f	324.2	324.2	0.02	0.0062%
SPe	324.2	324.2	0.02	0.0062%
4.1i	322.9	320.9	-1.96	-0.6107%
4.2i	304.1	302.3	-1.84	-0.6087%
4.3i	293.6	291.8	-1.81	-0.6203%
4.4e	293.6	291.8	-1.81	-0.6203%
5pe	305	303.3	-1.73	-0.5704%

Table 3.2: Model simulation results for different set-ups

Power calculations							
Simulation		Ebsilon		EES		Difference	
Link No	Component	kW	% kW of total	kW	% kW of total	kW	%
100% load,	W_HPT	-144581	27%	-144577	27%	-3	0,002%
WG=525750	W_MPT	-208246	39%	-208612	39%	366	-0,176%
kW	W_LPT	-176319	33%	-175243	33%	-1076	0,610%
	Total	-529146		-528432		-713	0,135%
103% load,	W_HPT	-135408	25%	-134760	25%	-648	0,479%
WG=542000	W_MPT	-228758	42%	-228758	42%	0	0,000%
kW	W_LPT	-185490	34%	-184601	34%	-889	0,479%
	Total	-549656		-548119		-1537	0,280%
HPT	W_HPT	-131345	24%	-131687	24%	342	-0,260%
efficiency,	W_MPT	-229860	42%	-230458	42%	598	-0,260%
3% drop.	W_LPT	-185490	34%	-185974	34%	484	-0,261%
	Total	-546695		-548119		1424	-0,260%

3.8 Conclusion

This chapter explains the methods and approaches used to develop the EES model. An in-depth discussion of the mathematical modelling of each component is given with references to the relevant literature. Each component is initially viewed separately when composing the necessary equations. Each components mathematics is then added systematically to the next component to form the system cycle as shown in the system schematics (Figure 3.3).

The model is validated by comparing the simulation results to that of a second model built in the software package Ebsilon. The model was built by the staff at the plant in Jeanschwalde. The software allows for fundamental principle models. This implies that the model is component based and will solve each of the components physical thermodynamic and performance parameters using mass-moment and energy balance equations. This kind of model allows for empirical correlation and component characteristics to be added in order for real life component simulation.

The results clearly indicate differences between the two models. The EES model does represent the system fairly well but would not be accurate enough for fault emulation for energy applications. From Table 3.2, even a difference as small as 0.17% will result in a difference of 366kW. This will have significant effects when the EGBV for FDI is applied since such an inaccurate result may contribute to a fault signature's size, resulting in wrongfully detecting a fault. Besides this, for some simulation cases, the EES model had differences from the Ebsilon model as large as 10%. The differences seen between the models can be attributed to a number of factors. The most likely and most prominent may be some of the following:

1) The assumptions made regarding no friction in the pipes and valves have large losses that are not accounted for in the model. At the 100% generator load requirement the effects might not be clear but from Table B.1 in appendix B section B.1 the difference in pressure can be seen for the 103% case.

2) The simplification of the actual system layout to obtain the schematic shown in Figure 3.3 results in a large neglecting of junctions and valves within the system. This results in mass flows not being accounted for in some sections of the system. The effects of this can especially be seen in the mass flow results shown in Table 3.1. For example, note the mass flow difference at the inlet of the reheat part of the boiler (link No: "1sre"), the difference in model results equal to 0.06 kg/sec, are due to a junction that was not accounted for. This also accounts for the consequent difference seen for link number "3.1i". The cumulative effects of this may account in part for the large differences seen at link number's "4.1i" and onwards.

3) The distribution of load between the turbines as well as the control that enables the boiler to deliver more or less power than required, is not taken into account in the EES model. This resulted in the model establishing a mathematical balance of power between the components with a pre-specified distribution of loads between the turbines. The Ebsilon model however, has built in limits and rules that accounts for the load distribution and gives more specific control in this regard. The effect being that the Ebsilon model would distribute the load between the turbines in a different matter than the EES model.

If the EES model is adjusted to account for these 3 points it is likely that the model's simulation would be more accurate. Due to the complexity of the physical system this would not be possible within the scope of this study since the model only forms a sub part of the research with the main focus on the energy based FDI. Due to this fact it is decided that the Ebsilon model will be used for fault simulation. Since the model is a validated representative model of the physical system it can easily be used to obtain more data for some fault condition that may occur in the plant but for which there does not exit any practical data. More details on the specifics of using the model are given in chapter 5.

CHAPTER 4

Energy characterisation

This chapter discusses the energy characterisation of the system. A thorough explanation of each aspect of energy characterisation is given. This includes a discussion of the methodology used in the process and the final resulting attributed graph.

4.1 Introduction

Since the model was developed and validated in the previous chapter, and it was concluded that the Ebsilon model will be used for the simulations of faults, the energy and energy attributes in the system can now be obtained for any component. The next objective would therefore be to characterise the system in terms of energy.

This chapter gives a detailed description of what is meant by energy characterisation and energy visualisation. Five guidelines are given that is used to establish the attributed graph.

This is accomplished by giving an overview of the energy graph based visualisation as applicable to this study. The evolution of the energy graph based visualisation at the NWU is shortly elaborated on and, as a final apex, the attributed graph of the system is composed.

4.2 Energy graph based visualisation

The principle of graph based energy visualisation in the context of this research was suggested by G van Schoor, K Uren, H Maraise et al [3], [15], [17]. They suggested an approach to analyse a system based on the energy changes in the system especially focusing on exergy change in components and the energy interactions between the components for FDI applications. These energy parameters are then visualized using the graphing and analytic approaches that modern graph theory [29] consists of. According to [19], the utilisation of energy as a characterisation variable is a method to reduce the dimensionality of the data used for FDI across multi-physical domains. This dimensional reduction is similar to what is achieved using techniques such as kernel PCA, but since energy is a universal concept, the energy characterisation approach is more simplistic in multi domain systems.

The utilizing of energy as a means of characterising a system for FDI applications, is demonstrated in [13] and [2]. In these articles the utilisation of Mollier diagrams, also known as enthalpy

vs. entropy graphs (h-s graphs), are applied in the energy characterisation for fault detection (FD) purposes in a generation IV nuclear high temperature gas-cooled reactor (HTGR). Two classifiers are devised to address the problem. The first classifier is called the error method and classifies faults based on the error seen in the h-s graphs when comparing the graphs under normal and fault conditions. This can be seen in Figure 4.1. In the figure the mathematically ideal cycle is compared to the practical tested cycle and the actual operation cycle by noting the change in entropy between the cases.

The second approach is called the area error method. This method addresses the FD problem via a multiple classifier ensemble approach that realized FD using a classification algorithm that is based on the area and trajectory of the residual shifts between the normal and fault states seen in the h-s graphs [2]. An example of this is shown in Figure 4.2. The energy characterisation in these cases consists of the derivation of energy attributes from the system in order to compose the h-s graphs under normal and fault conditions.

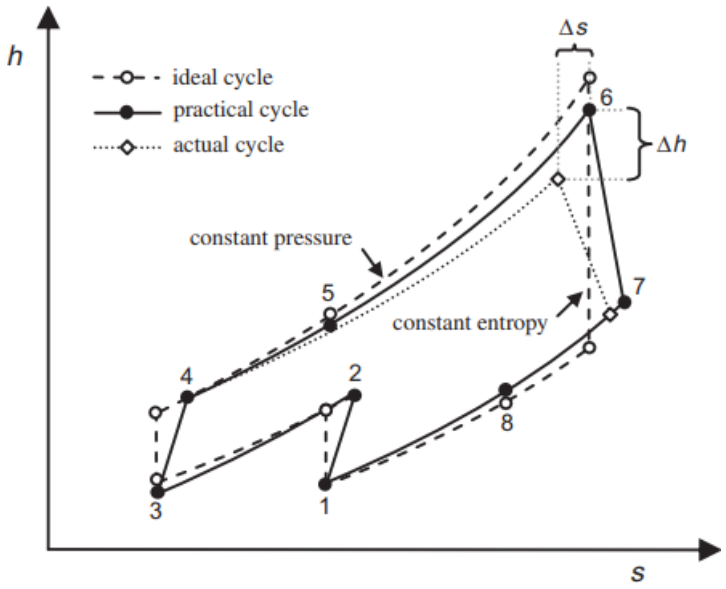


Figure 4.1: h-s graph for the error method in FD applications obtained from [2]

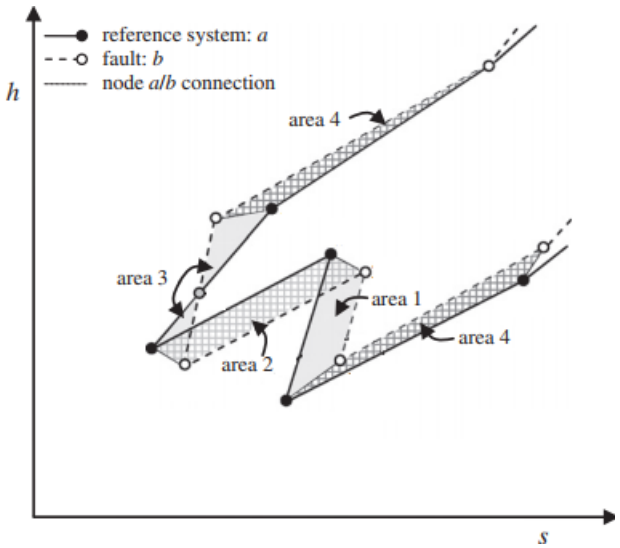


Figure 4.2: h-s graph for the area error method in FD applications obtained from [2]

The approaches by [13] and [2] allows for some structural information to be available for analysis but the analysis complexity can be largely reduced while also increasing the structural information, by rather using graphs consistent with modern graph theory as defined by [29]. This approach is illustrated in [3], [15] and [17]. In order to compose these graphs, the philosophy for energy characterisation of the system must broaden. The characterisation still includes the derivation of energy attributes as with the approach using h-s graphs but the structural layout of the system must also be taken into consideration. To this extent the energy characterisation refers to the calculation of energy attributes in the system, analysing the system structure to determine components of interest that must be represented by nodes (also called vertices) and the final composition of an attributed graph using the energy attributes obtained for each vertex. It is of value to answer the following possible questions when applying the energy characterisation.

1 Where should faults be detected?

2 Which components have many inflows or outflows in terms of energy?

3 Which components have a large effect on the control of the system? Thus, which components will affect the system if a small change in energy occurs at/in these components?

4 Which components have large energy attributed to them? Thus, at which components do large amounts of energy flow into/out-of?

5 How can the energy characterisation be done to compose the least complex graph possible?

An example of such an energy characterisation resulting in an attributed graph for a two tank system is depicted by K.R. Uren, G. van Schoor and L. Auret in [3] and can be seen in Figure 4.3. The figure shows the system schematics noting only the main components (1 to 5 with 5 being the environment) in Figure 4.3 (a). These components are then depicted as vertices being connected to each other by directed edges (represented by an arrow) in the attributed graph shown in Figure 4.3 (b). The direction of each edge indicated the direction of energy flow. The exergy change (Δb_i) over each component is shown next to the vertex, and the edges between the vertices, indicating the energy transfer in terms of heat transfer (q_{ij}), between the components are shown next to the edge. As seen from the attributed graph some components (e.g. valves) in the system diagram are not included while applying the energy characterisation. This is partly due to the consideration of the aforementioned questions.

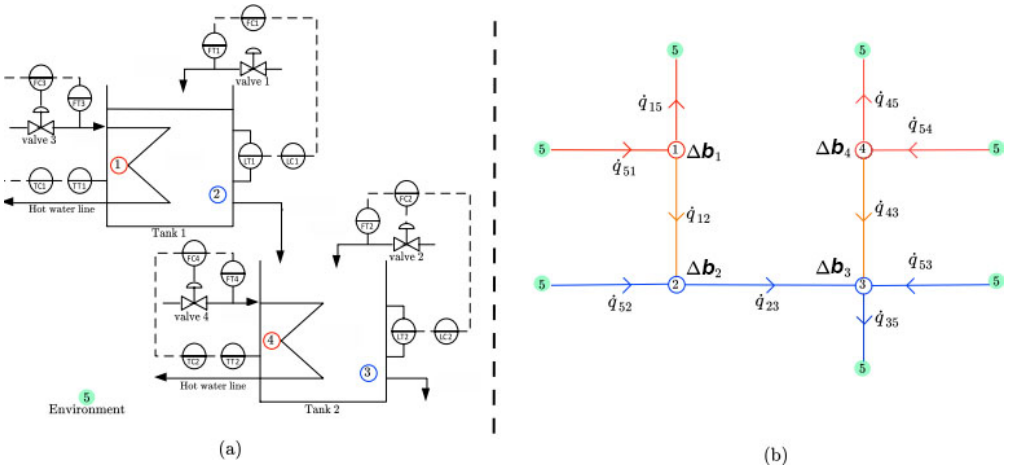


Figure 4.3: a) Two tank system schematics, (b) Energy attributed graph of the system as obtained from [3]

Utilizing this approach, the energy characterisation is applied to the system in Janschewalde. The calculation of energy attributes is done through the use of the Ebsilon model and through analysing the physical system data. The analysis of the system structure is done by considering the five questions. Since the faults that must be detected are attributed to the turbine and the main feeding pump, these two components are identified as necessary components to include in adherence to question 1 (4.2).

Considering that large energy conversion happens within each stage of the turbines (thermodynamic energy converted to mechanical energy), and each stage also has a fractional energy outlet used for the pre-heaters, it can be concluded to include each stage of the HPT, MPT and LPT as a vertex. This will satisfy questions 2 and 4 (4.2) with regards to the turbines.

Considering the rest of the system in terms of question 2 and 4, it is concluded that the boiler, generator, cooling tower, small water pump, low pressure pre-heaters 1 to 4, water storage tank, main pump, spat turbine and the high pressure pre-heaters 5 to 7 must be included as vertices. Each of these components can be classified as the main components in the system, attributing to large energy conversions, relative to the system's capacity. The boiler and cooling tower is indicated using two nodes, each representing the primary and secondary side of the components respectively. This allows for simplistic analysis when calculating the energy attributes since either the primary or secondary side of the component can be analysed independently.

The SPAT is included due to question 3. The main water pump, that is driven by the SPAT, is directly responsible for the mass flow through the system. This in turn affects the power generation of the whole plant. Due to the fact that the SPAT obtains a relatively large amount of energy from the reheated stream of the boiler, the effects that the SPAT has on the system's control and response is notable. The SPAT's energy requirements are in the order of 12MW. This accounts for +- 2.5% of the total plants generation capacity and is also in the same range as the energy required by the pre-heaters (10MW to 70MW).

In order to adhere to question 5 and keep the graph as simple as possible, the pre-heaters 1 to 4 and 5 tot 7 are combined into only four vertices, two for the primary and secondary sides of the 4 low pressure pre-heaters and two for the primary and secondary sides of the 3 high pressure pre-heaters. Since the energy attainment on the primary side is from the environment (see next paragraph for an explanation) the primary side is only indicated by an environmental vertex. In this sense the vertex allocation to these heat exchanging components differs from the boiler and cooling tower. This results in a graph with vertices and edges that correlates to the system schematics shown in Figure 3.3.

Since the graph complexity has a large impact on the mathematical analysis methods required [27], it would be helpful if the graph can be further reduced. One method to obtain this reduction is by neglecting or disconnecting some edges that connect vertices to each other. It is however, important that the energy balance at each vertex is coherent with the system. To accomplish this a general vertex can be added that represents the environment. An energy flow from such a disconnected edge out of the vertex can then be attributed to an energy flow into the environment and vice versa. An example of such an application is seen with the fractional energy flows, out of each of the turbine stages, that is utilized in the low and high pressure pre-heating heat exchangers. The resulting system schematic and the corresponding graph can be seen in Figure 4.4.

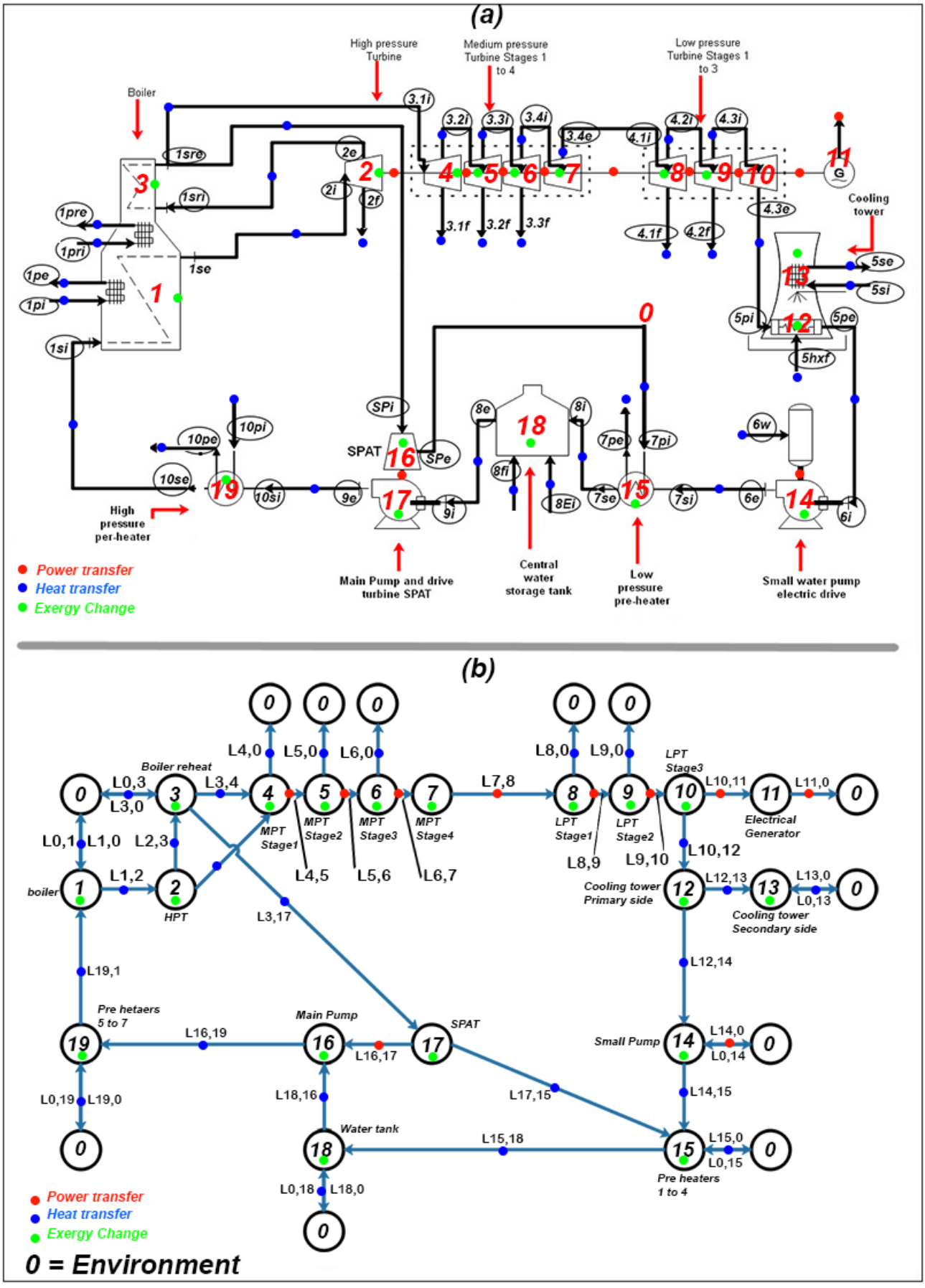


Figure 4.4: a) Reduced system schematic for graph applications (b) Attributed graph based on the reduced system schematic

This attributed graph can then be used in application of the FDI technique, using mathematical analysis methods. The next chapter will discuss how the required energy attributes are assigned to each vertex and edge and how a matrix representation of the attributed graph (the node signature matrix) can be composed.

4.3 Conclusion

As a conclusion it is noted that a short explanation of the term "energy graph based visualisation" was given with some examples of attributed graphs as shown in literature. A discussion of the development of the energy graph based visualisation, in the context of this research, at the NWU was also elaborated on. From relevant literature, five criteria were identified as a guide for choosing the nodes and links between the nodes. These criteria are only a reference and can be expanded if necessary for certain systems. With respect to these five criteria, the STS was analysed and finally the attributed graph in Figure 4.4 (b), was composed. This concluded the second high level objective.

CHAPTER 5

Experimental design and data collection

This chapter discusses the procurement of data through the use of the Ebsilon model and measurements from the physical system in Jeanschwalde. A short description of the fault types is given. The analysis of the data and the composition of time series from the simulation results and system measurements, for the purpose of energy-based graph matching FDI applications, are thoroughly explained.

5.1 Introduction

In chapters 3 and 4 the developments of the static thermodynamic model and the energy characterisation was described. This completed the first two high level objectives. The next and last objective is therefore to acquire the necessary energy based data.

In this chapter the process followed to obtain data from the physical system and the composing of time series data sets will be described. The end goal of the chapter is to obtain a single data set with energy derivative time series data representing each of the fault states, their variations and the normal state.

This is achieved through a clear description of the four fault types and how it were simulated with variations in the faults sizes and locations to obtain simulated system data. The procurement of the physical system data are also described after which the process of combining the simulated and physical data into a time series is discussed. The final result is the data set containing the time series data of each fault state as required.

5.2 Discussion of fault types

As mentioned, there are 4 types of faults being considered for this research. These faults were chosen based on the literature and information given by the staff working at the plant in Germany. As stated by [4] there are seven common faults occurring in a steam turbine system. Of these faults the following three are considered for this study: 1) Solid particle erosion on the high pressure and medium pressure turbine (FT 1); 2) Leakage of the overflow or control valve (FT 2); and 3) Overall ageing and wear of various turbine stages (FT 3). The fourth fault that is considered in the STS is cavitation on the main feeding pump (FT 4). This was chosen based on feedback from the staff at the plant that suggests these fault types are commonly found in the STS. Each of these

fault types are commonly occurring faults in large scale turbine systems, although, to the authors knowledge, these faults were not present in the physical system at the time of this research. Due to the cost, production effects and safety concerns, these faults could not be emulated in the practical system. They were, however, simulated in Ebsilon. A summary of the fault types as well as the size considerations for each fault will be described later. It is important to note, firstly that some of the same faults were simulated in different components. This was done in order to obtain data that could be used to evaluate the robustness of the FDI approach. Secondly, for some fault types the same fault was simulated in the same component but at different intensities (ie, 3%, 4% 6% etc.). This was done to obtain data that could be used to evaluate the sensitivity of the FDI approach. With this process the data obtained from the simulation would enable evaluation of the FDI approach in terms of both detection, isolation, robustness, sensitivity and false detection/isolation rate. With regards to each fault the description and effects caused by each fault in the STS are focussed on in the next subsections.

5.2.1 Solid-particle erosion on the high and medium pressure turbine stages

Solid-particle erosion refers to the phenomena where the impact of particles in the steam path of the first stages of the high and medium pressure turbine causes damage to the blades. The presence of these particles is commonly attributed to the exfoliation of iron oxide (rust) and magnetic particles from the high temperature part of the boiler [73]. According to [4, 23] the consequences of this erosion are an increase in the swallowing capacity of the turbine and a drop in the efficiency of the turbine stage. This can typically be modelled with a drop in turbine stage efficiency of 1% to 10% as well as an increase in mass flow rate through the first turbine stages. An example of Solid Particle Erosion (SPE) on a turbine blade can be seen in Figure 5.1.

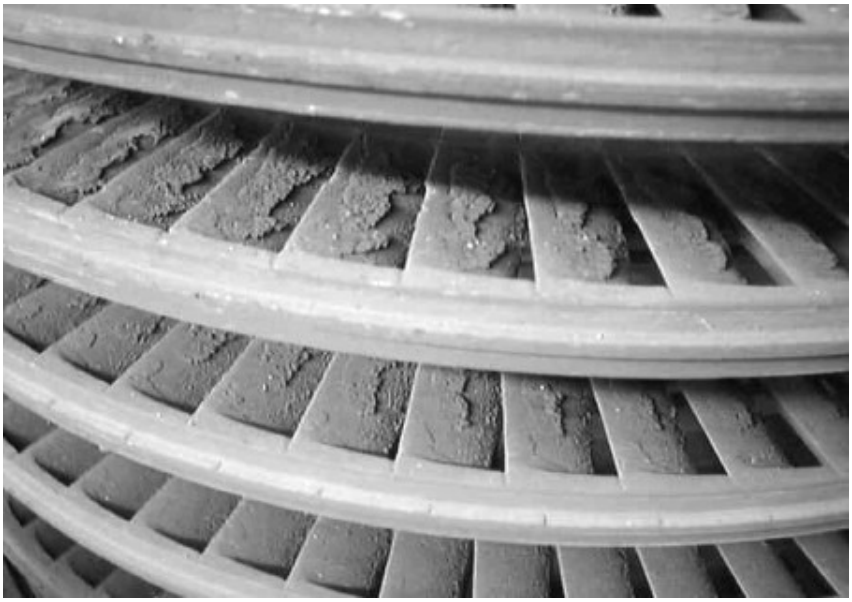


Figure 5.1: Solid particle erosion on a turbine blade [4]

5.2.2 Leakage of the overflow valve

The overflow valve is a de-super heating and decompression system in parallel with the turbine. The valve is used for bypassing steam from the turbine in part or as a whole. The function of the overflow valve as a bypassing system is to control the imbalance between the steam consumption and production of the turbine and boiler respectively [74]. Leakage of the overflow valve can be due to either a broken spindle or SPE [4]. This leads to high quality steam bypassing the turbine

stage to condensate in the turbine's leakage condenser or to enter the turbine at a later stage with lower pressure. The consequence of this fault is a drop in the turbine performance as well as in outlet temperature.

5.2.3 Overall ageing and wear of various turbine stages

Ageing and wear in the turbines are mostly noticeable in the increased surface roughness and degradation of the mechanical components. According to [4] the effects of ageing and wear is a drop in overall performance of the turbines and STS as a whole, and can be detected as a uniform degradation in performance. Temperature gradients, SPE and low-quality steam are al effects that contribute to ageing and wear and, in such, overall ageing and wear can be seen as the result of a number of undesirable effects in the system.

5.2.4 Cavitation in the main water feeding pump

The phenomenon called cavitation is described by [75] as the formation of a vapour cavity that moves along with the fluid stream when the pressure of the fluid falls below the vapour pressure of that fluid at some position in a component such as a pump. The cavity that forms contains a swirling mass of droplets and vapour that forms and collapses numerous times in a second. When this low-pressure cavity moves downstream to a higher-pressure region it suddenly collapses, resulting in a momentary localized pressure increase of significant size. When this happens close to, or against the boundary of the component, the component's wall receives a blow, stressing the wall surface, which in some cases may be beyond its elastic limit. A continuous occurrence of this eventually results in fatigue and failure of the component. According to [5] the range over which cavitation occurs in a pump is at the point where the NPSH (measure of pump pressure above vapour pressure) undergoes a significant drop, usually 3% and upwards. Thus, if the pump pressure is 3% or more below the vapour pressure, cavitation will occur. A typical image of cavitation in a blade as given by [5] is shown in Figure 5.2.

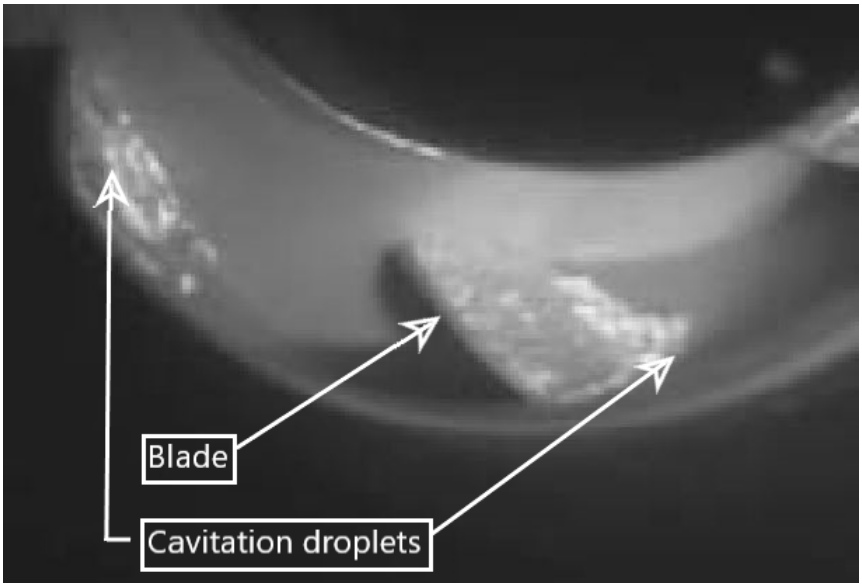


Figure 5.2: Cavitation bubbles in a pump forming on the impeller blades [5]

5.3 Ebsilon modelling of faults

The simulation of each fault was implemented by changing specific parameters in the Ebsilon model or by adding some valves that could be opened to simulate a small leak in the case of FT 2. All of the physical changes that were made in Ebsilon, can be found in appendix B. A typical example of a simulation result is shown in Figure 5.3. From the figure the steady state temperature, pressure, mass flow, heat flow and exergy calculated at a specific position in the system are shown. Due to the size of the system the results is not visible without zooming to the component for which the results is considered. To illustrate the typical results and layout, a zoomed partition is indicated on the figure.

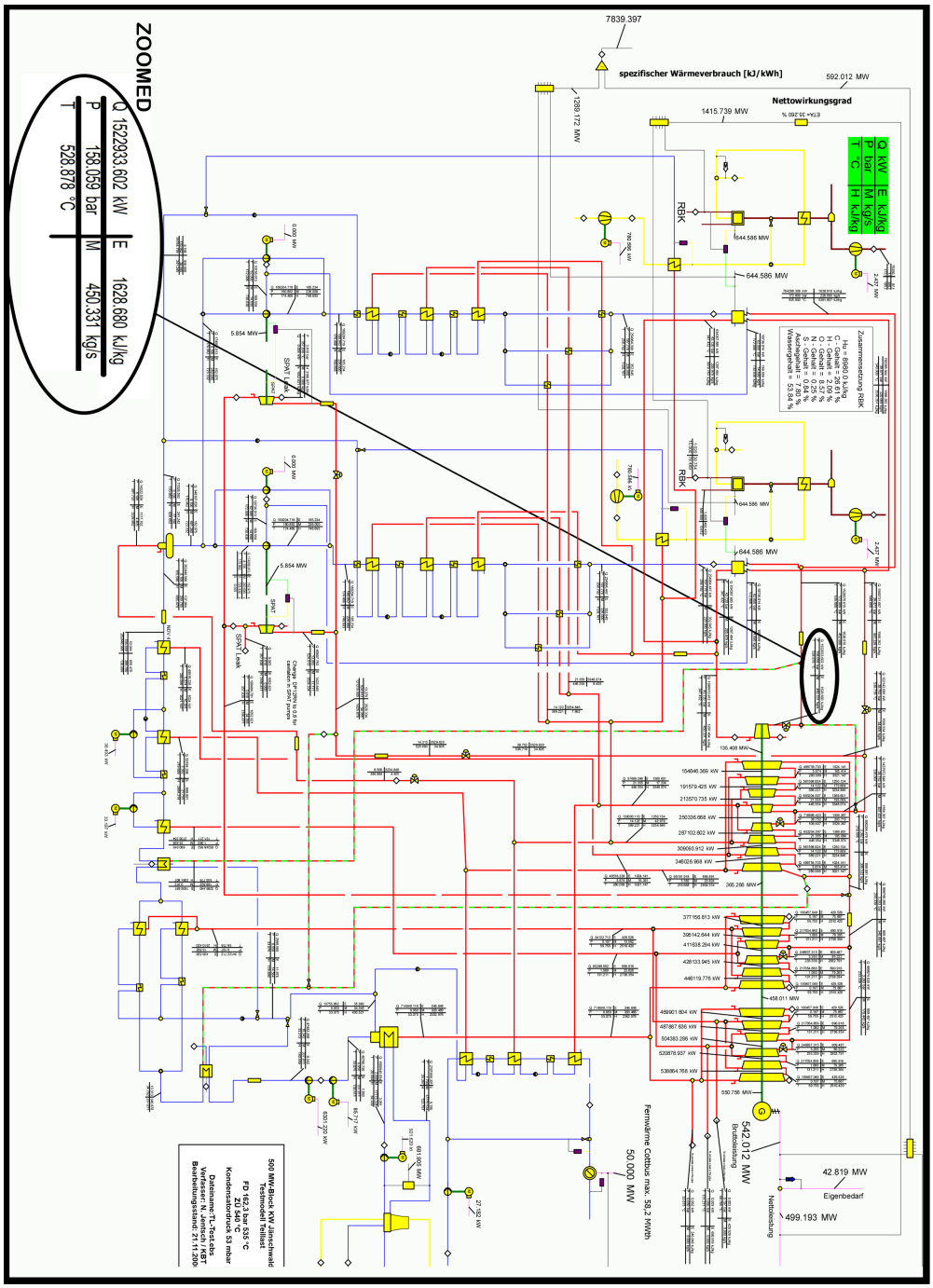


Figure 5.3: Example of a simulation result obtained with Ebsilon

With reference to Table 5.1, the effects of solid particle erosion was only simulated on the first stages of the high and medium pressure turbines. This is due to the fact that only these turbines are directly connected to the boiler, hence it was most likely to find SPE in these stages. The simulation entailed that the mechanical efficiency of the HPT and first stage of the MPT were dropped with 3% while also increasing the mass-flow rate with 3% in order to simulate an increase in suction capacity. This simulation was repeated for the first stage of the MPT with a drop in efficiency of 6% and an increase in mass flow rate of 6%. In terms of the mathematical effect this would have on the component, refer to (2.16).

Table 5.1: FT 1- Solid particle eroion

Designation	Description
FT1 _ HPT3%	High-pressure turbine mechanical efficiency dropped with 3% and mass-flow rate increased with 3%
FT1 _ MPT3%	Medium pressure turbine's first stage mechanical efficiency dropped with 3% and mass-flow rate increased with 3%
FT1 _ MPT6%	Medium pressure turbine's first stage mechanical efficiency dropped with 6% and mass-flow rate increased with 6%

The second FT as summarised in Table 5.2 was the leakage of the overflow or control valve. This fault was simulated on all three stages of the low-pressure turbine (LPT) and the SPAT. A loss in mass flow of 0.5 kg/s was simulated for each of the stages of the LPT and also for one simulation of the SPAT. The simulation was then repeated for 1 kg/s, 2 kg/s and 5 kg/s of mass flow loss in the SPAT. This is accomplished by simply specifying the mass flow rate that should exit the overflow valve. This would cause a lower mass flow rate being available to the following stages of the turbine. The result would be a decrease in power generation by the turbine stage according to (2.16). Considering the system layout (see Figure 3.3), it can be seen that each of the medium and low-pressure turbine stages has a split or fractional outflow. These outflows supply energy in the form of hot fluid to the pre-heaters. A leak in the overflow valve of the turbine will change the mass flow rate of the fluid exiting the turbine stages and the fractional outflows that are used to supply energy to the pre-heaters. This will cause the pre-heater to have less energy available to preheat the fluid stream and result in an overall temperature drop in the system (see (2.34)).

Table 5.2: FT2- Leakage of the overflow valve

Designation	Description
FT2_LPT1	0.5 kg/s Leak on the overflow valve of the low-pressure turbine's first stage
FT2_LPT2	0.5 kg/s Leak on the overflow valve of the low-pressure turbine's second stage
FT2_LPT3	0.5 kg/s Leak on the overflow valve of the low-pressure turbine's third stage
FT2_0,5SPAT	0.5 kg/s Leak on the overflow valve of the SPAT
FT2_1SPAT	1 kg/s Leak on the overflow valve of the SPAT
FT2_2SPAT	2 kg/s Leak on the overflow valve of the SPAT
FT2_5SPAT	5 kg/s Leak on the overflow valve of the SPAT

FT 3 consisted of 5 different simulations done to simulate the effects of overall wear and ageing. These simulations are summarised in Table 5.3. The specific changes that were made are listed in appendix B, subsection B.2.3. Since overall wear and ageing on the turbines were simulated, the decrease in efficiency similar to FT1 was implemented, but simultaneously on each 3 stages of the turbine under consideration. The "per-component" effect of this fault will therefore be similar to FT1 but will have different effects on a system level. Due to the fact that the simulation gives a steady state result, it was not possible to simulate a gradual introduction of wear. This implies that the simulation results should rather be evaluated as the results that could be expected as some stage of the turbine's operational lifetime when wear and ageing already occurred.

Table 5.3: FT3- Overall wear and ageing

Designation	Description
FT3_LPT3%	3% drop in efficiency of all the low-pressure turbine stages
FT3_LPT4%	4% drop in efficiency of all the low-pressure turbine stages
FT3_LPT6%	6% drop in efficiency of all the low-pressure turbine stages
FT3_MPT3%	3% drop in efficiency of all the medium pressure turbine stages
FT3_MPT4%	4% drop in efficiency of all the medium pressure turbine stages

The last FT was cavitation in the main feeding pump driven by the SPAT. The simulation in Ebsilon was set-up to simulate a condition or state in the system under which cavitation would occur rather than simulating the effects of cavitation on the fluid. This way the system parameters and energy state under conditions of cavitation would be obtained. This was implemented by utilizing the ability of the "piping" component in Ebsilon to specify a pressure drop that would result in the inlet fluid properties of the pump to be of such that cavitation would occur. The settings of the piping component can be seen in appendix B, subsection B.2.4. It is important to note that the piping component was added before the pump, this created the required fluid properties. The simulation was completed for two cases as shown in summary in Table 5.4. In each case the pressure drop of the piping component was specified to cause a pressure value at the inlet

of the pump that is a specified percentage lower than the NPSH.

Table 5.4: FT4- Cavitation in the main pump

Designation	Description
FT4_6%NPSH	Pressure drop such that the fluid pressure is 6% below the vapour pressure
FT4_9%NPSH	Pressure drop such that the fluid pressure is 9% below the vapour pressure

5.4 Data procurement from the physical system

As mentioned in the previous chapters the data requirements from the system had to be either in the form of energy attribute variables or variables like temperature, pressure and mass flow (that are measured in the system) from which energy attributes can be derived. Through examination of the system diagram and the attributed graph shown in Figure 4.4, the locations where the variables had to be measured were identified. Figure 5.4 gives the system schematic with the locations of the measurement points indicated. With these locations identified an operator at the facility in Jeanschwalde then extracted the data from the monitoring system at the plant.

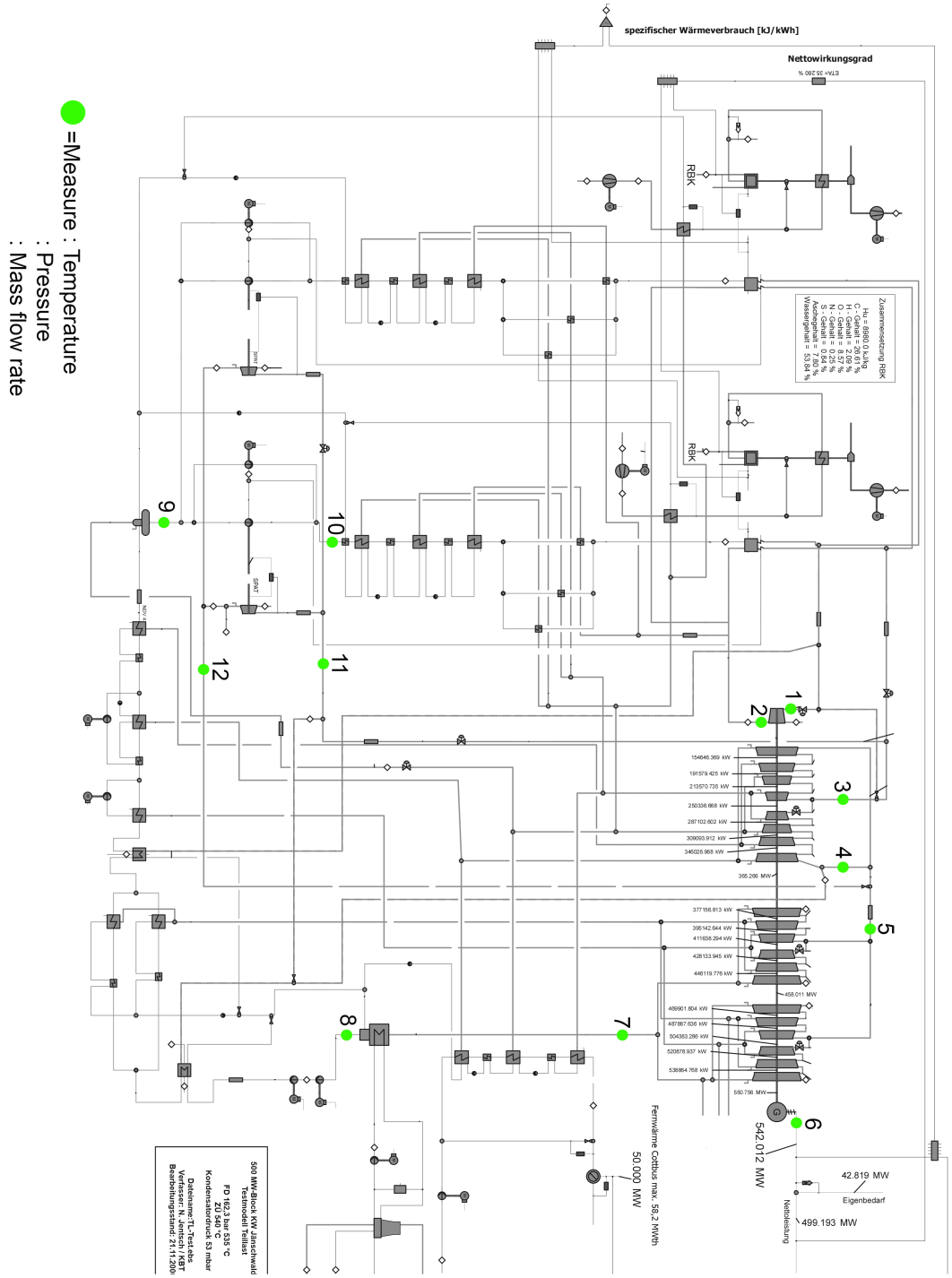


Figure 5.4: System schematic with the locations indicated where measurements were taken

The data were obtained in two different sets, each containing the variables, temperature, pressure and mass flow. The first set of data were measured for nodes 1 to 10 as indicated in Figure 5.4 with the green dots. Unfortunately, due to a lack of physical instrumentation the data for nodes 4, 5 and 9, (Figure 5.4), were not measured. This implied that physical data for the outlet of the MPT, inlet of the LPT and inlet of the SPAT pump could not be obtained. The data from the remaining 7 nodes will be referred to as the "system data". The second set of data were measured for nodes 11 and 12 as shown in Figure 5.4. This data set will be referred to as the SPAT data. A summary of the nodes that measurements was obtained for and the variables measured is shown in Table 5.5. The system data were obtained by taking 96 measurements at 15-minute intervals during two distinct operating points. The transient operation was neglected and only the data from operational state 2 were used since only the variations and noise of the system were of interest in order to obtain some indication of the extend by which a specific variable (like pressure, temperature, mass flow rate or power) would vary under steady operations. An example of this can be seen in Figure 5.5 for the generator output power. The two operating points are clearly visible. This resulted in one 24 hour day of operational data. The SPAT data were obtained by taking 1920 measurements at 45-second intervals. This also resulted in a 24-hour day of operational data. In Table 5.6 a general symbolic illustration of the data for the two sets can be seen.

Table 5.5: Summary of nodes for which physical measurements were obtained

Node:	**Description:**	**Variables measured:**
1	High pressure turbine inlet	Pressure, Temperature, Mass flow rate
2	High pressure turbine outlet	Pressure, Temperature, Mass flow rate
3	Medium pressure turbine inlet	Pressure, Temperature, Mass flow rate
4	Medium pressure turbine outlet	NOT Measured
5	Low pressure turbine inlet	NOT Measured
6	Electrical Generator	Power output
7	Low pressure turbine outlet	Pressure, Temperature, Mass flow rate
8	Condenser outlet	Pressure, Temperature, Mass flow rate
9	SPAT pump inlet	Pressure, Temperature, Mass flow rate
10	SPAT pump outlet	Pressure, Temperature, Mass flow rate
11	SPAT turbine inlet	Pressure, Temperature, Mass flow rate
12	SPAT turbine outlet	Pressure, Temperature, Mass flow rate

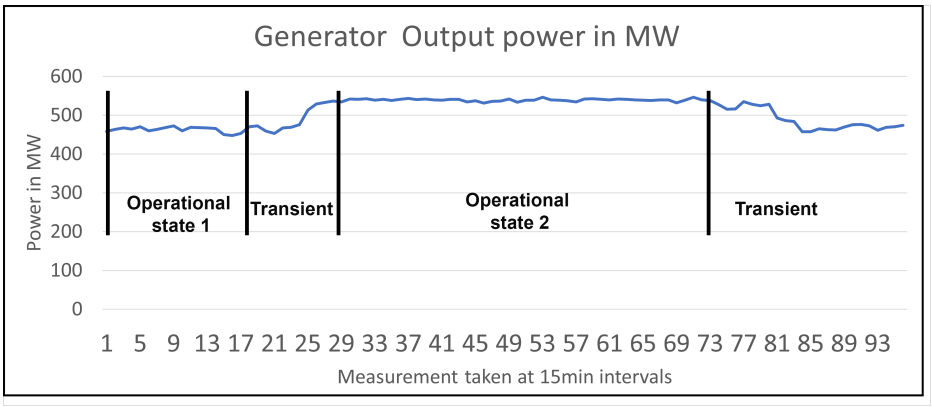


Figure 5.5: Physical data measured for the generator output.

In order to allow for a number of data points that represents a 24 hour operational period, both the system data set and the SPAT data set were artificially expanded to obtain a total number of

Table 5.6: Example of the general symbolic data sets obtained from the plant in Jeanschwalde

Data measurements taken in 15 min intervals							**Data measurements taken in 45 sec intervals**					
System data (nodes 1 to 10)							**SPAT data (nodes 11 and 12)**					
Node1			...	Node10			Node 11			Node 12		
T [°C]	P [kPa]	$\dot{m}$ [kg/s]	...	T [°C]	P [kPa]	$\dot{m}$ [kg/s]	T [°C]	P [kPa]	$\dot{m}$ [kg/s]	T [°C]	P [kPa]	$\dot{m}$ [kg/s]
T_1	P_1	$\dot{m}_1$	...	T_1	P_1	$\dot{m}_1$	T_1	P_1	$\dot{m}_1$	T_1	P_1	$\dot{m}_1$
.	.	.	...	.	.	.	.	.	.	.	.	.
.	.	.	...	.	.	.	.	.	.	.	.	.
.	.	.	...	.	.	.	.	.	.	.	.	.
T_{96}	P_{96}	$\dot{m}_{96}$	...	T_{96}	P_{96}	$\dot{m}_{96}$	T_{1920}	P_{1920}	$\dot{m}_{1920}$	T_{1920}	P_{1920}	$\dot{m}_{1920}$

2880 measurements (equivalent to 30 second intervals between each measurement). This was accomplished by "randomly" selecting data for each variable from the existing dataset of that variable. The method used, was to write a function in Matlab that takes the distribution of the measured dataset (the input data) that must be expanded, and map that distribution to a vector with data points equal to the length of the required dataset (i.e. 2880) minus the length of the original datasets. This resulted in a vector (the output vector) that has the distribution of the input dataset and a length equal to the missing number of data points ($2880 - lengt(inputdata)$). The function then mapped the values of the original measured dataset using the "rand" function to this output vector and then added this vector to the original measured dataset. The resulting output vector was a vector with dimensions 1 x 2880 that has the same distribution as the original dataset with the minimum and maximum values of this new dataset equal to the minimum and maximum values of the original measured datasets. For example, the temperature values measured for node 1 in Table 5.6 (Column 1) were expanded by selecting $(2880 - 96) = 2784$ measurements from that same column to generate a total number of temperature measurements for node 1 equal to 2880. Similarly, $(2880 - 1920) = 960$ random measurements with the same distribution as the original SPAT dataset were chosen from the respective columns for the variables in the SPAT dataset. A representation of the resulting data set can be seen in Table 5.7.

Table 5.7: Symbolic example of the data sets after it were artificially expanded

Data measurements taken in 15 min intervals							**Data measurements taken in 45 sec intervals**					
System data (nodes 1 to 10)							**SPAT data (nodes 11 and 12)**					
Node1			...	Node10			Node 11			Node 12		
T [°C]	P [kPa]	$\dot{m}$ [kg/s]	...	T [°C]	P [kPa]	$\dot{m}$ [kg/s]	T [°C]	P [kPa]	$\dot{m}$ [kg/s]	T [°C]	P [kPa]	$\dot{m}$ [kg/s]
T_1	P_1	$\dot{m}_1$	...	T_1	P_1	$\dot{m}_1$	T_1	P_1	$\dot{m}_1$	T_1	P_1	$\dot{m}_1$
.	.	.	...	.	.	.	.	.	.	.	.	.
.	.	.	...	.	.	.	.	.	.	.	.	.
T_{96}	P_{96}	$\dot{m}_{96}$	...	T_{96}	P_{96}	$\dot{m}_{96}$	T_{1920}	P_{1920}	$\dot{m}_{1920}$	T_{1920}	P_{1920}	$\dot{m}_{1920}$
T_{rand}	P_{rand}	$\dot{m}_{rand}$	...	T_{rand}	P_{rand}	$\dot{m}_{rand}$	T_{rand}	P_{rand}	$\dot{m}_{rand}$	T_{rand}	P_{rand}	$\dot{m}_{rand}$
.	.	.	...	.	.	.	.	.	.	.	.	.
.	.	.	...	.	.	.	.	.	.	.	.	.
T_{2880}	P_{2880}	$\dot{m}_{2880}$	...	T_{2880}	P_{2880}	$\dot{m}_{2880}$	T_{2880}	P_{2880}	$\dot{m}_{2880}$	T_{2880}	P_{2880}	$\dot{m}_{2880}$

From these data sets and the simulation results, a set of time series could be composed that gave an approximation of how the variables at the respective locations in the system would change over time during real live conditions for each of the simulated states (i.e. Normal operation, FT 1, FT 2 etc.). The next section will describe the process of obtaining the time series data sets and the assumptions made in cases where data from the plant in Jeanschwalde could not be obtained.

5.5 Data analysis and composing of time series

To be able to evaluate the systems health over time it was necessary to know what the fluctuations due to the noise and control variations in the system might be. A method to obtain data that would allow this, would be to compose time series data by combining the simulation results from Ebsilon and the data sets obtained from the plant. Through extraction of the noise in the system and SPAT data, and then combining this noise data of each variable with the simulation result, a time series data set for both the normal and faulty simulations are obtained. The flow diagram in Figure 5.6 illustrates the basic idea.

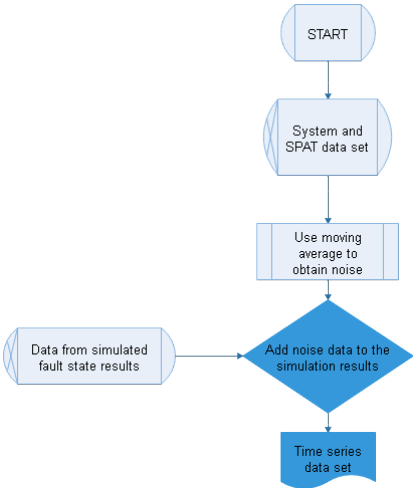


Figure 5.6: Flow diagram showing how the time series data sets are composed

To determine the window size for the moving average a simple analysis of the original data was done via the help of a Matlab program. The program imported each of the data sets at a time and then calculated the moving average of the data while sweeping through a specified window size range from 2 to 100. After each loop, the sum of the absolute difference (SAD) between the smoothed dataset and the original dataset was calculated and stored in a variable. Once the sweep was completed the resulting values of the SAD were plotted against the window size. This resulted in a graph that illustrated the error between the smoothed and original data as the window size increased. For all of the measurements, a knee was found in-between a window size of 7 to 10. This implied that a window size larger than 10 would be inaccurate and hence 10 was chosen as the maximum window size. In figure 5.7 the plot for the SPAT outlet temperature is shown. To illustrate as an example how the process in Figure 5.6 was applied, consider the system data set and select the temperature measurements taken at the inlet of the HPT (i.e. the temperature at node 1 in Figure 5.4). By subtracting the 10-point moving average ((2.45), with $(2N+1) = 10$) from the original data set, the noise present in the temperature measurements can be obtained. These results will be referred to as the noise data set. Next, consider each of the simulation results as obtained for each variation of the four fault states and one normal operation. Select the temperature value calculated at the inlet of the HPT for each of the states simulated and add this value to each of the noise values in the noise data set. The resulting data sets are the generated time series for the temperature at

the inlet of the HPT for each of the simulated states. This is then repeated for each of the variables of all of the components that were identified as necessary for the attributed graph as illustrated in 4.4. The flow diagram in Figure 5.8 illustrates this example.

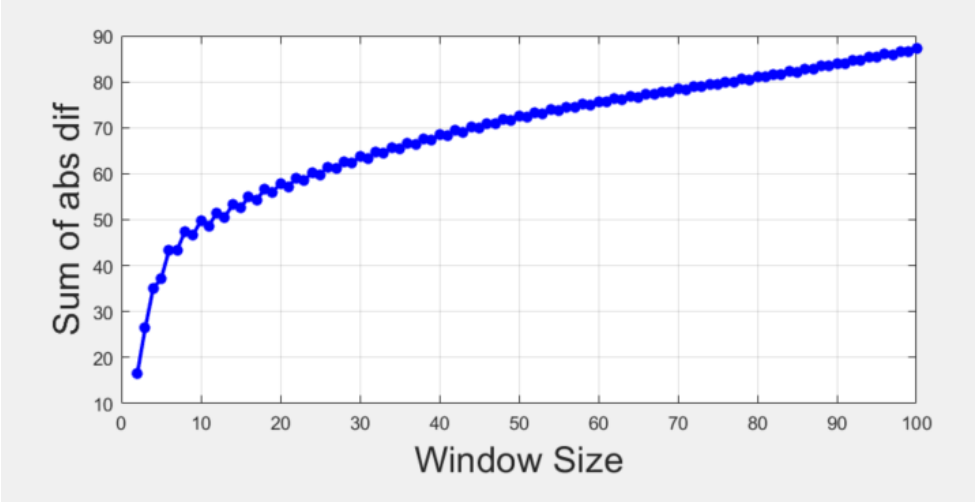


Figure 5.7: Example of why a window of 10 was chosen for the moving average filter.

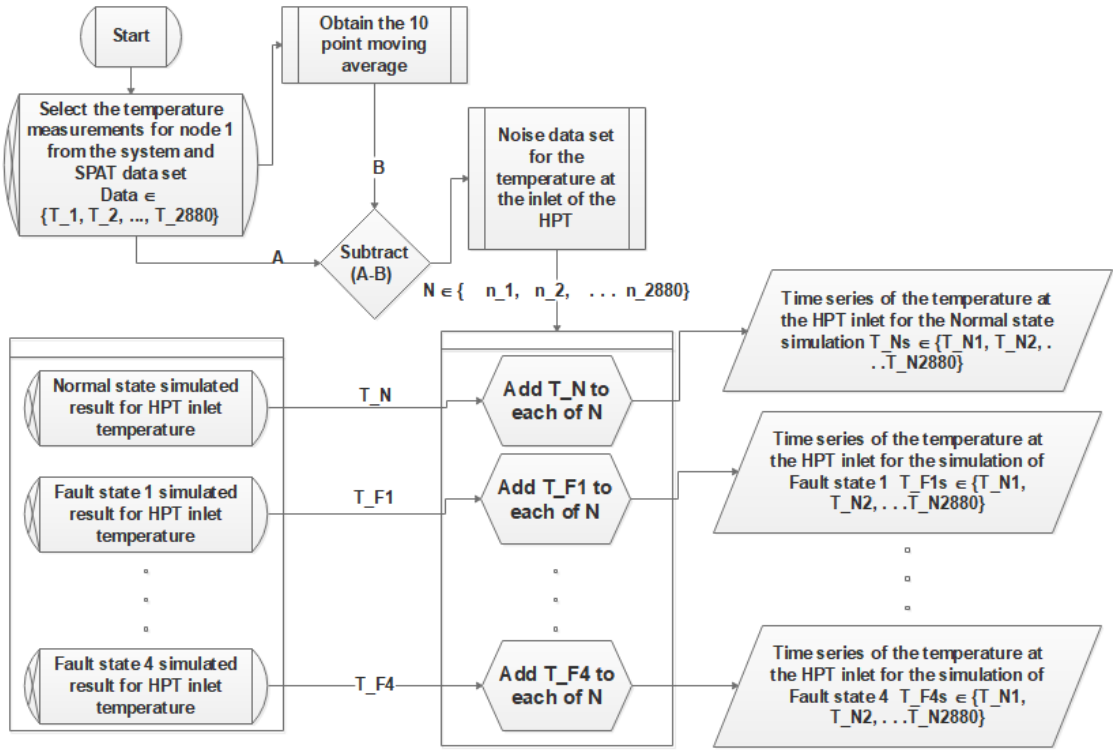


Figure 5.8: Flow diagram illustrating how the time series was generated for the temperature variable example

Using Ebsilon, simulated temperature, pressure and mass flow rate, results were obtained for the input and output of each of the components as represented by the nodes 1 to 19 in the attributed graph (Figure 4.4). However, there is only noise data for the inputs and outputs of some components (nodes 1 to 3, 6 to 8 and 10 to 12) in Figure 5.4. This necessitates the generation of some noise data to be used when compositing the time series data sets for the remaining variables in the attributed graph.

In order to obtain representative noise, it was decided to use the average noise of the components that enclosed the component for which the noise data were missing. For example, the noise data at the inlet of the MPT (node 3, Figure 5.4) and the outlet of the LPT (node 7, Figure 5.4) were known but there were no noise data for any of the consecutive stages of the MPT or LPT which are represented by attributed nodes 5,6,7,8 and 9 in Figure 4.4. In order to obtain representative noise data for these nodes, the average noise between the inlet of the MPT and the outlet of the LPT were used. Figure 5.9 gives a visual illustration of this approach.

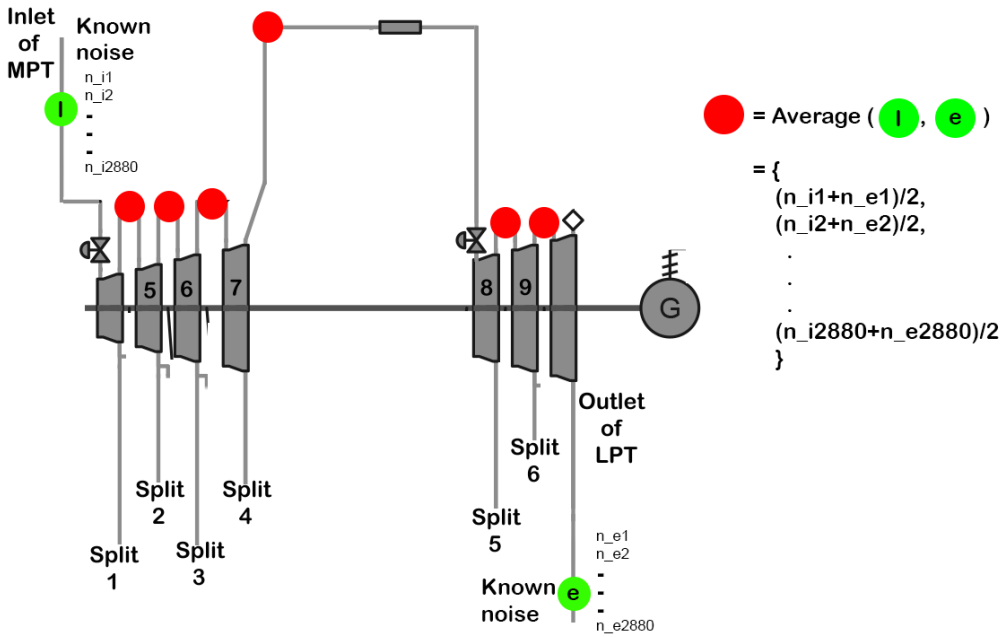


Figure 5.9: Generating the noise data for the unknown (red) nodes by using the known noise data (green nodes)

The resulting data set obtained from the approach to generate the time series for each of the variables (temperature, pressure and mass flow rate) and each of the simulated states (Normal operation, Fault state 1, . . ., Fault state 4) consisted of 18 tables of data each with 2880 measurements (rows) of temperature, pressure and mass flow rates for the input and output of each of the 19 attributed nodes in Figure 4.4 (thus 38 columns). This is depicted in Figure 5.10.

Note that each of the heat flow connections between the components that is represented by links in the attributed graph, are numbered in Figure 4.4 (b). Rather than referring to the input or output of a node it will be referred to by the link number (i.e. L1.2 , L2.3, ... , L19.1).

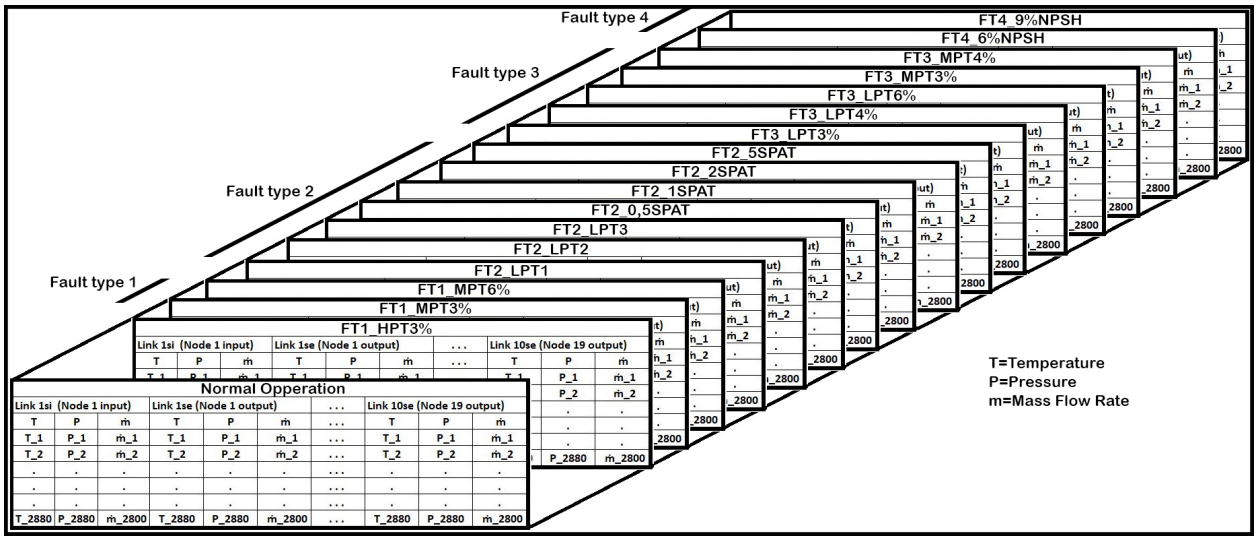


Figure 5.10: Illustration of the final time series data obtained for each of the simulated fault states and the normal operating state

Once the time series dataset was obtained the data analysis could be done to generate energy data sets that would be applicable for the EGBV approach for FDI. The next section will specifically focus on this.

5.6 Generation of energy graph-based fault detection and isolation applicable data sets

As discussed in chapter 2, section 2.4 the energy attributes that will be considered are heat and power transfer between components as well as change in exergy over the components. To obtain these energy attributes a simple function in Matlab was written. The function takes as input, the pressure and temperature time series data from the previous section and, through the use of the lookup IAPWS water and steam tables, determines the specific entropy and enthalpy that correlates with those inputs. The function can be mathematically expressed as seen in (5.1) with the input variable $y1$ and $y2$ the given time series variables and S the variable specifying which thermodynamic variable needs to be allocated in the lookup tables (i.e. enthalpy entropy etc.) . The resulting specific entropy and enthalpy values (the specific energy time series data sets) could be used to calculate the heat transfer between the components and change in exergy over the components.

$$X(y1, y2, S) = lookup(IAPWS, a = y1, b = y2) \tag{5.1}$$

First the specific heat transfer to each of the components were calculated, where applicable, through the use of (2.3). This was done in accordance to the approaches described in chapter 3, depending on the component under evaluation. The resulting specific heat was then simply multiplied by the corresponding mass flow rates to obtain the heat transfer in kilowatts. A symbolic example of the resulting heat flow dataset is shown in Table 5.8. Note that the headers $L1.2, L2.3, ..., L0.19$ refer to the links as indicated in the attributed graph (Figure 4.4). Also note that data sets like these, exist for each of the variations on the fault states (tables 5.1, 5.2, 5.3 and 5.4) and the normal operating state. Since the mechanical work done by the turbines in the physical system were not measured, their power were simply taken as steady (noise and variation free) values equal to the results of the simulation.

Table 5.8: Symbolic example of the heat transfer time series data set to be used in the attributed matrices for FDI

Heat transfer (Q) in [kW]											
L1.2	L2.3	L2.4	L3.4	L3.16	L4.5	L4.0	L5.6	L5.0	L6.7	L6.0	L7.8	L8.9
Q1.2,1	Q2.3,1	Q2.4,1	Q3.4,1	Q3.16,1	Q4.5,1	Q4,1	Q5.6,1	Q5,1	Q6.7,1	Q6,1	Q7.8,1	Q8.9,1
Q1.2,2	Q2.3,2	Q2.4,2	Q3.4,2	Q3.16,2	Q4.5,2	Q4,2	Q5.6,2	Q5,2	Q6.7,2	Q6,2	Q7.8,2	Q8.9,2
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
Q1.2,2880	Q2.3,2880	Q2.4,2880	Q3.4,2880	Q3.16,2880	Q4.5,2880	Q4,2880	Q5.6,2880	Q5,2880	Q6.7,2880	Q6,2880	Q7.8,2880	Q8.9,2880
Data Continue												
L8.0	L9.10	L9.0	L10.11	L11.0	L10.12	L12.13	L12.14	L13.0	L14.15	L15.15	L15.0	L16.17
Q8,1	Q9.1,1	Q9,1	Q10.11,1	Q11,1	Q10.12,1	Q12.13,1	Q12.14,1	Q13,1	Q14.15,1	Q15.15,1	Q15,1	Q16.17,1
Q8,2	Q9.1,2	Q9,2	Q10.11,2	Q11,2	Q10.12,2	Q12.13,2	Q12.14,2	Q13,2	Q14.15,2	Q15.15,2	Q15,2	Q16.17,2
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
Q8,2880	Q9.1,2880	Q9,2880	Q10.11,2880	Q11,2880	Q10.12,2880	Q12.13,2880	Q12.14,2880	Q13,2880	Q14.15,2880	Q15.15,2880	Q15,2880	Q16.17,2880
Data Continue												
L16.0	L17.19	L18.17	L18.0	L19.1	L19.0	L0.1	L0.3	L0.13	L0.14	L0.15	L0.18	L0.19
Q16,1	Q17.19,1	Q18.17,1	Q18,1	Q19.1,1	Q19,1	Q0.1,1	Q0.3,1	Q0.13,1	Q0.14,1	Q0.15,1	Q0.18,1	Q0.19,1
Q16,2	Q17.19,2	Q18.17,2	Q18,2	Q19.1,2	Q19,2	Q0.1,2	Q0.3,2	Q0.13,2	Q0.14,2	Q0.15,2	Q0.18,2	Q0.19,2
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
Q16,2880	Q17.19,2880	Q18.17,2880	Q18,2880	Q19.1,2880	Q19,2880	Q0.1,2880	Q0.3,2880	Q0.13,2880	Q0.14,2880	Q0.15,2880	Q0.18,2880	Q0.19,2880

The same approach was followed to obtain a time series in terms of the exergy change over each of the components. The dead state variables (T_0, s_0 and h_0) were chosen to be equal to the average values in Jeanschwalde of $T_0 = 15$ C° and $P_0 = 1.015$ kPa. From this pressure and temperature values the enthalpy (h_0) and entropy (s_0) were calculated. In Figure 5.11 a generic example of how this calculations were applied to a generic component is given.

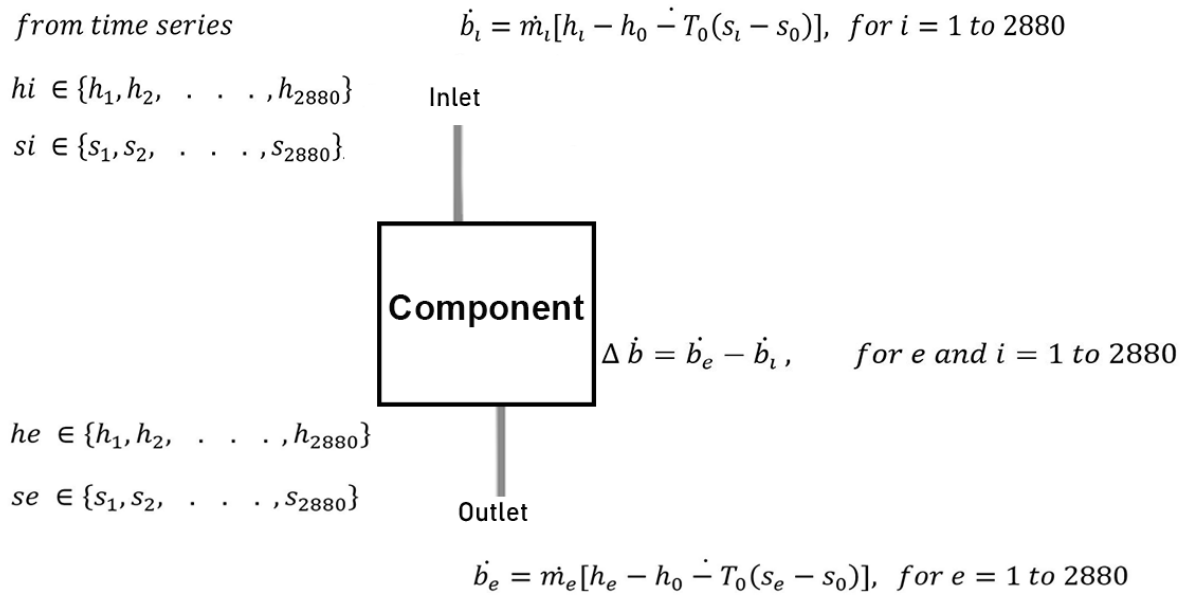


Figure 5.11: Generic example of how the exergy and exergy change over a component were calculated with the dead state variable known

The end results of calculating the exergy change over all of the components for each of the fault states and the normal operational state were 18 data sets (the exergy time series data sets) each with a structure as illustrated symbolically in Table 5.9. The headers $N0, N1, ..., N19$ represent each node in the attributed graph. The variable bNx, i with $x \in \{0, ..., 19\}$ and $i \in \{1, ..., 2880\}$ represents the exergy change over each of the "N" components or nodes.

Table 5.9: Symbolic example of the exergy time series data

Change in exergy (Db) in [kW]																			
N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N0
bN 1,1	bN 2,1	bN 3,1	bN 4,1	bN 5,1	bN 6,1	bN 7,1	bN 8,1	bN 9,1	bN1 0,1	bN1 1,1	bN1 2,1	bN1 3,1	bN1 4,1	bN1 5,1	bN1 6,1	bN1 7,1	bN1 8,1	bN1 9,1	bN 0,1
bN 1,2	bN 2,2	bN 3,2	bN 4,2	bN 5,2	bN 6,2	bN 7,2	bN 8,2	bN 9,2	bN1 0,2	bN1 1,2	bN1 2,2	bN1 3,2	bN1 4,2	bN1 5,2	bN1 6,2	bN1 7,2	bN1 8,2	bN1 9,2	bN 0,2
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
bN 1, 2880	bN 2, 2880	bN 3, 2880	bN 4, 2880	bN 5, 2880	bN 6, 2880	bN 7, 2880	bN 8, 2880	bN 9, 2880	bN1 0, 2880	bN1 1, 2880	bN1 2, 2880	bN1 3, 2880	bN1 4, 2880	bN1 5, 2880	bN1 6, 2880	bN1 7, 2880	bN1 8, 2880	bN1 9, 2880	bN 0, 2880

A link to the actual data for both the heat transfer time series and the exergy change time series can be found in appendix B. Using this data, the FDI method could be applied. This will be discussed in the next chapter along with the final results obtained for the EGBV, FDI approach.

5.7 Conclusion

In chapter 5 each FT was discussed in detail, noting the changes that each FT will cause in the STS. A detailed explanation of how each FT was simulated using the Ebsilon model is given. The FT's is summarised in four tables showing the variation in size or location of each of the faults. The data procurement from the physical system are discussed with reference to how and where in the STS the data were measured. A layout of the resulting data tables are given in Table 5.6. The adjustments to the data in order to obtain an expanded data set with 2880 variables are explained from which time series data were developed for which the process was explained in the flow diagram of Figure 5.8. The noise data generation for components from which physical measurements were not obtained, are also explained and the final time series data sets for each FT are then illustrated in Figure 5.10. The final section of chapter 5 was dedicated to the discussion and explanation of composing the energy graph applicable data sets. A clear explanation on the process is given with an example of who the heat flow and exergy data for each component were obtained. The chapter is concluded with the final data sets as shown in Tables 5.8 and 5.9.

CHAPTER 6

Energy-based fault detection and isolation

This chapter describes the energy-based graph matching approach followed for FDI. The three approaches considered, and the reasons behind the decisions are explained with specific focus on the differences between them. A detailed elaboration on the process followed is given, with the FDI results obtained in terms of detection, isolation and the performance of the various approaches in terms of sensitivity, robustness and false detection rates are given.

6.1 Introduction

In the previous 3 chapters the required preparation for energy-based FDI was discussed with a clear focus on the three high level objectives; Obtain a static thermodynamic model, compositing an attributed graph of the system and acquisition of time series data. Since these objectives were met, the next step would be to apply the energy based FDI.

In this chapter the composition of attributed matrices from the attributed graph is discussed. The focus is placed on the detection and isolation of the fault types, applying the time series data from the previous chapter in three different approaches.

This is accomplished through composing the attributed matrices, calculating the cost and residual matrices respectively, evaluating the distance parameters, eigenvalues and residuals of the cost or residual matrices and applying an analysis sequence to enable detecting and isolation of the faults.

6.2 Overview

The initial graph matching approach considered was based on the so-called distance parameter approach as mentioned by [76]. In the article an approach was proposed where two attributed graphs were matched through applying the HEOM function to the attribute matrices of the attributed graphs. This generates a square cost matrix, indicating the dissimilarities between the two graphs. A further analysis for FDI applications is then performed by calculating the average diagonal value of this cost matrix, called the distance parameter. The method allows for large reduction of data since the dissimilarity information between the initial attributed graphs are reduced to a single parameter. By following a specific sequence when analysing all of the attributed graphs, both fault detection and fault isolation can be accomplished. The method however did not reveal satisfactory

isolation results and hence a second approach (the eigenvalue approach) suggested by [3, 51, 77] was evaluated.

In this method the eigenvalues of the cost matrix is determined rather than computing only the distance parameter. Utilizing the information given by the eigenvalues, more information on the dissimilarities between the two graphs were obtained and hence it was reasoned that the isolation of faults might improve. The approach, however, increased the amount of data to be processed by a number, at most, n^2, with n the number of nodes in the attributed graph. This approach was also unable to effectively distinguish between the different intensities and locations of FT 2, although detection of all faults is achieved along with reasonable isolation for FT's 1, 3 and 4.

It was therefore decided to follow a third approach called the residual approach suggested in the PhD of H.Nesser [78]. This residual approach simply compared each of the values in the attributed matrices directly after the matrices were normalized. This approach disposes of using the HEOM function, simplifying the mathematical procedure while also keeping a direct link to the original attributed graph structure since the residual matrix is of the same dimensions as the node signature matrix and can easily be traced back to the attributed graphs. The data reduction advantage of the other two methods is, however, lost and hence a matrix of the same size as the node signature matrix is analysed for the FDI.

As an overview the flow diagram in Figure 6.1 explains the process flow of the three approaches. For future reference the three approaches are uniquely designated as follows:

Approach 1 : Distance parameter approach
Approach 2 : Eigenvalue approach
Approach 3 : Residual matrix approach

For all three of the mentioned approaches, the compositing of a node signature matrix using the attributed graph, illustrated in Figure 4.4, and the energy data as discussed in the previous chapter, are required. It is important to note the following regarding the data sets discussed in Chapter 5: 1) The 24-hour physical data are used as an indication of the amount of variation the system has under normal steady state operation. 2) The 2880 data points of the time series are therefore seen as variation in the data that is used as an indication of the amount that a parameter can vary under fault conditions due to normal system operational fluctuations. 3) The result of the FDI methods will thus be an indication of how effective the methods are to detect any of the 2880 possible energy of exergy variations that might occur once a fault is intruded in the system. The next section is specifically focussed on describing the process through which these node signature matrices were obtained.

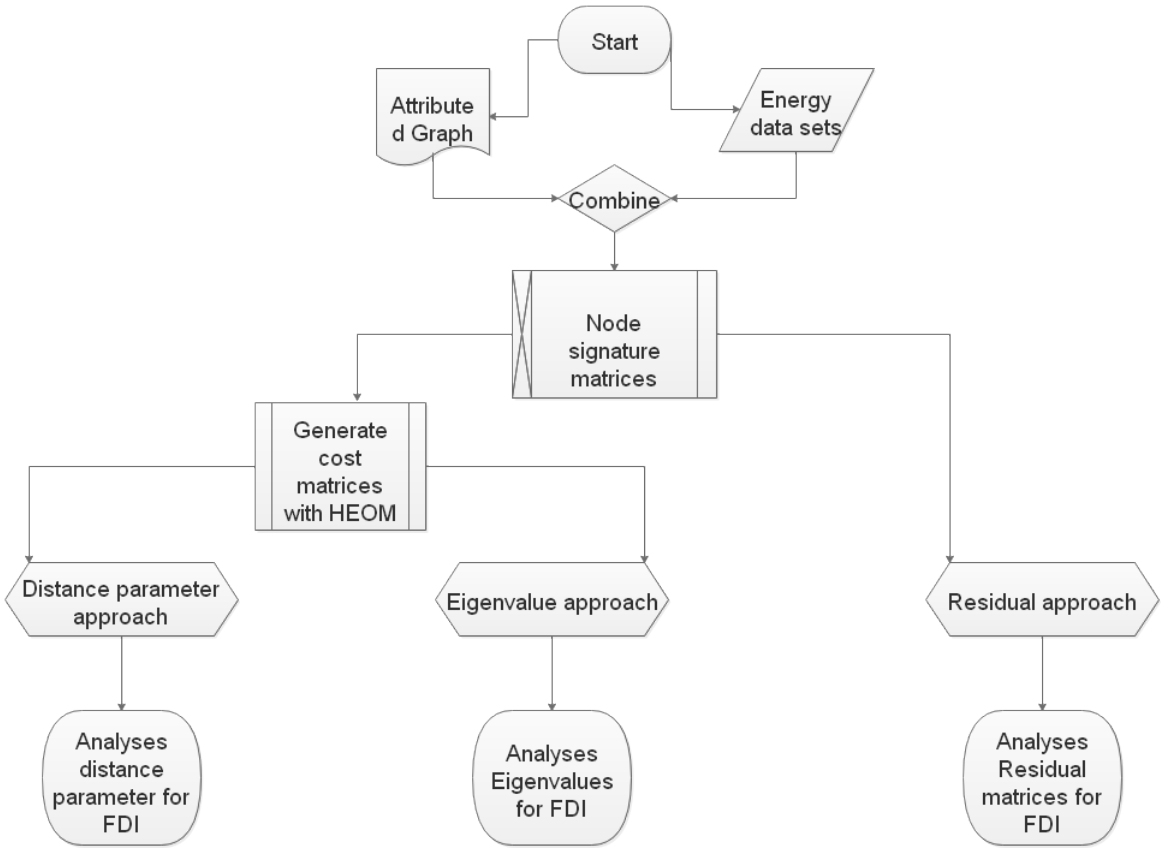


Figure 6.1: Flow diagram illustrating the process flow for the three graph matching approaches

6.3 Composing of the node signature matrices

6.3.1 Node signature matrix

To compose a node signature matrix, refer to the attributed graph in Figure 4.4 (b). Using the approach in [28] an adjacency matrix can be composed from the graph by composing a matrix with rows equal to n and columns equal to $n+1$, where n = the number of nodes in the attributed graph. The columns are compiled equal to $n+1$ to allow one column for the exergy change attribute over each of the nodes. The consequences of this is a nxm node signature matrix indicating in the first column the change in exergy over each node and in each row the energy transfer from one node to another. In Figure 6.2 a symbolic example of this process is given. As indicated with the arrows, the "db" values, representing change in exergy, are placed in the first column and the energy transferred from one node to another is placed in the respective cells, according to the attributed graph.

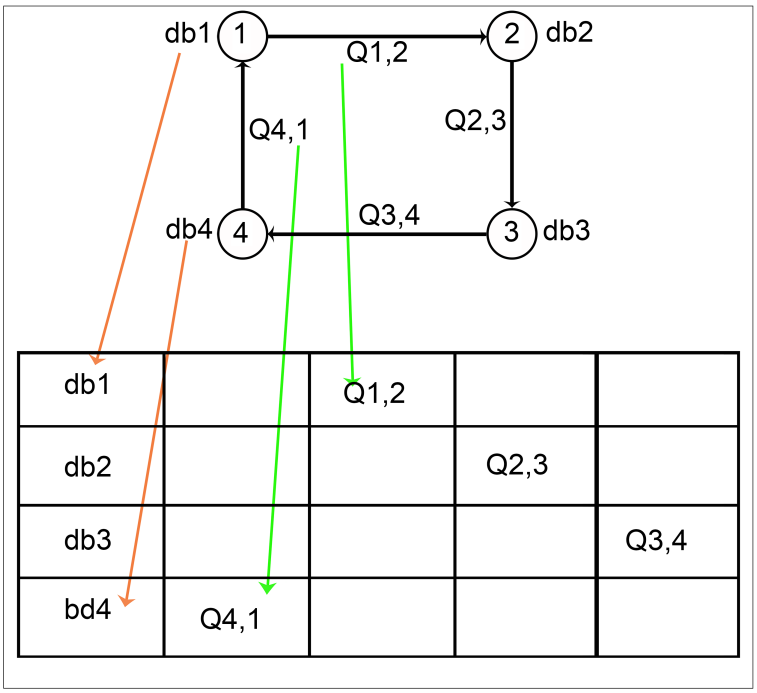


Figure 6.2: Example of how a node signature matrix is composed from an attributed graph

The node signature matrix obtained for the attributed graph of the STS as shown in Figure 4.4, is shown in Figure 6.3. Note that the values, representing the energy transfer from one node to another, is indicated using the node numbers, with the first number indicating the node from which the energy is transferred, and the second number indicating the node to which the energy is transferred. Thus, the value 1,2 indicates an energy transfer from node 1 to node 2.

		N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N0
	Δb	$\dot{Q}_{N1}$	$\dot{Q}_{N2}$	$\dot{Q}_{N3}$	$\dot{Q}_{N4}$	$\dot{Q}_{N5}$	$\dot{Q}_{N6}$	$\dot{Q}_{N7}$	$\dot{Q}_{N8}$	$\dot{Q}_{N9}$	$\dot{Q}_{N10}$	$\dot{Q}_{N11}$	$\dot{Q}_{N12}$	$\dot{Q}_{N13}$	$\dot{Q}_{N14}$	$\dot{Q}_{N15}$	$\dot{Q}_{N16}$	$\dot{Q}_{N17}$	$\dot{Q}_{N18}$	$\dot{Q}_{N19}$	$\dot{Q}_{N0}$
N1	Δb_1		1.2																		
N2	Δb_2			2.3	2.4																
N3	Δb_3				3.4												3.16				
N4	Δb_4					4.5															4.0
N5	Δb_5						5.6														5.0
N6	Δb_6							6.7													6.0
N7	Δb_7								7.8												
N8	Δb_8									8.9											8.0
N9	Δb_9										9.10										9.0
N10	Δb_{10}											10.11	10.12								
N11	Δb_{11}																				11.0
N12	Δb_{12}													12.13	12.14						
N13	Δb_{13}																				13.0
N14	Δb_{14}															14.15					
N15	Δb_{15}																		15.15		15.0
N16	Δb_{16}																	16.17			16.0
N17	Δb_{17}																			17.19	
N18	Δb_{18}																	18.17			18.0
N19	Δb_{19}	19.1																			19.0
N0	Δb_0	0.1		0.3										0.3	0.14	0.15			0.18	0.19	

Figure 6.3: Node signature matrix of the attributed graph

Once this matrix is composed, the time series energy attributes illustrated in Tables 5.8 and 5.9 are used to populate the matrix. A simple program in Matlab is written that runs a loop, taking each row of these respective time series tables and writing them to the cells of the node signature matrix. The resulting matrix is then saved in an array. This leads to an array of matrices corresponding to each row of the time series. Repeating this for all 18 states (4 Fault states, variations of the fault states and one normal operating state) results in an array with 18 columns and 2880 rows with each index in this array representing a node signature matrix. An illustration of the 18 x 2880 array,

consisting of attributed matrices, is shown in Table 6.1. The next step would be to apply the three respective graph matching techniques in a specific sequence.

Table 6.1: Example of the array of attributed matrices for each state

	Normal Operation	**TF1_HPT3%**	...	**FT2_LPT1**	...	**FT3_LPT3%**	...	**FT4_9%NPSH**
1	20X21 Matrix	20X21 Matrix	...	20X21 Matrix	...	20X21 Matrix	...	20X21 Matrix
2	20X21 Matrix	20X21 Matrix	...	20X21 Matrix	...	20X21 Matrix	...	20X21 Matrix
.	.	.	...	.	...	.	...	.
.	.	.	...	.	...	.	...	.
.	.	.	...	.	...	.	...	.
2880	20X21 Matrix	20X21 Matrix	...	20X21 Matrix	...	20X21 Matrix	...	20X21 Matrix

6.4 Approach 1 and 2

6.4.1 Composing the cost matrix with the HEOM function

The idea behind a cost matrix is simply a matrix that is calculated as the result of the comparison of two other matrices, indicating the dissimilarities or "distance" between the compared matrices. There is a number of different methods to accomplish this, with even variations of these methods, as mentioned in chapter 2. The method chosen for this was the Heterogeneous Euclidean-Overlap Metric (HEOM) as given in equation (2.5). This function allows for comparison of two matrices when either categorical or continuous attributes are used. A variety of distance operators is given by [79] ranging from obtaining a distance parameter (d_r) by simply using the $range$ as illustrated in (2.5), or by substituting the $range$ with either one or four standard deviations (σ or 4σ) of the dataset. For the applications of this study the slight differences obtained when using a different $range$ is of minimalistic effect, and thus only the $range$ as defined in (2.5) is considered. The advantage of this approach is that the HEOM normalizes the data as part of the function using this $range$ parameter.

In order to compose a cost matrix for any two matrices, consider the following example for matrices A and B as shown in Figure 6.4. In this example, matrix A is used as the reference matrix and hence, the range is calculated with the values of A. For the application of FDI, A and B would obviously be equated to a specific energy node signature matrix in Table 6.1.

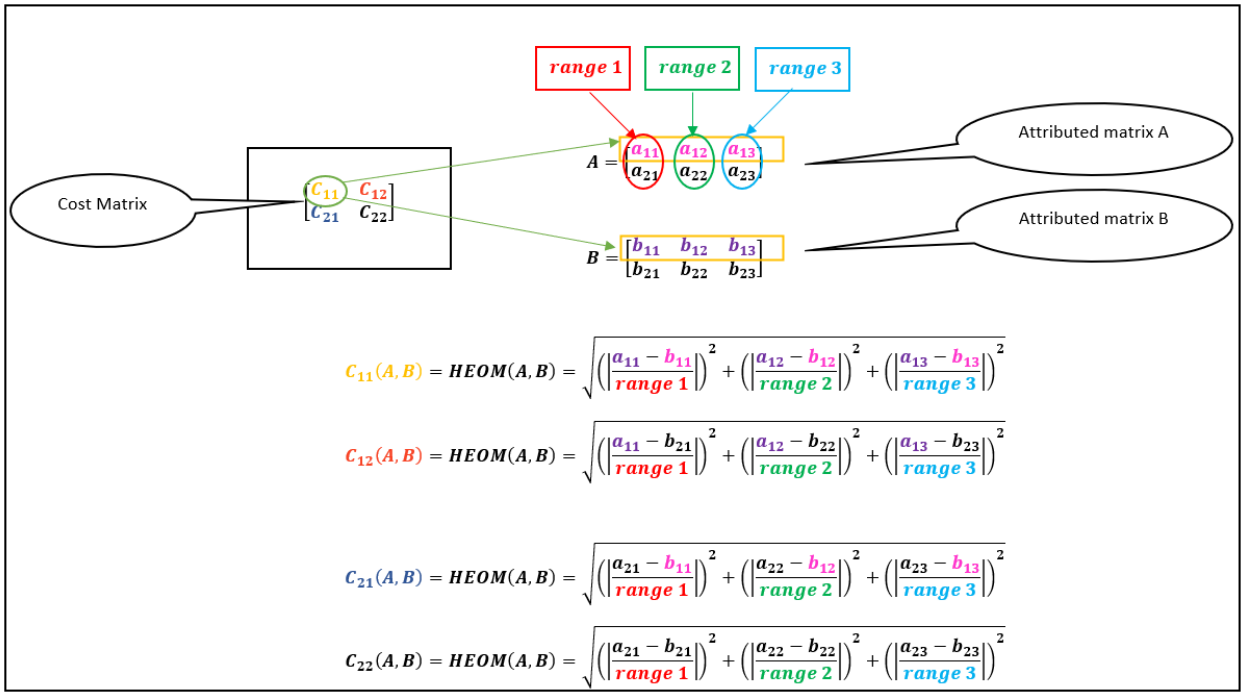


Figure 6.4: Illustration of how the cost matrix is calculated

Since faults must be detected and isolated the rationale used to determine which cost matrices to calculate can be explained as follows: Consider Figure 6.5 showing a hypothetical energy variable, measured at some location in the system (for example at position $L1,2$, the HPT inlet). Suppose the system initially operates under normal conditions (i.e. the variable is within the specified acceptable limits) until $time = T$. At $time = T$ a fault is introduced causing the energy variable to change more than the designed acceptable specifications. To detect this fault, the attributed matrices containing attributes measured during $time < T$, should not raise concerns, but any attributed matrices containing measurements taken at $time >= T$, should be flagged to indicate a fault condition. Keep in mind that the attributed matrices illustrated in Table 6.1, represent the normal operational energy measurements ($time < T$) in column 1. Each the fault state measurements ($time >= T$) are represented in the subsequent columns 2 to 18, for all the locations in the system, where measurements were taken during fault conditions.

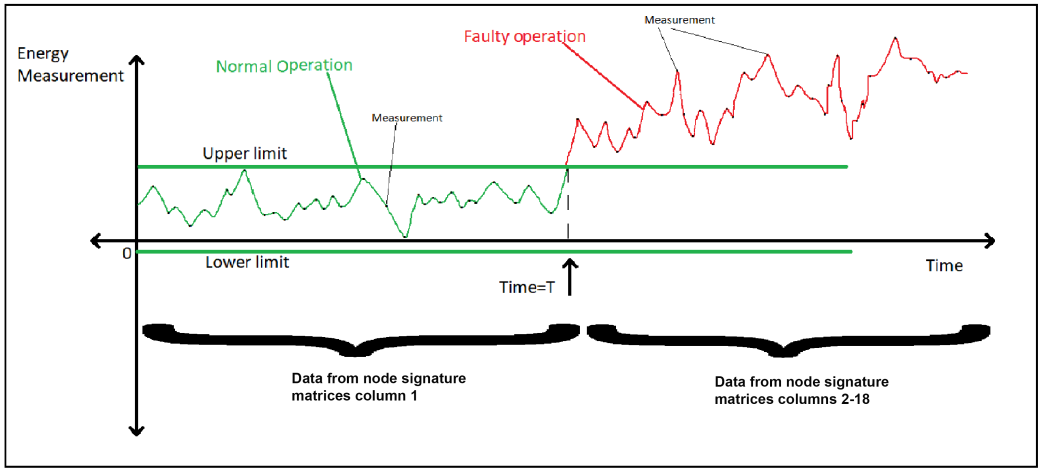


Figure 6.5: Example of an energy measurement in the system changes over time as the system state change from a normal to faulty

For clarity purposes, any node signature matrix obtained from measurements taken under

normal conditions (measurements in column 1 Table 6.1) is referred to as a normal node signature matrix and any matrix obtained from measurements taken under faulty conditions (column 2-18 in Table 6.1) is referred to as a fault node signature matrix.

Next, consider that for real life applications, the measurements that are taken in the system and used as input to the FDI method, are unspecified (normal or faulty is unknown) from the perspective of the FDI method. These operational state measurements, captured in an operational graph, must therefore be matched to some normal state reference graph, to determine if a fault or normal operation is occurring and thus specifying the system state. Since the cost matrices give an indication of the cumulative "distance" between the graphs that are matched, it can be reasoned that, if the operational state represents a normal state, the resulting values in the cost matrix would be small when a match to a normal reference matrix is calculated. Opposed to this, the values in the cost matrix would be larger when the operational graph represents a fault condition.

Since the system states are known under which the attributed matrices in Table 6.1 were obtained, the variation in the cost matrix values for the operational state, being either normal or faulty, can be obtained. Therefore, to detect and isolate faults the operational graph can be matched to a graph representing the normal state as well as graphs representing the fault states. The match with the shortest distance will then indicate the largest probability of the operational state being equal to that state. This approach however implies that the possible faults in the system must be known in order to have a reference of that fault signature to which the operational graph can be matched. This might not always be the case in real live applications since novel faults may arise.
If an operational graph represents a fault that no reference is known for, a match to any of the known fault graphs would be inconclusive. To obtain these reference graphs a graph for each of the states is composed using the average energy and exergy measurements from Tables 5.8 and 5.9. Therefore, for example the normal state reference graph (G_{Navg}) is composed using the average energy data from each node and link as calculated with $AVG = \frac{1}{i}\sum_{1}^{2880} N_{ei}$. The variable N_{ei} refers to the energy data as shown in figure 6.6. The same process is followed for each of the fault reference graphs. This is different from conventional FDI methods that commonly use only 70% of the data set to "train" the method while using the other 30% to test the method. The reason why this was not done in a typical way is that the reference graphs are constructed with the average data. Since the average of any statistical reliable sample (70% of the data in the typical case) should have the characteristic that the average of the sample and the initial population is approximately the same, it does not make a difference to use the whole data set. Once the reference graphs were constructed the graph matching could be done to obtain the cost matrices. The sequence of matching was proposed by [76] and is illustrated in Figure 6.6 along with the process of obtaining the reference graphs. The operational graph illustrated as $G\_op \in matrices\{N, F1, ..., F4\}$ represents all the matrices in Table 6.1.

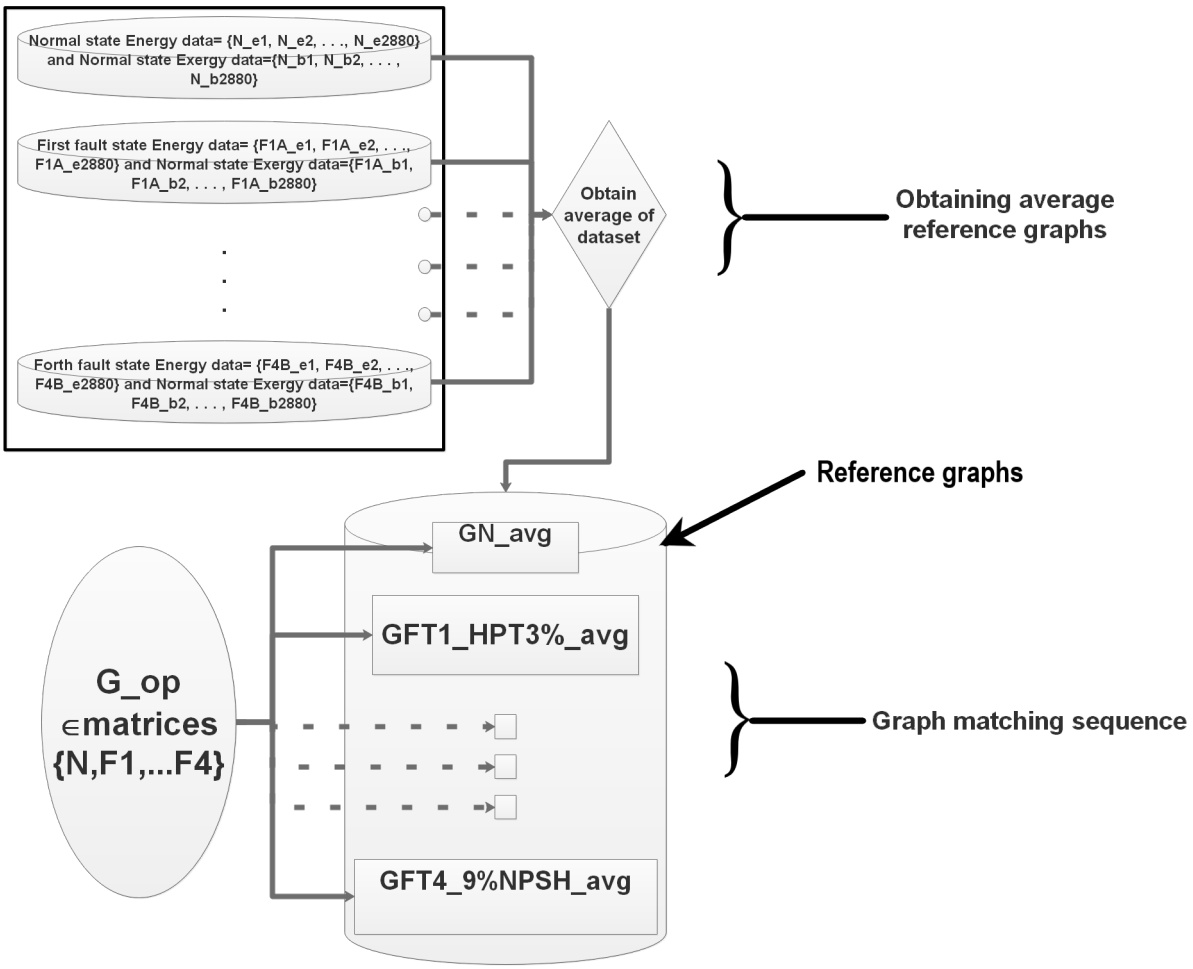
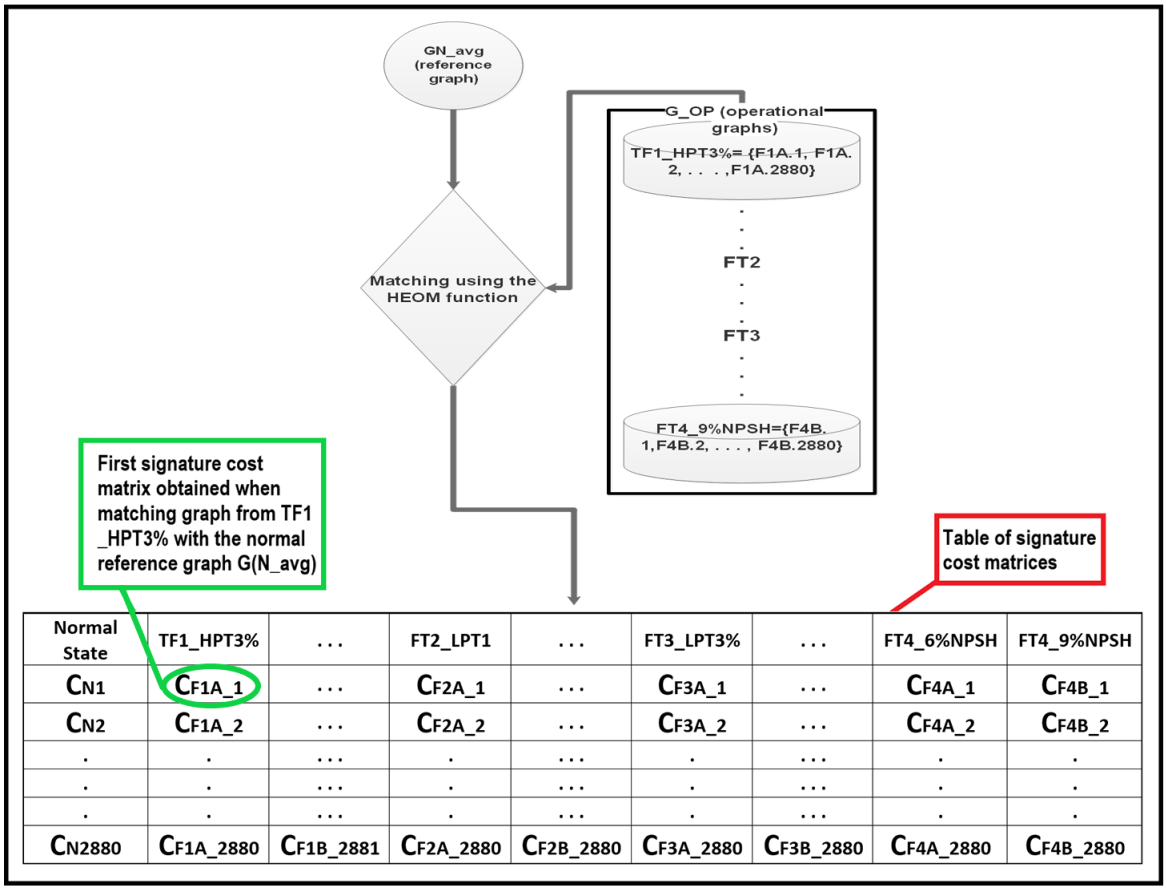


Figure 6.6: Obtaining average reference graphs and matching sequence for calculating the cost matrices

As shown in Figure 6.6 the operational graph can be composed for any of the system states (Normal (N), Fault 1 (F1), . . ., Fault 4 (F4)). In order to execute the matching, a simple Matlab program was written that iteratively composes the operational graph for each of the node signature matrices in Table 6.1 and then match the operational graph to the reference graphs using the HEOM function.

A signature cost matrix is then obtained for each match. An example of this set of signature cost matrices, for all operational graphs, matched with only the graph of the normal reference state (GN_{avg}), is shown in Figure 6.7. Similar tables consisting of signature cost matrices are obtained for each of the operational graphs matched to each of the 18 reference graphs. In this way a total of 18 tables of signature cost matrices are obtained.



Normal State	TF1_HPT3%	...	FT2_LPT1	...	FT3_LPT3%	...	FT4_6%NPSH	FT4_9%NPSH
C_{N1}	C_{F1A_1}	...	C_{F2A_1}	...	C_{F3A_1}	...	C_{F4A_1}	C_{F4B_1}
C_{N2}	C_{F1A_2}	...	C_{F2A_2}	...	C_{F3A_2}	...	C_{F4A_2}	C_{F4B_2}
.	.	...	.	...	.	...	.	.
.	.	...	.	...	.	...	.	.
.	.	...	.	...	.	...	.	.
C_{N2880}	C_{F1A_2880}	C_{F1B_2881}	C_{F2A_2880}	C_{F2B_2880}	C_{F3A_2880}	C_{F3B_2880}	C_{F4A_2880}	C_{F4B_2880}

Figure 6.7: Illustration of the fault signature cost matrices data set

Thus for fault detection, an operational graph is matched to the reference graphs and 18 resulting signature cost matrices are obtained. The 18 signature cost matrices should then be evaluated and if it is found that the signature cost matrix, showing the smallest dissimilarities, were not the signature cost matrix obtained when matching to the normal state reference graph, a fault is detected.

An example of the actual signature cost matrix obtained (C_{F1A_1}) for a graph of the $TF1\_HPT3\%$ matched to the normal operational reference (GN_{avg}) (shown in Figure 6.7) can be seen in Figure 6.8 along with the cost matrix (CN_1) obtained when the normal reference graph was matched to a graph representing the normal state.

Note that the diagonal values of CN_1 are all close to zero as opposed to some of the diagonal values of C_{F1A_1} not being relatively close to zero. This is an indication that CN_1 represents a normal operating state and that C_{F1A_1} represents a fault state.

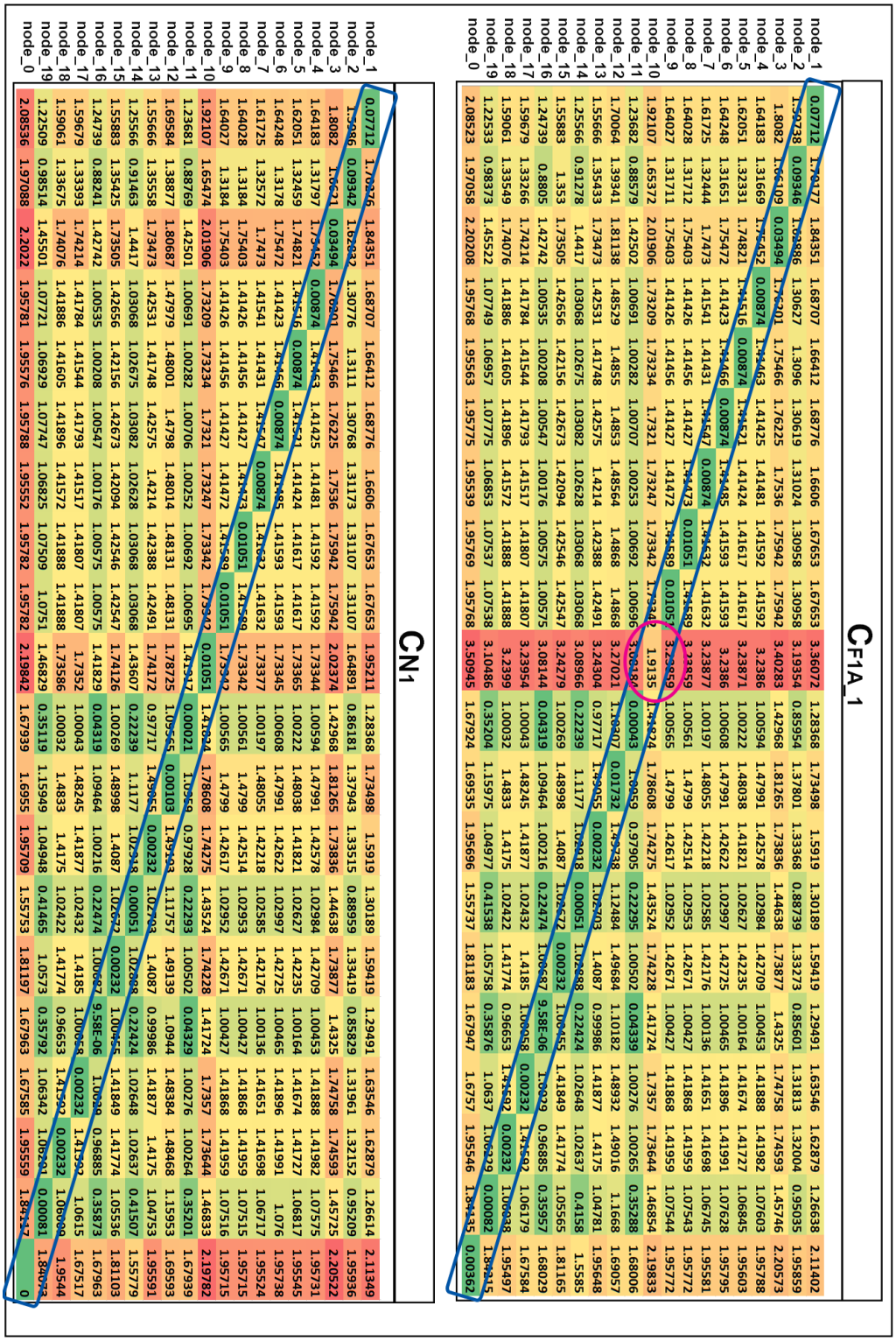


Figure 6.8: Cost matrix obtained for $G(N_{avg})$ matched to graph from $FT1\_HPT3\%$ as well as the normal state

Regarding isolation of the faults, the basic idea would be that the fault signature cost matrices of each of the fault states are unique. Thus no one fault signature set of cost matrices are similar to any other fault signature set of cost matrices. This implies that, if the operational graph represents a fault state and is then matched to each of the reference states (as for fault detection), 18 resulting signature cost matrices will be obtained. These matrices can be analysed to show that the signature cost matrix, with the least dissimilarities, is the closest match to the fault reference graph which is equal to the state represented by the operational graph.

An example of this is shown in Figure 6.9. The first matrix is the signature cost matrix obtained when the operational state represented $FT1\_HPT3\%$ was matched to the normal reference graph $G(N_{avg})$. The second matrix is the signature cost matrix obtained when the operational state representing $FT1\_HPT3\%$ was matched to the graph of the reference state, representing the same FT ($G(FT1\_HPT3\%)\_avg$). It can clearly be seen from the diagonal values that the operational graph, matched to the reference graph G(FT1_HPT3%_avg), gave the closest match. These signature cost matrices can now be evaluated to determine detection and isolation of faults. The next subsection will specifically focus on approach 1 (distance parameter) as a method of analysing the signature cost matrices.

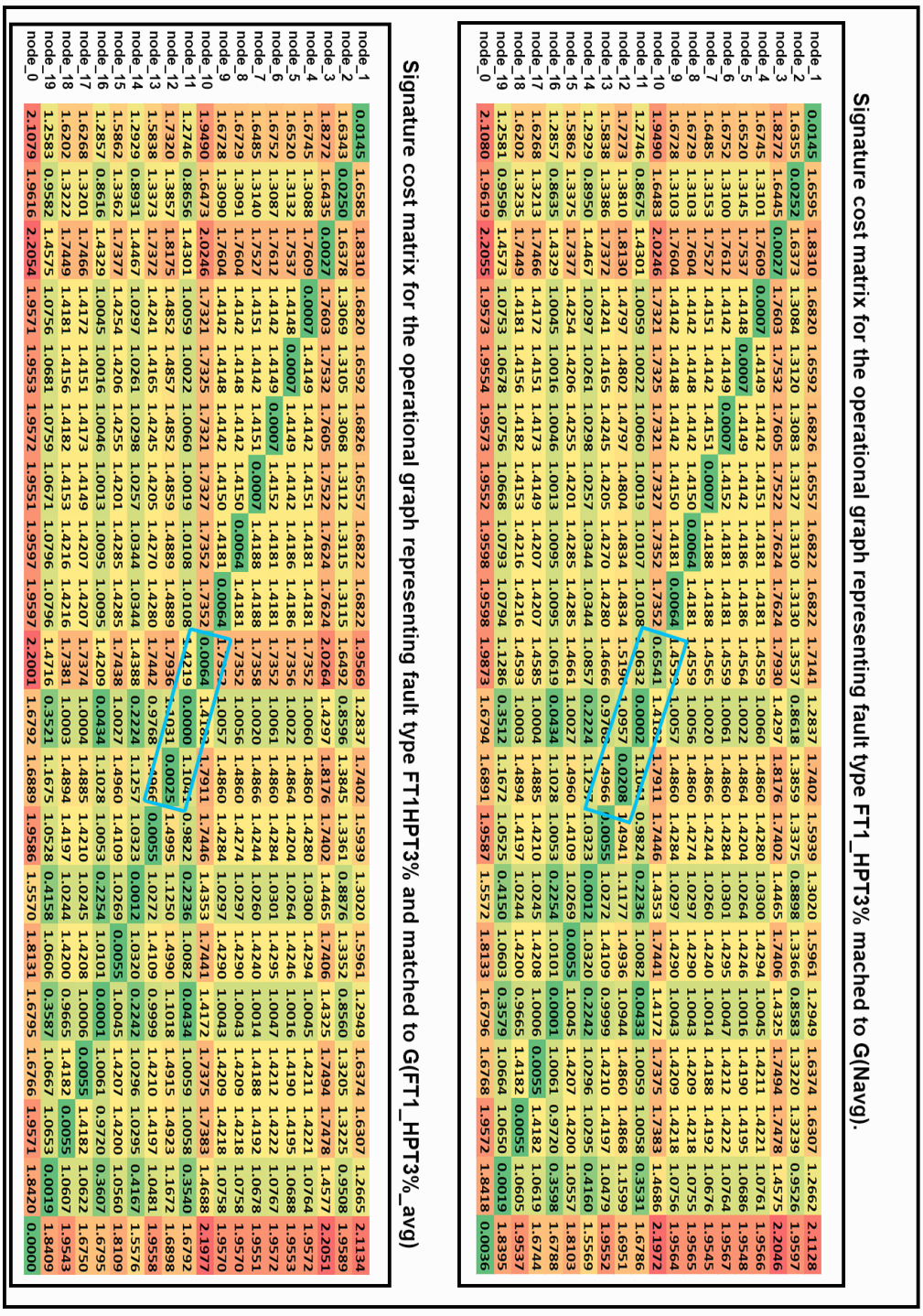


Figure 6.9: Example of isolation of a fault based on there signature cost matrices

6.4.2 Approach 1 (Distance parameter)

When considering the signature cost matrices shown in Figures 6.8 and 6.9, the basic idea of using the distance parameter approach is logic and clear. Since the diagonal values of the signature cost matrices encapsulate the important information stored in the cost matrices it would be of value to simply examine these diagonal values. The approach accomplishes this by computing the average value of the diagonal. This results in reducing the cost matrix as a whole to a single value that indicates the distance between the two graphs encapsulated in the cost matrix. The equation proposed by [76] to calculate the distance parameter is given by:

$$DC = \frac{\sum_{k=1}^{n} C_{kk}}{n}, \tag{6.1}$$

with kk refers to the relevant diagonal elements in the cost matrix. n represents the number of diagonal elements in the square cost matrix or then also the number of nodes of the graph.

If this approach is successful in terms of fault detection, the results expected would indicate that the distance parameters for all 2880 signature cost matrices obtained when the operational graph, represents a normal state, is matched to the normal reference graph ($G(N_{avg})$, are small in comparison to the distance parameters obtained for any of the fault state signature cost matrices when the operational graph, representing a fault state, is matched to the same normal reference graph. This implies that if one would plot these 2880 distance parameters on the same axis a clear distinction in distance parameter size would be expected. This would only be the case if each parameter was obtained from matching the normal reference graph ($G(N_{avg})$ with the operational graph representing all possible states (1 normal and 4 fault states). This result is plotted in Figure 6.10. The blue graph shows the distance parameter obtained from matching the operational graph, when representing a normal state, to the normal state reference graph.

The other graphs in the figure represent the distance parameter obtained for the same matching to the normal reference graph, but with the operational graph representing the different fault states (TF1_HPT3%,. . ., FT2_LPT1,. . ., FT3_LPT3%, . . ., FT4_6%NPSH, FT4_9%NPSH). It is clear from the figure that each of the fault states can easily be detected regardless of the variation in the size of the fault or the location of the fault. There are also no distance parameter results for any of the fault states that overlap with the distance parameters of the normal state, thus no false detection of faults occurred.

In terms of sensitivity and robustness, it is clear that the approach can detect any of the faults regardless of the change in intensity or location. This implies good robustness and low sensitivity to fault size changes. Thus, the method is not sensitive to changes in the fault size of location and can therefore still detect all the faults.

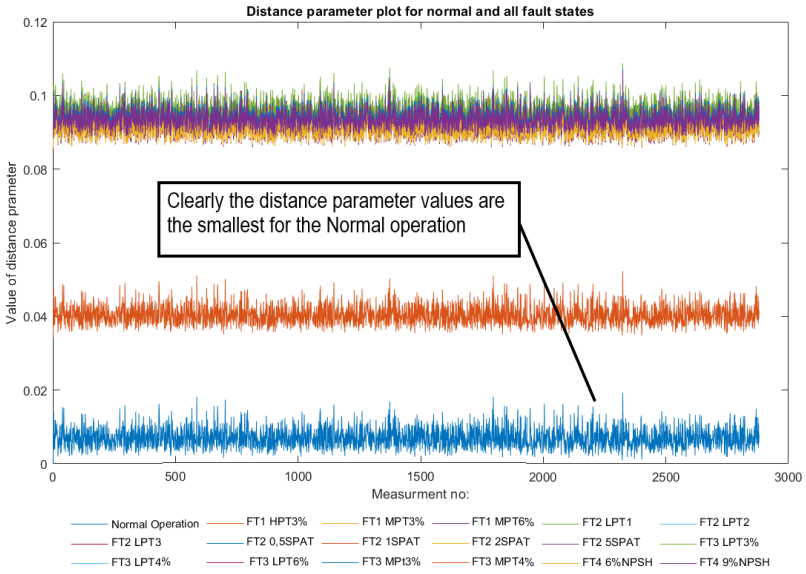


Figure 6.10: Distance parameter results- Normal operational graphs matched with the normal reference graph

In order to isolate the faults, it is necessary to determine if the signature cost matrices of each of the faults are unique in comparison to any other fault's signature cost matrices. This implies that the smallest distance parameter values would be obtained when an operational graph, representing a specific fault, is matched with the reference graph of that same fault. The results can also be plotted as shown for the detection of faults. In Figure 6.11 the resulting graph for fault state TF1_HPT3% can be seen. It is clear that the smallest distance is obtained for TF1_HPT3%. Since the operational graph represented TF1_HPT3%, and was also matched to the reference graph TF1_HPT3%_avg, the results are as expected.

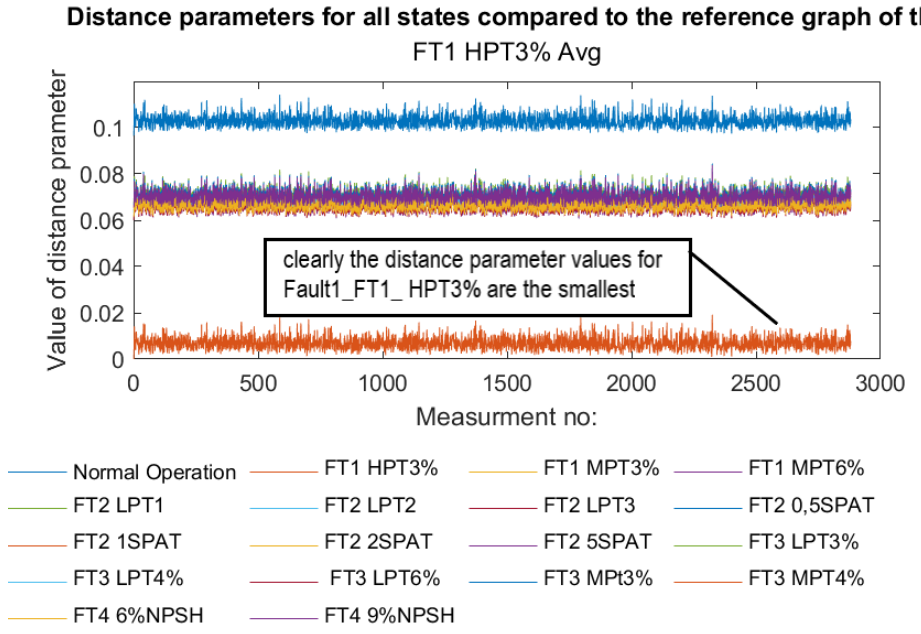


Figure 6.11: Distance parameter result - G(TF1_HPT3%_avg)

These results seem to be accurate in terms of isolation of FT FT1_HPT3%, however if the same process is repeated for all the other fault types, the results are inconclusive. An example of this is shown in Figure 6.12 for FT FT2_2SPAT. It can obviously be seen that the smallest distance parameter values of some faults are in the same size range, resulting in an inconclusive isolation of the faults. Graphs of all of the isolation results can be found in appendix C, section C.1, subsection C.1.1.

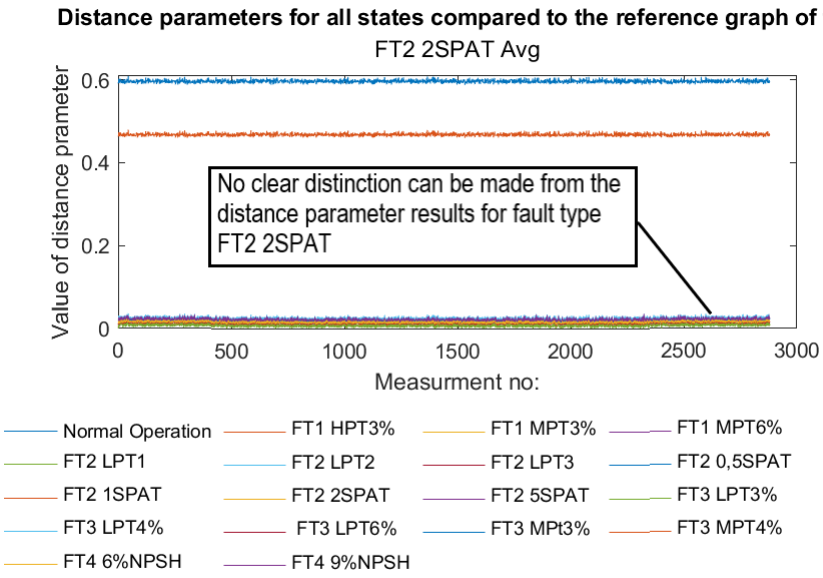


Figure 6.12: Distance parameter results- G(FT2_ 2SPAT_ avg)

The reason for this is due to the fact that some of the fault types cause energy changes in the system that are within the same range as those caused by other fault types, with exception of the HPT which delivers 3 to 5 times more power (approximately 130 MW) than the stages of the MPT (approximately 40 MW) and LPT (approximately 25 MW). These changes happen in specific locations in the system of which the locational information is encapsulated in the attributed graph. The distance parameter approach, however, does not consider this locational information since all the information is reduced to one value. Accompanying this, is the fact that the noise fluctuations in the system also increased the amount by which the size of energy values under one FT would overlap the size in energy values of another FT. This lead to the distance parameter's value distribution, for some fault types, to overlaps with those of other. If the distribution of the distance parameters obtained when matching each FT to the normal reference graph is known, an indication of the extent that isolation can be accomplish is obtained. By plotting the distribution of the distance parameter for fault types FT1_MPT3%, FT2_LPT1, FT3_LPT3% and FT4_6%NPSH compared to the normal reference graph on the same axis, the overlapping of the distance parameters distribution is clearly illustrated in Figure 6.13. This can be repeated for any of the fault types and the same result will be obtained. Only these four faults are plotted for illustrative purposes.

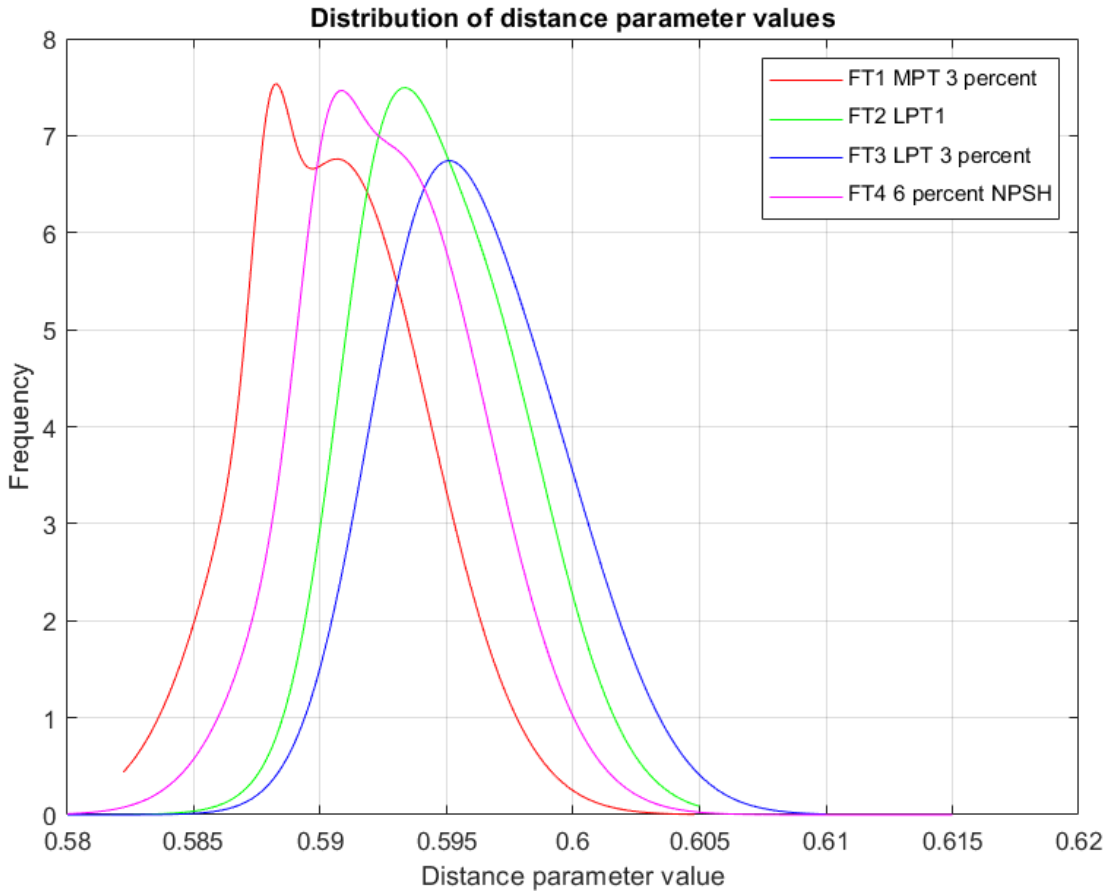


Figure 6.13: Distributions plotted for TF1_MPT3%, FT2_LPT1, FT3_LPT3% and FT4_6%NPSH

If each of the 2880 measurements of each fault state are evaluated and the correct isolation as a percentage of the number of measurements of each fault state is counted, Table 6.2 is obtained. From the table it is clear that the method works well to detect the faults and can also isolate FT1_HPT3% for 100% of the 2880 measurements. For all of the other faults types, however, isolation is not achieved. A slight increase in accurate isolation (29% to 34%) is seen for FT2_LPT3, FT3_LPT4% and FT4_6%NPSH. The results, however are still inconclusive for most of the measurements with no specific pattern identifiable. This results also corresponds with the distribution plots shown in Figure 6.13.

From these results it can be concluded that the only evaluating a distance parameter leads to ineffective isolation of faults. If the energy changes in the system, due to a fault, are closely related, the distance parameter will not retain enough information to allow for fault isolation. It can thus be state that the reduction in information caused by reducing the cost matrices to a single distance parameter are ineffective. This would also imply that if some of the faults would result in small changes that do not necessarily differ much from the normal state, both detection and isolation of the faults might be ineffective. If one can therefore include more information than a single distance parameter or the location information of the fault, as encapsulated by the attributed graph, this error might be solved. A method that seems to retain more of the information would be the eigenvalue approach. The next subsection will be dedicated to discussing this approach.

Table 6.2: Distance parameter approach- detection and isolation rate results

	Summary	Normal Operation	FT1_HPT3%	FT1_MPT3%	FT1_MPT6%	FT2_LPT1	FT2_LPT2	FT2_LPT3	FT2_0.5SPAT	FT2_1SPAT	FT2_2SPAT	FT2_5SPAT	FT3_LPT3%	FT3_LPT4%	FT3_LPT6%	FT3_MPt3%	FT3_MPT4%	FT4_6%NPSH	FT4_9%NPSH
Detection	100%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Isolate fault	**FT1_HPT3%**	100%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Isolate fault	**FT1_MPT3%**	9%	0%	9%	7%	4%	6%	6%	6%	7%	6%	5%	6%	5%	6%	6%	6%	6%	7%
Isolate fault	**FT1_MPT6%**	12%	0%	6%	12%	6%	3%	7%	5%	6%	7%	5%	8%	5%	5%	7%	6%	6%	4%
Isolate fault	**FT2_LPT1**	15%	0%	5%	6%	15%	7%	3%	6%	7%	4%	5%	6%	4%	7%	6%	6%	7%	7%
Isolate fault	**FT2_LPT2**	10%	0%	6%	6%	6%	10%	7%	6%	7%	8%	6%	6%	7%	4%	3%	3%	8%	7%
Isolate fault	**FT2_LPT3**	30%	0%	3%	4%	3%	3%	30%	6%	4%	3%	5%	3%	12%	4%	2%	3%	13%	2%
Isolate fault	**FT2_0,5SPAT**	10%	0%	6%	5%	6%	7%	7%	10%	5%	5%	6%	6%	7%	6%	6%	6%	6%	5%
Isolate fault	**FT2_1SPAT**	12%	0%	5%	6%	6%	7%	7%	6%	12%	10%	6%	6%	6%	5%	5%	4%	4%	4%
Isolate fault	**FT2_2SPAT**	19%	0%	6%	7%	6%	7%	7%	6%	7%	19%	6%	6%	4%	6%	4%	3%	3%	3%
Isolate fault	**FT2_5SPAT**	19%	0%	4%	4%	6%	7%	7%	6%	7%	7%	19%	4%	3%	3%	4%	4%	7%	8%
Isolate fault	**FT3_LPT3%**	6%	0%	6%	8%	6%	7%	7%	6%	7%	7%	6%	6%	7%	4%	6%	4%	6%	7%
Isolate fault	**FT3_LPT4%**	34%	0%	3%	3%	4%	3%	12%	2%	2%	1%	4%	3%	34%	5%	5%	5%	10%	4%
Isolate fault	**FT3_LPT6%**	16%	0%	3%	3%	4%	4%	5%	5%	7%	7%	6%	6%	7%	16%	7%	7%	7%	7%
Isolate fault	**FT3_MPt3%**	16%	0%	5%	5%	4%	3%	4%	4%	6%	7%	6%	6%	7%	7%	16%	7%	7%	7%
Isolate fault	**FT3_MPT4%**	17%	0%	5%	5%	5%	3%	4%	6%	5%	5%	5%	6%	7%	7%	7%	17%	7%	7%
Isolate fault	**FT4_6%NPSH**	29%	0%	4%	4%	5%	5%	8%	3%	4%	3%	3%	4%	8%	5%	5%	5%	29%	5%
Isolate fault	**FT4_9%NPSH**	9%	0%	7%	4%	6%	8%	5%	6%	8%	7%	6%	4%	7%	7%	6%	7%	4%	9%

6.4.3 Approach 2 (Eigenvalues)

As mentioned in the overview of this chapter, the eigenvalue approach simply differs from the distance parameter approach in the sense that the eigenvalues of the cost matrices are calculated rather than the distance parameters. In order to understand why this method is considered the basic properties of eigenvalues needs to be recapped [80]. For more information on these fundamentals refer to appendix D. From the fundamentals it is clear that the eigenvalues of a matrix are linked to the diagonal of that matrix and since it seems like the important information of the cost matrix for FDI applications are encapsulated within the diagonal of the cost matrix; this approach can be of value. The basic methodology is thus to follow the same sequence as with the distance parameter approach, but rather calculating the eigenvalues of each of the cost matrices. This implies that for each of the cost matrices a set of 20 eigenvalues is obtained. A symbolic illustration of this is shown in Figure 6.14.

The figure shows both the sequence of comparison used to obtain the eigenvalues and the structure of the resulting data set. The sequence followed are the same as the sequence that was followed with the distance parameter approach. Thus the operational graph is composed for each of the signature node matrices of each state and then matched to the 18 reference graph respectively, using the HEOM function. The reference graphs, as illustrated in Figure 6.14, consists of the node signature matrices composed form the average energy values of each state. Note that the symbols in bold represent matrices and arrays of data with $n \times m$ dimensions (n is number of rows and m is number of columns). From this data set further analysis is done in order to effectively evaluate the results.

When considering the array with sets of eigenvalues, note that each of the eigenvalues in bold from Tables 0 to 17 represents an array with 20 eigenvalue elements. Similarly, as with the distance parameter approach, each of the 2880 rows from Tables 0 to 17 is analysed to determine either detection or isolation. The analysis entails comparing each of the sets of eigenvalues in the particular row with each other in order to find the eigenvalue set representing the smallest distance between the graphs. Table 0 illustrates the results obtained when the comparison was done between the operational graph and the normal reference graph. This implies that Table 0 represents the results that must be analysed to determine the detection of faults. In the same manner Tables 1 to 17 represent the eigenvalue results that must be analysed to determine isolation of the various fault types. For the analysis of these eigenvalues, the approach used by [3], suggested a procedure that can be summarised as follows:

1) Obtain the fault eigenvalues matrix as represented by each row of the eigenvalue set tables in Figure 6.14. An example of this fault eigenvalues matrix is shown in Figure 6.15. Only the eigenvalues from Table 0 (the detection table) are used for this suggested approach and none of the other tables are necessary.

2) Standardise or normalise each of the eigenvalues in this fault eigenvalue matrix to obtain the normalized fault eigenvalue matrix. This allows for both ease of interpretation and fairness.

3) Find the qualitative fault signature by comparing each of the fault eigenvalues (columns 2 to 18 of the fault eigenvalue matrix) to the normal eigenvalues (column 1 of the fault eigenvalue matrix) in a qualitative way. This implies that a + sign is used if a fault eigenvalue is larger than that of a normal eigenvalue and a − sign for the inverse with an 0 sign if no change is seen.

4) Identify the patterns of + and − signs to determine detection and isolation of faults. If a fault is

detected there will be at least one + or one − sign in the column of that fault. If a fault is isolable each of the entries of + and − signs will be unique compared to that of the other faults. This method is in accordance with the method suggested in [3].

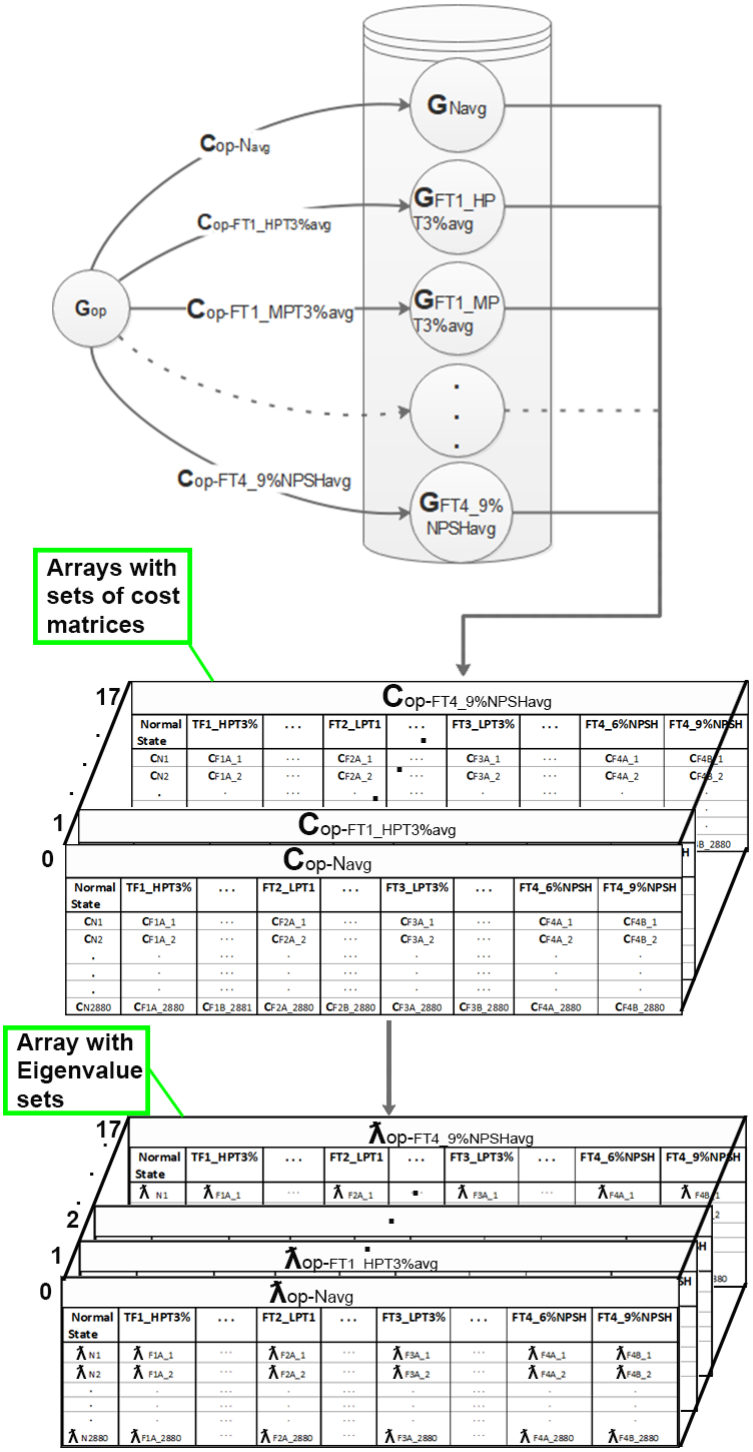


Figure 6.14: Illustration of how the eigenvalues are obtained and how the final eigenvalue dataset looks

In Figure 6.15 the normalised fault eigenvalues matrix are obtained by normalising the original eigenvalues using (6.2)

$$Z_i = \frac{x_i - \bar{x}}{\sigma}, i \in \{1, 2, ..., 2880\}, \tag{6.2}$$

with x_i the value that must be normalised, $\bar{x}$ is the average and σ the standard deviation of the set of values that is normalised. Since this is a statistical approach to normalising data the resulting range of the normalised data is not between 0 and 1 but rather between some maximum and minimum value dependant on the distribution of the dataset.

From Figure 6.15 it is clearly seen that the faults can be detected regardless of the intensity or location of the fault in the system. This is concluded since each of the columns of each fault contains one or more + or − signs (see point 3 from the aforementioned list of procedure). In terms of isolation it is seen that fault 1 is isolable regardless of the fault size and location. This is concluded by noting that the columns of the 3 cases for fault 1 have different qualitative patterns (see red + and − signs in columns 1 to 3 of Figure 6.15). The same can be said for fault 4. In terms of faults 2 and 3, however, only some of the variations in the faults are unique and can thus be isolated from any other variations of the same fault. FT's 2 and 3 can, however, still be isolated from the other fault types (FT 1 to 4) regardless of the variations in size and location of those fault types. In other words, FT 2 has a unique qualitative pattern in comparison with FT 1, 3 and 4, regardless of the size and locations of FT 2 but it can't always be isolated in terms of variation in fault size and location of fault 2 itself. The same applies to FT 3.

In order to determine the false detection and isolation rate, this pattern analysis must be repeated for all 2880 measurements in the data set. This proved to be time consuming and inefficient as can obviously be deducted from the qualitative fault signature matrix. To overcome this problem an alternative quantitative approach was applied. For this approach rather than comparing each of the substantive columns of the eigenvalue signature matrix with column one, the following was done:

1) Obtain the fault eigenvalue matrices as represented in Figure 6.14 for each row of each of the 18 tables (0 to 17). The fault eigenvalue matrices from Table 0 is used for fault detection. The fault eigenvalue matrices from Table 1 is used for isolation of fault 1, Table 2 for isolation of fault 2 and so forth until Table 17.

2) Rather than comparing each of columns 2 to 18 of the fault eigenvalue matrices with column 1, each of the columns is averaged to obtain a single value representing the size of that column. This is shown in Figure 6.16.

3) Evaluate each of the averages obtained and find the smallest average value. This value represents the system state that is most likely the state represented by the operational graph, when the cost matrix was calculated.

4) By simply noting which column has the smallest average value for each of the 2880 sets of eigenvalues in Table 0 of Figure 6.14 and summing the number of times that this was the case for column 1, a result smaller or equal to 2880 is obtained. This result is then represented as a percentage of 2880 to indicate the true detection rate. With this approach, rather than detecting a fault, it can be considered as trying to isolate the normal operational state.

5) Apply the same approach as in point 4 to each other eigenvalue set (Tables 1 to 17) in Figure 6.14 to determine the percentage of correct isolation of the faults. The resulting detection and isolation can be seen in Figure 6.17.

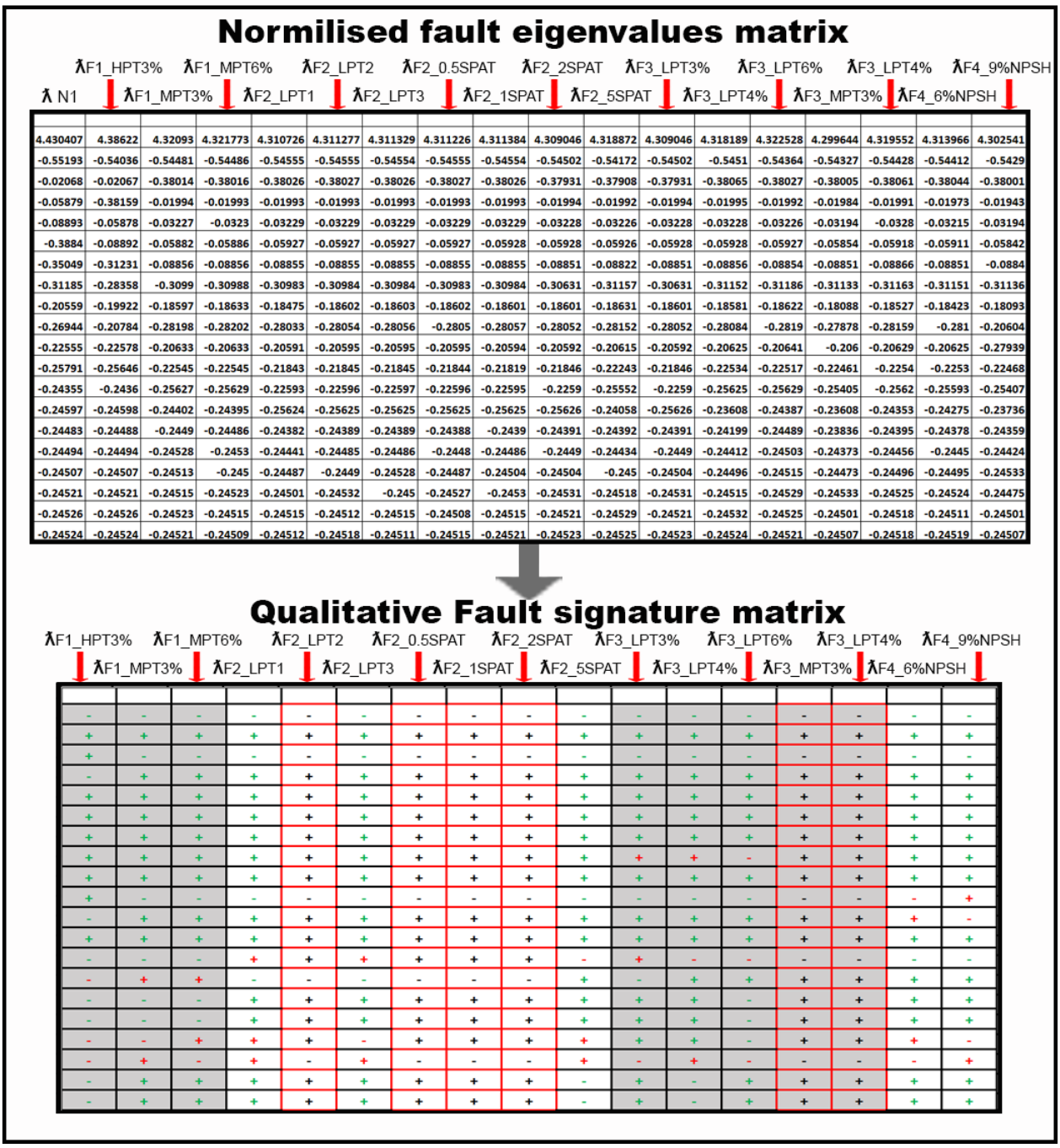


Figure 6.15: Example of the results obtained for the eigenvalue approach

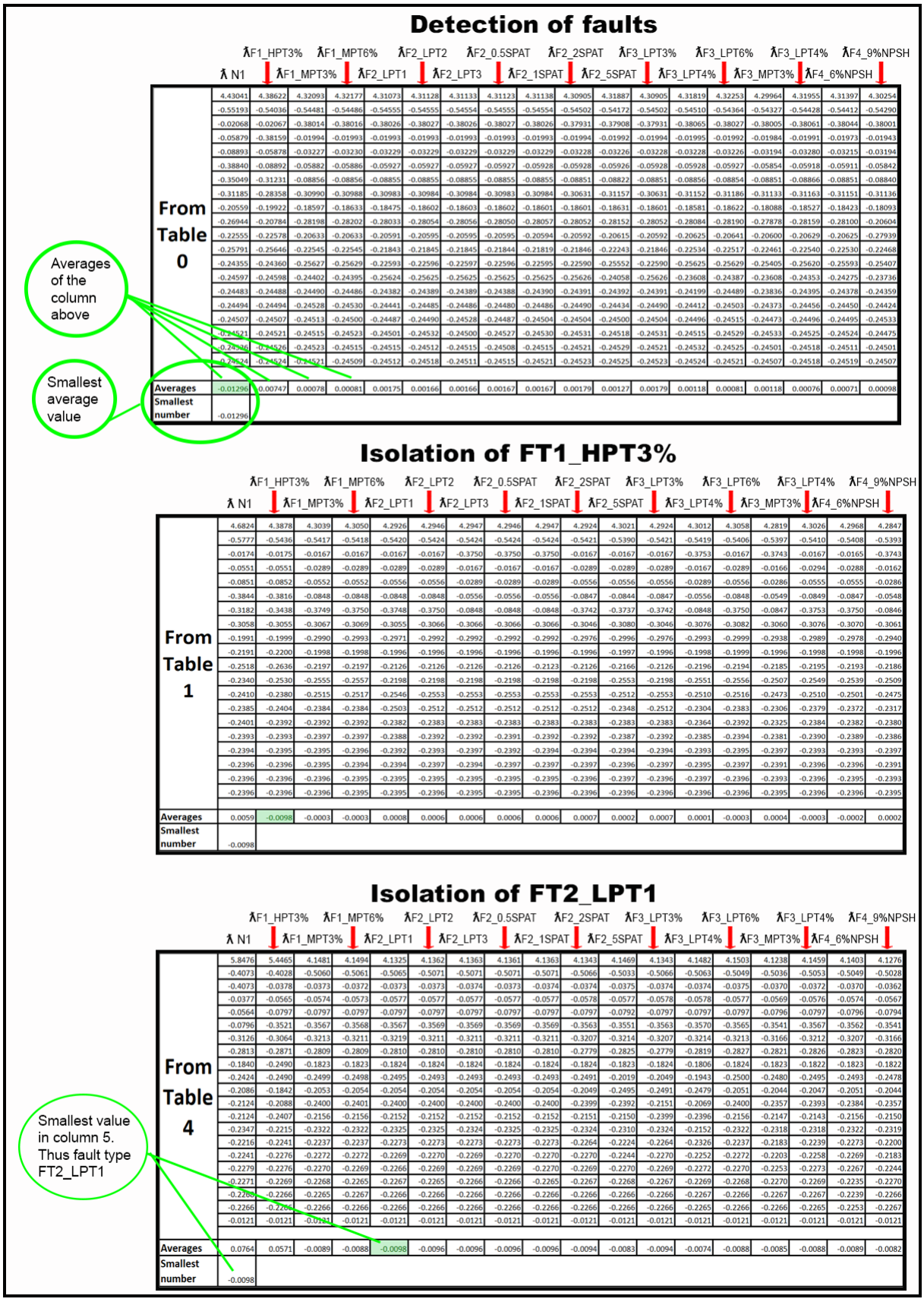


Figure 6.16: Results when the average size of each column of eigenvalues is calculated from the eigenvalue set.

Since the normal state is isolated 100% of the time it implies all faults can be detected

FT1MPT3% was correctly identified as FT1_MPT3% for 71% of the 2880 measurements taken. For 26% of the measurements it was wrongly identified as TF1_MPT6% and for 2.36% and 0.52% it was thought to be FT3_LPT6% and FT4_6%NPSH respectivly

		Summary	Normal Operation	FT1_ HPT3%	FT1_ MPT3%	FT1_ MPT6%	FT2_ LPT1	FT2_ LPT2	FT2_ LPT3	FT2_0, 5SPAT	FT2_ 1SPAT	FT2_ 2SPAT	FT2_ 5SPAT	FT3_ LPT3%	FT3_ LPT4%	FT3_ LPT6%	FT3_ MPt3%	FT3_ MPT4%	FT4_6 %NPSH	FT4_9 %NPSH
Detection	**Normal Operation**	100%	100%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Isolate fault	**FT1_HPT3%**	100%	0	100%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Isolate fault	**FT1_MPT3%**	71%	0	0	71%	26%	0	0	0	0	0	0	0	0	0	2.36%	0	0	0.52%	0
Isolate fault	**FT1_MPT6%**	68%	0	0	30%	68%	0	0	0	0	0	0	0	0	0	1.49%	0	0	0.07%	0
Isolate fault	**FT2_LPT1**	99.97%	0	0	0	0	99.97%	0	0	0	0	0.035%	0	0	0	0	0	0	0	0
Isolate fault	**FT2_LPT2**	32%	0	0	0	0	0	32%	10%	8%	17%	21%	12%	0	0	0	0	0	0	0
Isolate fault	**FT2_LPT3**	34%	0	0	0	0	0	7%	34%	9%	17%	21%	12%	0	0	0	0	0	0	0
Isolate fault	**FT2_0,5SPAT**	36%	0	0	0	0	0	7%	14%	36%	10%	21%	12%	0	0	0	0	0	0	0
Isolate fault	**FT2_1SPAT**	36%	0	0	0	0	0	10%	15%	6%	36%	21%	12%	0	0	0	0	0	0	0
Isolate fault	**FT2_2SPAT**	48%	0	0	0	0	0	4%	9%	4%	9%	48%	26%	0	0	0	0	0	0	0
Isolate fault	**FT2_5SPAT**	44%	0	0	0	0	0	4%	9%	4%	9%	30%	44%	0	0	0	0	0	0	0
Isolate fault	**FT3_LPT3%**	100%	0	0	0	0	0	0	0	0	0	0	0	100%	0	0	0	0	0	0
Isolate fault	**FT3_LPT4%**	100%	0	0	0	0	0	0	0	0	0	0	0	0	100%	0	0	0	0	0
Isolate fault	**FT3_LPT6%**	95%	0	0	4%	1%	0	0	0	0	0	0	0	0	0	95%	0	0	0	0
Isolate fault	**FT3_MPt3%**	100%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%	0	0	0
Isolate fault	**FT3_MPT4%**	99%	0	0	0.10%	0.03%	0	0	0	0	0	0	0	0	0	0	0	99%	0.63%	0
Isolate fault	**FT4_6%NPSH**	99%	0	0	0.38%	0	0	0	0	0	0	0	0	0	0	0	0	0.42%	99%	0
Isolate fault	**FT4_9%NPSH**	100%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%

Figure 6.17: Eigenvalue approach-Detection and isolation rate results

When considering the results in Figure 6.17 it is clear that all faults are detected for 100% of the 2880 measurements. In terms of isolation all four fault types were 100% of the 2880 measurements correctly isolated but only under specific fault conditions. This is especially seen when considering fault types 1 and 2. The solid particle erosion (FT 1) can easily be isolated for the high-pressure turbine (FT1_HPT3%). For the medium pressure turbine in the 3% case of SPE (FT1_MPT3%), only 71% of the 2880 measurements were correctly identified as FT1_MPT3%. The other 29% of the 2880 measurements were wrongly identified to be either FT1_MPT6% (26%), FT3_LPt6% (2.4%) or FT4_6%NPSH (0.52%) respectively. Similarly, for an increase in SPE in the MPT the isolation decreased to 68%. For 30% of the remaining 32% isolation, the technique wrongly identified FT1_MPT6% as FT1_MPT3%. Since FT1_MPT3% is the same fault as FT1_MP6%, smaller in intensity but still within the same component in the system, it is reasonable that some overlapping of the data would occur. This explains the 30% missed isolation. The remaining 2% isolation is attributed to wrongly identifying fault types FT3_LPT6% and FT4_6%NPSH as FT1_MPT6%. This suggests that an increase in SPE in the medium pressure turbine can sometimes have similar effects, in terms of energy changes in the system, as seen with overall wear/ageing and cavitation. These results for FT 1, suggest that the technique is well adjusted in terms of sensitivity and robustness since a small change in the fault intensity was clearly detected and a change in the fault location was also flagged for 100% of the times measured.

Following the same reasoning for FT 2, it can be concluded that a leak in the overflow valve of the first stage of the LPT can easily be isolated. However, the isolation decreased significantly for leaks with other intensity and/or in other parts of the system. For example, a leak in the overflow valve of the second LPT stage was correctly isolated for only 32% of the 2880 measurements. For 21% of the measurements the leak was identified as a leak in the SPAT under the conditions where its overflow valve had a leak resulting in 0.5 kg/s loss in mass flow rate. These values imply that the technique is not robust enough to effectively differentiate between overflow valve leaks in different parts of the system (between 32% and 44% of the time the technique will be accurate). Similarly, the technique failed to differentiate between this type of fault when different intensities of the fault occurred, only being accurate between 36% and 44% of the time. Thus, reasonable sensitivity to overflow valve leak size changes does exist, but would still result, at most, in approximately 54% incorrect identification of the fault size and/or location. The interestingly accurate isolation seen for FT2_LPT1 relative to the rest of the variations in FT two, can be attributed to the following facts: Firstly, stage one of the LPT is fed with significantly higher quality steam than the subsequent stages. It is therefore reasonable to argue that a leak in the overflow valve of the first stage of the LPT would result in a significantly higher change in the energy flow of the system than the subsequent stages. Secondly a leak in the overflow valve of stage one of the LPT directly results in a lower mass flow rate of high-quality steam available for one of the pre-heaters in the system, fed from a partial stream of the LPT first stage. Since the heat transferred from this fluid stream to the pre-heater is approximately double that of the subsequent LPT stages, the effects of a leak in the overflow valve of stage 1 is more intense and can therefore be more effectively isolated since larger energy changes occurs. This pre-heater also has an impact on the energy of the water that is fed to the SPAT pump. This might explain the small confusion in distinguishing between FT2_LPT1 and FT2_2SPAT. Regarding the isolation of FT 3 and 4, not much discussion is required. The small inaccuracy in the isolation of FT FT3_LPT6% can possibly be attributed to the fact that this fault represents overall wear and ageing in the medium pressure turbine resulting in a 6% overall drop in the efficiency of the whole medium pressure turbine. On the other hand, FT1_MPT3% and FT1_MPT6% represent a drop in efficiency and mass flow rate due to solid particle erosion. Since both these faults affects the turbine's efficiency it is reasonable to expect some overlap in the energy changes seen under these two fault conditions, which explains the 4% and 1% incorrect isolation results.

The next section will be dedicated to the residual approach as the third and final approach evaluated in this study. Since this approach directly evaluates the attributed graphs it is expected that the preservation of the structural information will positively contribute to the success of the FDI approach especially in terms of fault isolation.

6.5 Approach 3 (Residual Approach)

As mentioned in the overview of this chapter, the residual approach was suggested by [78] as a method of evaluating the change in the attributed graphs when a fault occurred. The idea of using residual based methods in FDI is common practise and would be, for this study, an obvious candidate for evaluation. In a fault free or normal operation, the residuals would ideally be zero. However, due to the fact that some variation in the normal operating condition exists, this would not be the case for this study. For fault detection, however, it can be argued that, for any of the normal operational graphs, compared to a normal state reference graph, the residuals will be small in comparison to a residual calculated for any fault operational graph compared to the normal reference graph. In terms of isolation of a fault the same argument can be made for a fault operational graph compared to that same fault's reference graph. The residuals are also expected to have a distinct pattern for each of the fault conditions. The approach by [78] suggested a direct index by index comparison of the attributed graphs. Equation (6.3) illustrates this idea for a residual graph $\mathbf{G}_{res}$ calculated by comparing an operational graph $\mathbf{G}_{op}$ to a reference graph $\mathbf{G}_{ref}$.

$$\mathbf{G}_{res} = \frac{g_{ref(i,j)} - g_{op(i,j)}}{g_{ref(i,j)}}, \tag{6.3}$$

with, $i \in \{1,2,...,n\}$ and $j \in \{1,2,...,m\}$ for all the graphs representing a $n \times m$ matrix. The division with $\mathbf{G}_{ref}$ has the effect of normalizing the residual matrix with respect to the specific attribute. This implies that for any of the attributed graphs obtained in this study the residual graph will have the form,

$$\mathbf{G}_{res} = \begin{pmatrix} \frac{\Delta\dot{b}_{ref1} - \Delta\dot{b}_{op1}}{\Delta\dot{b}_{ref1}} & | & \frac{\dot{q}_{ref(1,1)} - \dot{q}_{op(1,1)}}{\dot{q}_{ref(1,1)}} & \cdots & \frac{\dot{q}_{ref(1,20)} - \dot{q}_{op(1,20)}}{\dot{q}_{ref(1,20)}} \\ \frac{\Delta\dot{b}_{ref2} - \Delta\dot{b}_{op2}}{\Delta\dot{b}_{ref2}} & | & \frac{\dot{q}_{ref(2,1)} - \dot{q}_{op(2,1)}}{\dot{q}_{ref(2,1)}} & \cdots & \frac{\dot{q}_{ref(2,20)} - \dot{q}_{op(2,20)}}{\dot{q}_{ref(2,20)}} \\ . & | & . & \cdots & . \\ . & | & . & \cdots & . \\ . & | & . & \cdots & . \\ \frac{\Delta\dot{b}_{ref20} - \Delta\dot{b}_{op20}}{\Delta\dot{b}_{ref20}} & | & \frac{\dot{q}_{ref(20,1)} - \dot{q}_{op(20,1)}}{\dot{q}_{ref(20,1)}} & \cdots & \frac{\dot{q}_{ref(20,20)} - \dot{q}_{op(20,20)}}{\dot{q}_{ref(20,20)}} \end{pmatrix},$$

with some of the indexes equal to zero depending on the node signature graphs.

In order to determine the amount of variation in the residual of the operational graphs for a specific state, two basic methods can be followed. The first would be to use the reference graphs for a state as calculated for approach 1 and 2. Thus, a reference graph is composed with the average energy attributes as illustrated in Figure 6.6. When the operational graph is then composed for each of the system fault states and the normal state respectively, and the residual is calculated, for example, when the operational graphs are compared to the average normal reference, the results obtained will indicate the residual difference from this average reference. The resulting range will thus be spread around the average of the reference state considered, with the maximum residual value given by either the distance between the average and the maximum or the average and the minimum value (see Figure 6.18 for a graphic illustration). The second method increase the size of this maximum residual distance value by rather using the maximum energy attributes to compose

a reference graph for each of the operational states. The result of this is that the size of the residual values obtained for any of the states, compared to any of the references, are maximised. This allows for larger residual values that leads to easier visual interpretation when the data are analysed. In Figure 6.18 the two methods are illustrated with 5 random data values. These data values can be thought of as energy attributes, like $\Delta\dot{b}_{op1}$ or $\dot{q}_{op(1,1)}$, in the system that was measured at a specific location while the system is operating in a specific state (normal or faulty). The figure clearly indicates that using the maximal value method would maximize the residual values obtained, since the residual obtained when comparing the largest value (measurement 2) with the smallest value (measurement 4), resulted in a larger residual value than when comparing the average value (measurement 3) with the largest or smallest values (measurements 2 and 4) as shown for the average value method.

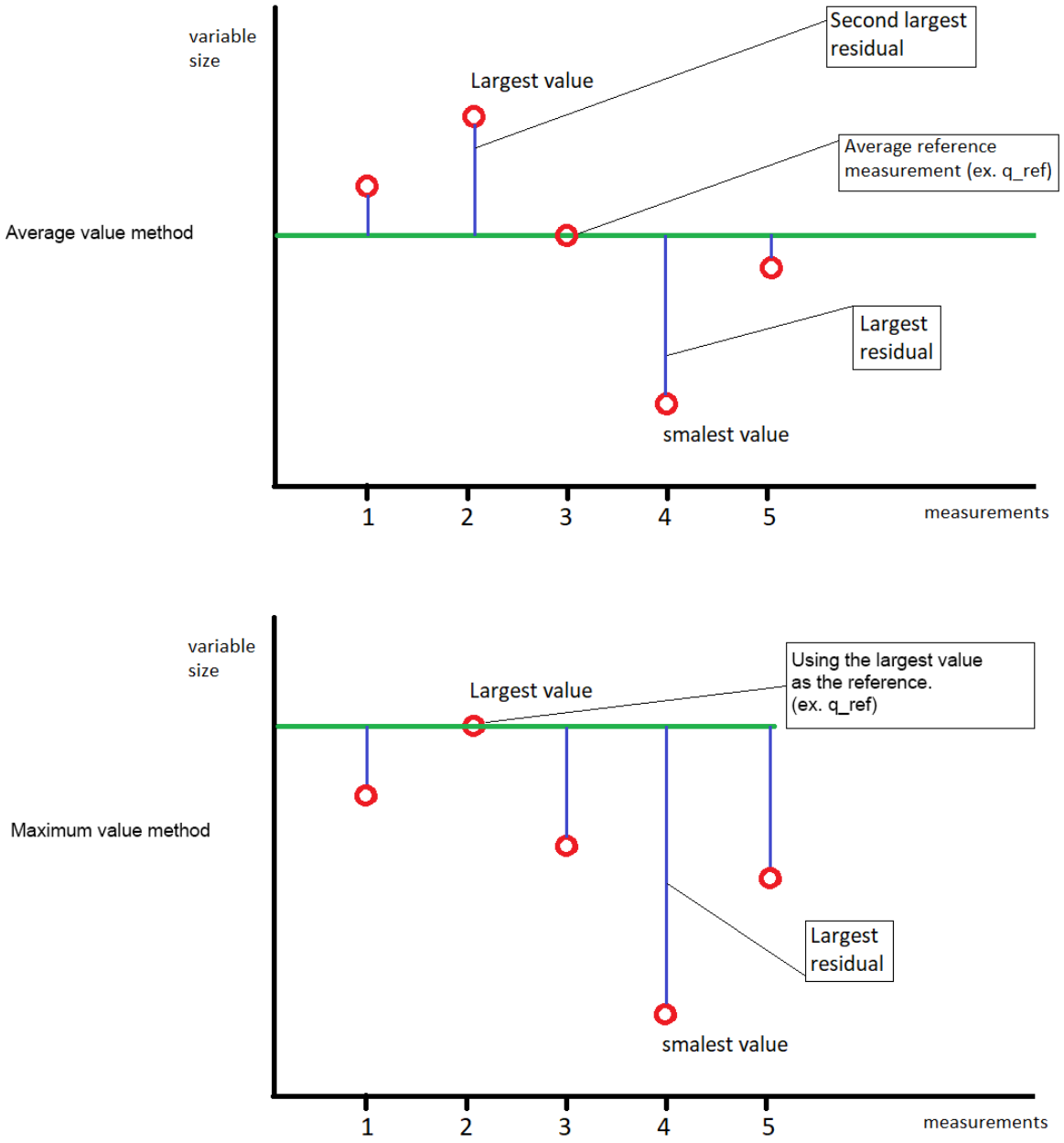
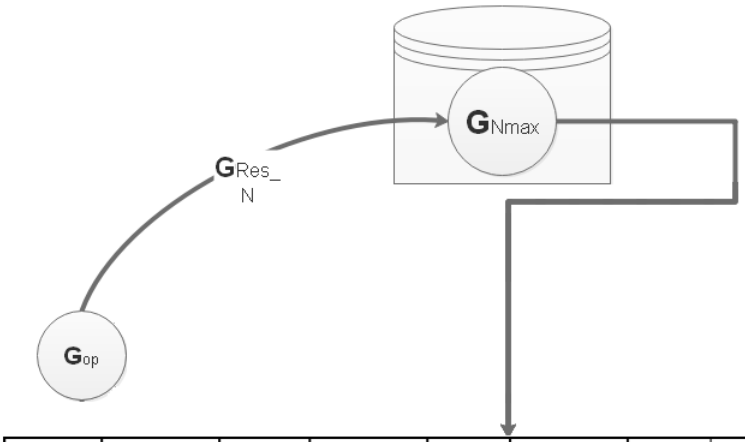


Figure 6.18: Illustration of residual value's size difference when using the average or maximum values as reference respectively

Through the use of the maximum value method in Figure 6.18, it is ensured that the residual result calculated for an operational graph, compared to a reference graph, would be maximized.

This will apply to any of the state graphs compared to the any of the reference graphs. By comparing each of the operational graphs to the normal reference graph, and calculating the residual matrices the following sets of data are obtained as shown in Figure 6.19. This can be done for any of the reference graphs of any fault state as well.



Normal State	TF1_HPT3%	...	FT2_LPT1	...	FT3_LPT3%	...	FT4_6%NPSH	FT4_9%NPSH
G_{Res_N1}	$G_{Res\_F1A\_1}$	...	$G_{Res\_F2A\_1}$	...	$G_{Res\_F3A\_1}$	...	$G_{Res\_F4A\_1}$	$G_{Res\_F4B\_1}$
G_{Res_N2}	$G_{Res\_F1A\_2}$	...	$G_{Res\_F2A\_2}$	...	$G_{Res\_F3A\_2}$	...	$G_{Res\_F4A\_2}$	$G_{Res\_F4B\_2}$
.	.	...	.	...	.	...	.	.
.	.	...	.	...	.	...	.	.
.	.	...	.	...	.	...	.	.
G_{Res_N2880}	$G_{Res\_F1A\_2880}$	...	$G_{Res\_F2A\_2880}$	...	$G_{Res\_F3A\_2880}$	...	$G_{Res\_F4A\_2880}$	$G_{Res\_F4B\_2880}$

Figure 6.19: Illustration of the residual data set obtained

The normal reference graph $\mathbf{G}_{Nmax}$ used for this study is shown in Table 6.3 accompanied by one of the residual graphs obtained for the normal operational state $\mathbf{G}_{ResN1}$ compared to the normal reference graph. Similar graphs exist for each of the 2880 normal states and fault states compared to the normal and fault state reference graphs. Figure 6.19 only illustrate the results for the operational graphs compared to the normal reference graph for illustrative purposes. In order to use the results for fault detection and isolation the residual data should be normalized with regards to the attributes represented by each column of the graph. The method suggested by [78] was simply to divide the first column of a residual graph, representing the residuals of the change in exergy, by the maximum value of that column. Similarly, the values in columns 2 to 21 is divided by the maximum value of all the residual values representing the heat transfer. The equation used for the normalising of the exergy residuals is given in (6.4) and for the heat flow residuals is given in (6.5).

$$g_{res(i,1)} = \frac{g_{res(i,1)}}{max|g_{res(i,1)}|}, \tag{6.4}$$

with $i \in \{1,2,...,m\}$ and m the number of columns in the residual graph.

$$g_{res(i,j)} = \frac{g_{res(i,j)}}{max|g_{res(i,j)}|}, \tag{6.5}$$

with $i \in \{1,2,...,m\}$, $j \in \{2,...,n\}$ and m the number of columns and n the number of rows in the residual graph.

Table 6.3: Reference graph $\mathbf{G}_{Nmax}$ and residual graph $\mathbf{G}_{ResN1}$

Normal reference graph G_Nmax

5.2925e+05	0	1.5286e+06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.2582e+05	0	0	1.3869e+06	1.3545e+05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.2214e+05	0	0	0	0	0	0
8.4327e+04	0	0	0	1.4380e+06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.7704e+04	0
9.4407e+04	0	0	0	0	1.3799e+06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6.3247e+04	0	0	0	0	0	1.1672e+06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.3904e+05	0
9.5329e+04	0	0	0	0	0	0	1.0734e+06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.0013e+04	0
5.8391e+04	0	0	0	0	0	0	0	1.0382e+06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8.7891e+04	0	0	0	0	0	0	0	0	9.4354e+05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6.5901e+04	0
8.7891e+04	0	0	0	0	0	0	0	0	0	8.1749e+05	0	0	0	0	0	0	0	0	0	0	0	0	0	3.4454e+04	0
8.7891e+04	0	0	0	0	0	0	0	0	0	0	1.8734e+05	7.1407e+05	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	6.6004e+05	4.6890e+04	0	0	0	0	0	0	0	0	0	6.2933e+05	0
7.3579e+04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5.5865e+04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.1394e+07	0
-562.5179	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5.1653e+04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.7957e+04	0	0	0	2.6507e+05	0	2.1711e+05	0
1.3386e+04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.1708e+04	0	1.0948e+05	0
1.0357e+04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.4179e+05	0
517.6693	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.5168e+05	0	1.4364e+04	0
7.4913e+04	5.1503e+05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.7323e+05	0
0	-1.0124e+06	0	-2.7678e+05	0	0	0	0	0	0	0	0	0	0	1.5009e+06	6.3869e+03	2.1496e+05	0	0	7.1529e+04	1.7152e+05	0				

Residual graph G_resN1

-0.1060	0	1.8182e-05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.3699	0	0	1.4221e-05	2.9866e-04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-0.7959	0	0	0	2.0396e-05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.1738e-06	0	0	0	0	0	0
0.1777	0	0	0	0	5.3136e-06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.2707e-04	0
0.2653	0	0	0	0	0	6.2821e-06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.2738e-05	0
0.1760	0	0	0	0	0	0	6.8313e-06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0
0.2874	0	0	0	0	0	0	0	1.0256e-05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.1422	0	0	0	0	0	0	0	0	0.0099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0
0.1422	0	0	0	0	0	0	0	0	0	0.0099	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0
0.1422	0	0	0	0	0	0	0	0	0	0	0.0099	1.9987e-05	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	3.0596e-04	0	0	0	0	0	0	0	0	0	0.1229	0
2.9288e-04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0	0	0	0	0	0	0	0
-3.8568e-04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0	0	0	0	0	0.0099	0
-0.0101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0	0	0	0	0	0
-0.0101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0	0.0099	0
0.0476	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.3109e-05	0	0	7.1987e-10	0
0.0099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0	0
0.0099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0	0.0099	0
-0.0101	0.0099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0099	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

When (6.4) and (6.5) are applied the resulting matrix is,

$$\mathbf{GN}_{res(State\_x)} = \left(\begin{array}{c|c|ccc} \frac{g_{res(1,1)}}{max|g_{res(1,1)}|} & g\frac{g_{res(1,2)}}{max|g_{res(1,2)}|} & \cdots & \frac{g_{res(1,20)}}{max|g_{res(1,20)}|} \\ \frac{g_{res(2,1)}}{max|g_{res(2,1)}|} & \frac{g_{res(2,2)}}{max|g_{res(2,2)}|} & \cdots & \frac{g_{res(2,20)}}{max|g_{res(2,20)}|} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \frac{g_{res(20,1)}}{max|g_{res(20,1)}|} & \frac{g_{res(20,2)}}{max|g_{res(20,2)}|} & \cdots & \frac{g_{res(20,20)}}{max|g_{res(20,20)}|} \end{array}\right),$$

with $State\_x$ the $x'th$ number of the normal or faulty states graph for which the residual is calculated.

The method for visualizing the detection of a fault as suggested by [78] for an attributed graph with 6 nodes, is to compose a bar plot of the residual heat flow in and out of a node (R_{Qin} and R_{Qout}) along with the residual of the exergy change ($R_{\Delta d}$) and the heat transferred from and to the environmental node (R_{Qext})). The pattern encapsulated within each bar plot are then analysed to determine if the bar plot represents a normal or fault state. This allows for detecting faults as well as isolating then by only using the residual graphs obtained from comparing the operational graphs with the normal reference graph. In this case the other residual graphs obtained when the operational graphs are compared to the fault reference graphs is not used. However, due to a different system and a resulting different attributed graph with 20 nodes, this method would result in ineffectively large graphs that won't be practical for interpretation. The visualisation was therefore obtained by simply generating a bar plot of each of the columns of a residual graph. The result of this is that the residual bar plots for the variables R_{Qin}, R_{Qout}, $R_{\Delta d}$ and R_{Qext}) are plotted onto each other. By making the bars transparent the size of each variable at each node can be seen but without an indication of what variable are attributed to each of the transparent bars. This method therefore only gives an indication of the size of each energy attribute at each node and how it changes when the system conditions changes form normal to faulty, but does not identify the individual variables. If the four variables (R_{Qin}, R_{Qout}, $R_{\Delta d}$ and R_{Qext})) are to be identified, one would have to analyse the respective residual graphs along with the bar plots. Regardless, the resulting bar plots can be viewed as a signature of the system state. These bar plot signatures are therefore only used as a detection and isolation method based on analysing the patterns of the individual bar plots for each state. In Figures 6.20, 6.21 and 6.22 the residual signatures obtained for the operational graph composed to two measurements of the normal state, FT1_HPT3% and FT2_LPT1, and compared to the normal reference graph, are shown as an example. Similar graphs for each of the fault states and the variation of each fault can be seen in appendix C section C.1.2. Note that in these bar plots the node number, "20" represents the environmental node indicated as node "0" in the attributed graph of Figure 4.4. The bar labels named "Db" represents the first column of the residual graphs, hence indicating the residual exergy values. Similarly, the bar labels named "col1" till "col20", represents the second till last columns of the residual graphs, hence indicating the heat flow residual values.

When analysing the graphs, a clear pattern can be observed. From Figure 6.20 node 3, 13, 16 and the environmental node 20 shows a distinct pattern. However, due to variation in the normal operational state, the sizes of these residual values may vary for each of the 2880 residual graphs.

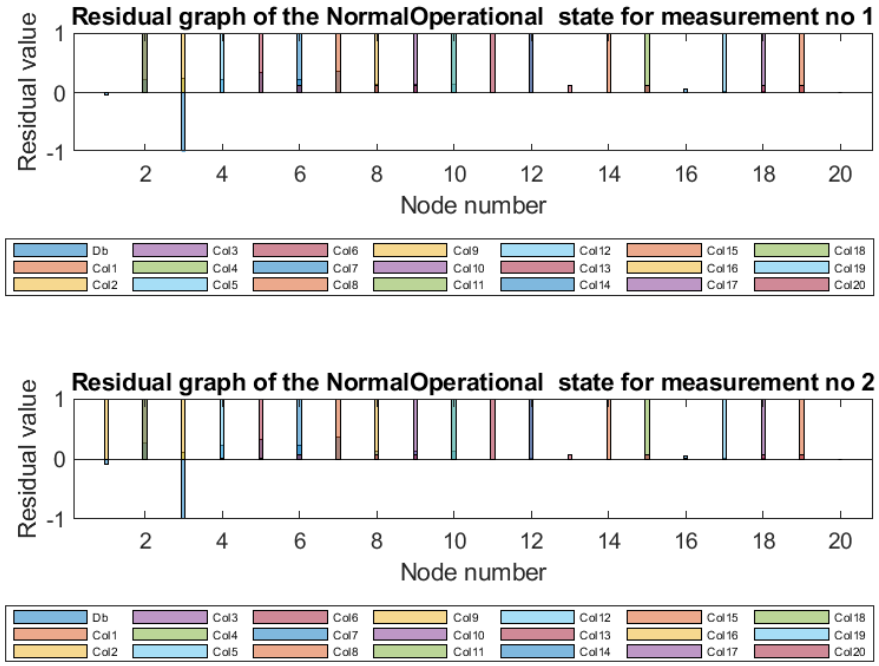


Figure 6.20: Residual fault signatures obtained for 2 of the normal graphs

Similarly, to the normal state signature, the signature of the first variation of fault 1 shows a distinct pattern. Nodes 3, 12, 13, 16, and 20 can especially be attributed to this fault. Once again variations in the residual values can be seen. This same approach is followed when examining the residual signatures of the other faults.

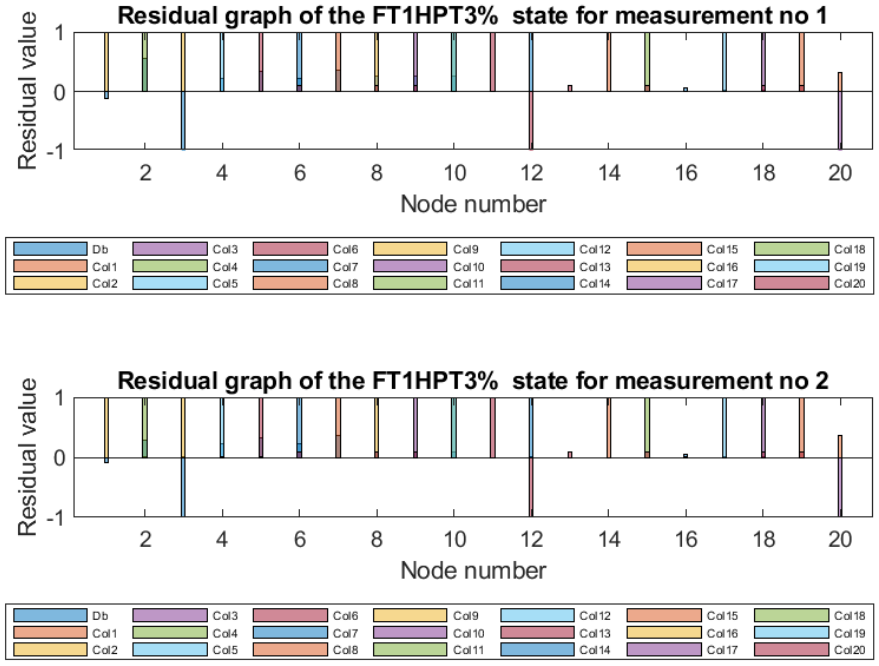


Figure 6.21: Residual fault signatures obtained for 2 of the graphs of FT1_HPT3%

From the figures in appendix C section C.1.2 the other results for variations in intensity and location of fault 2 can be seen. When comparing these residual fault signatures with each other, it is clear that fault 2 does not have a distinct signature that remains the same when the size and location of the fault are changed. To obtain the range of variation of these residual values each of the signatures for all 2880 graphs of each state must be analysed. This is also required in order to determine false detection and isolation.

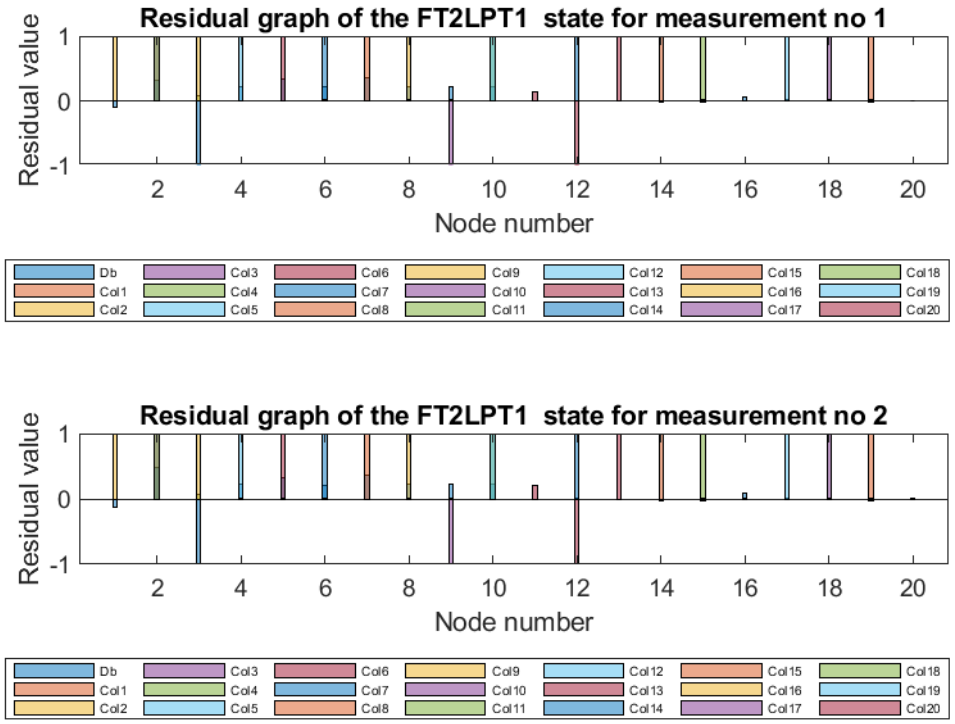


Figure 6.22: Residual fault signatures obtained for 2 of the graphs of FT2_LPT1

Since a graphic inspection of all of the graphs would be inefficient and time consuming a qualitative approach was followed to determine the number of residual signatures that matches a specific pattern associated with a specific state. The procedure can be described as follows:
1) Obtain the average and standard deviation of the first column of each of the normalized residual matrices. This will represent the average and standard deviation of each of the nodes' exergy change residuals (Δb_{RESavg} and Δb_{RESstd}).
2) Obtain the average and standard deviation for the range from column 2 till the end of each of the residual matrices for each state. This represents the average and standard deviation of all of the heat transfer values (Q_{RESavg} and Q_{RESstd}.
3) Run the following "if" statement to determine the number of values from each column that falls within two standard deviations from the mean.
If an exergy residual in column one is larger than $\Delta b_{RESavg} + 2 \times \Delta b_{RESstd}$ then change that variable to 1. Otherwise, if the variable is smaller than $\Delta b_{RESavg} - 2 \times \Delta b_{RESstd}$ change the variable to 1, otherwise change the variable to 0. Repeat this for all of the heat transfer residuals using Q_{RESavg} and Q_{RESstd}. This forms a new quantitative residual signature containing only zero's and one's.
4) When this is done for all of the normalized residual signatures of each state, the number of one's in each column is counted and 18 sets of 2880 frequency vectors containing a single value for each exergy and heat flow residual is obtained. Select one of the qualitative residual signatures for each state as a quantitative residual reference signature. Hence a total of 18 quantitative residual

reference signatures is obtained, one representing each fault state and the normal state.
5) Compare all of the quantitative residual signatures (2880 in total) for each state (thus 18 x 2880 total) to each of the qualitative residual references. The comparison can simply be done by determining if the a frequency vector is identical to the reference and returning a Boolean result.
6) By simply counting the number of true results obtained from the Boolean comparison for each of the states the number of total correctly identified signatures are obtained along with the number of unidentified signatures.

These 6 steps can be explained at hand of the flow diagram in Figure 6.23 and continued in Figure 6.24. An enlarged table of the final results is shown in Figure 6.25.

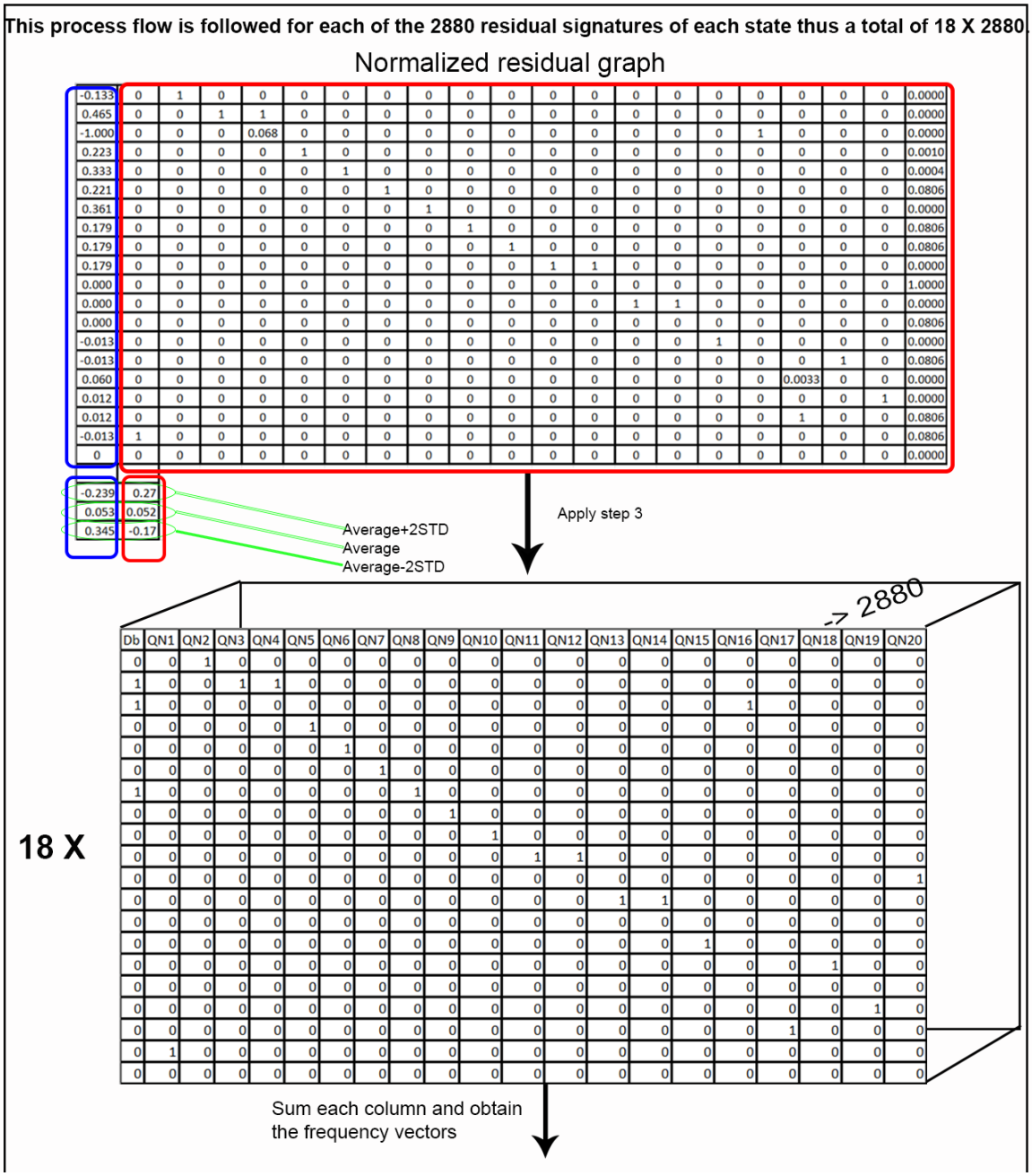
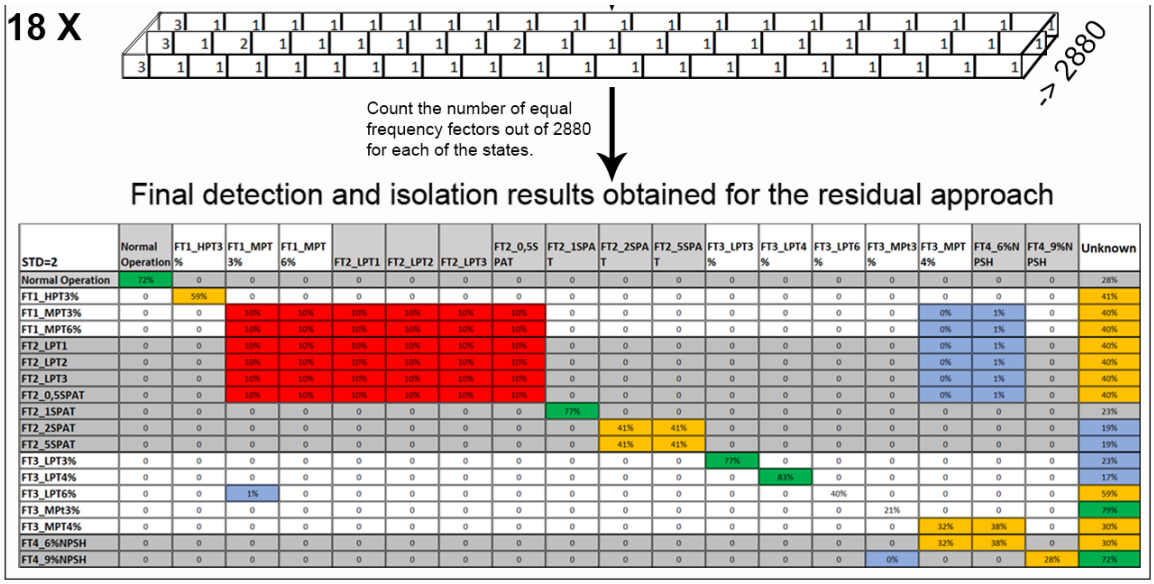


Figure 6.23: Process flow for determining the isolation rate of a state for the residual signatures.



STD=2	Normal Operation	FT1_HPT3%	FT1_MPT3%	FT1_MPT6%	FT2_LPT1	FT2_LPT2	FT2_LPT3	FT2_0,5SPAT	FT2_1SPAT	FT2_2SPAT	FT2_5SPAT	FT3_LPT3%	FT3_LPT4%	FT3_LPT6%	FT3_MPt3%	FT3_MPT4%	FT4_6%NPSH	FT4_9%NPSH	Unknown
Normal Operation	72%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28%
FT1_HPT3%	0	59%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	41%
FT1_MPT3%	0	0	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0%	1%	0	40%
FT1_MPT6%	0	0	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0%	1%	0	40%
FT2_LPT1	0	0	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0%	1%	0	40%
FT2_LPT2	0	0	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0%	1%	0	40%
FT2_LPT3	0	0	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0%	1%	0	40%
FT2_0,5SPAT	0	0	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0%	1%	0	40%
FT2_1SPAT	0	0	0	0	0	0	0	0	77%	0	0	0	0	0	0	0	0	0	23%
FT2_2SPAT	0	0	0	0	0	0	0	0	0	41%	41%	0	0	0	0	0	0	0	19%
FT2_5SPAT	0	0	0	0	0	0	0	0	0	41%	41%	0	0	0	0	0	0	0	19%
FT3_LPT3%	0	0	0	0	0	0	0	0	0	0	0	77%	0	0	0	0	0	0	23%
FT3_LPT4%	0	0	0	0	0	0	0	0	0	0	0	0	83%	0	0	0	0	0	17%
FT3_LPT6%	0	0	1%	0	0	0	0	0	0	0	0	0	0	40%	0	0	0	0	59%
FT3_MPt3%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21%	0	0	0	79%
FT3_MPT4%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32%	38%	0	30%
FT4_6%NPSH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32%	38%	0	30%
FT4_9%NPSH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	0	0	28%	72%

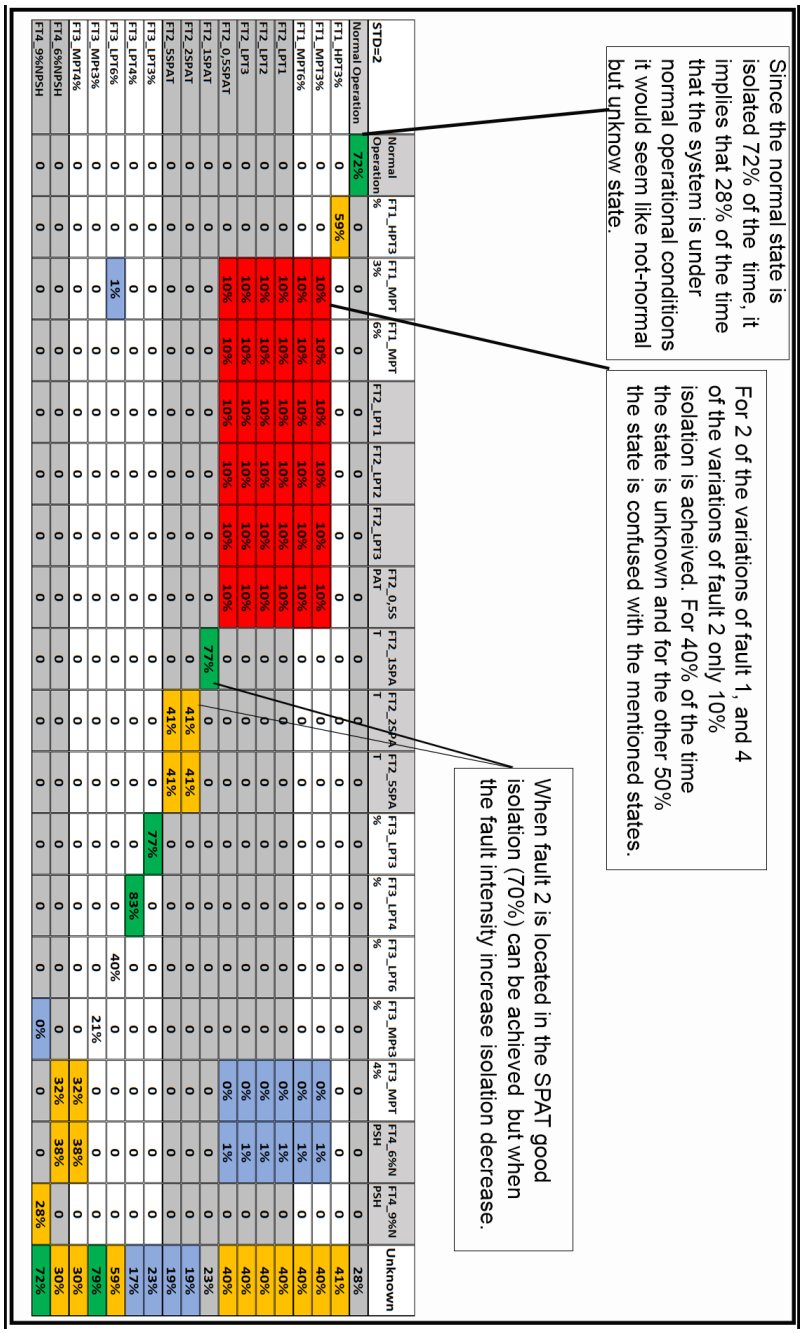
Figure 6.24: Process flow continue

When analysing the results in Figure 6.25 it is clear that this approach did not perform as well as the eigenvalue approach. It is noted that only 72% of the normal state measurements were identified due to the fact that it had the same residual signature. This number can easily be explained by referring to the fact that the frequency vector was generated using an interval consisting of only 2 standard deviations from the mean (see point 3). Enlarging the number of standard deviations to 3 or 4 standard deviations, would increase the number of identified normal states, but it will also increase the false identification of faults. Thus, the effectiveness of the technique will decrease.
In terms of FT 1, only when the SPE occurred in the HPT the fault could be isolated for 59% of the time, while 49% of the time the state was unidentifiable. For the SPE in the MPT the method could not differentiate well between changes in the size of fault 1 or some of the variations of fault 2. Only for 10% of the time the fault could be isolated. Some improvement in isolation was seen for fault 2 when the valve leak was simulated in the SPAT. The isolation effectiveness decreased for an increase in the fault size in the leak of the overflow valve in the SPAT. This might be due to the fact that an increased leak in the overflow valve from 2kg/s to 5kg/s is still the same fault and thus will result in the energy changes in the system to appear in the same locations. This explains the confusion between the changes in the fault size.

For fault 3, only two of the variations could be isolated more than 50% of the time for the LPT. An increase in the overall wear and ageing in the LPT gave inconclusive isolation. For wear and ageing in the MPT the approach gave even worse results and also could not be isolated from fault 4 when the ageing and wear were increased in the MPT. In summary, the method performed worse than approach 2 but better than approach 1.

Note that these results are based ons using 2 standard deviations above and below the average to of the residual matrices. When using 0.5, 1, 1.5 or 3 standard deviations the results changes. It was, however, found that 2 standard deviations gives the best results. It is also necessary to note that the results are based on analysing the patterns of the bar plot signatures, using only the residual results when the operational graphs were compared to the normal reference graph. The isolation might increase if analysing residual bar plot signatures obtained from comparing the operational graphs to the respective fault state reference graphs. This implies that the described process must be repeated for each of the sets of residual graphs obtained from comparing the operational graph with the different fault state reference graphs. The "result" table will be similar to that shown in

Figure 6.25 but with the expectation that each FT will be detected when the reference graph of that specific FT is used to calculate the residuals. Since a similar approach as described was followed, it would be repetitive to give a detail description of the process and therefore it is not described in this dissertation. The results obtained can be found by following the link in appendix C section C.1.2. The results however did not improve the isolation of any fault while even slightly decreasing the correct isolation in some cases. The reason for this can be attributed to the range of variation in the energy values of the system when a fault occurs as well as the system noise. As mentioned by [78] the variation in the energy values of each fault due to noise should not over lap the variations of other faults. If this occurs, isolation would not be able. From the results obtained for this approach and approach 1 it is, however, clear that large overlaps in the energy value of most faults occurs thus explaining the ineffective isolation.



STD=2	Normal Operation	FT1_HPT3%	FT1_MPT3%	FT1_MPT5%	FT2_LPT1	FT2_LPT2	FT2_LPT3	FT2_0.5SPAT	FT2_1SPAT	FT2_2SPAT	FT2_5SPAT	FT3_LPT3%	FT3_LPT4%	FT3_LPT6%	FT3_MPT3%	FT3_MPT4%	FT4_6%NPSH	FT4_9%NPSH	Unknown
Normal Operation	72%	59%	3%	6%											4%		PSH	PSH	
FT1_HPT3%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28%
FT1_MPT3%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	41%
FT1_MPT6%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40%
FT2_LPT1	0	10%	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0	0	0	40%
FT2_LPT2	0	10%	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0	0	0	40%
FT2_LPT3	0	10%	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0	0	0	40%
FT2_0.5SPAT	0	10%	10%	10%	10%	10%	10%	10%	0	0	0	0	0	0	0	0	0	0	40%
FT2_1SPAT	0	0	0	0	0	0	0	0	77%	0	0	0	0	0	0	0	0	0	23%
FT2_2SPAT	0	0	0	0	0	0	0	0	0	41%	41%	0	0	0	0	0	0	0	19%
FT2_5SPAT	0	0	0	0	0	0	0	0	0	41%	41%	0	0	0	0	0	0	0	19%
FT3_LPT3%	0	0	0	0	0	0	0	0	0	0	0	77%	0	0	0	0	0	0	23%
FT3_LPT4%	0	0	0	0	0	0	0	0	0	0	0	0	83%	0	0	0	0	0	17%
FT3_LPT6%	0	0	1%	0	0	0	0	0	0	0	0	0	0	40%	0	0	0	0	59%
FT3_MPT3%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21%	0	0	0	79%
FT3_MPT4%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32%	38%	0	30%
FT4_6%NPSH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32%	38%	0	30%
FT4_9%NPSH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28%	72%

Figure 6.25: Final FDI results obtained for the residual approach

6.6 Summary of results

For ease of interpretation the results obtained for approaches 1 to 3 is summarised in this section. In order to compare the 3 approaches a confusion matrix for each approach was composed as suggested by [6] . In Figure 6.26 the structure and rules for the confusion matrix is shown.

CONFUSION MATRIX

		True condition	
		Fault-free	Fault
Detection/ Isolation condition	Fault-free	True negative ***a*** 1	False negative ***b*** 0
	Fault	False positive ***c*** 0	True positive ***d*** 11

DETECTION AND ISOLATION RATES

Rate	Formula	%
r_{FP}	$= \frac{c}{(a+c)} \times 100$	0
r_{FN}	$= \frac{b}{(b+d)} \times 100$	0
r_{TP}	$= \frac{d}{(b+d)} \times 100$	100
Accuracy	$= \frac{(a+d)}{(a+b+c+d)} \times 100$	100

Figure 6.26: Confusion matrix for true and false detection and isolation [6]

Applying these rules to the results of the three approaches as shown in the respective "result" tables in Table 6.2 and Figures 6.17 and 6.25, the confusion matrices are obtained for detection and isolation as shown in Tables 6.4 and 6.5. To explain the method of applying the rules in Figure 6.26, consider the results obtained for approach 2 as indicated in Figure 6.17. Since the diagonal values represents the correct detection (first row) or isolation (succeeding rows) results the percentages on the diagonal indicates the rate of true positive detections or isolations (r_{TP}). Similarly the sum total of the percentages in each row of the table in Figure 6.17 excluding the diagonal values, indicates the rate of false negative (r_{FN}) detections or isolations. The rate of false positive (r_{FP}) detections or isolations is obtained by the sum of each column excluding the diagonal value. The accuracy is then simply calculated using the equation as given in Figure 6.26. From the results in Tables 6.4 and 6.5 it is clear that approach 1 and 2 are accurate in terms of detection. Approach 3 did not achieve 100% detection although 72% detection are still an acceptable detection rate. In terms of fault isolation, approach 2 did the best overall with consistent results showing a clear pattern when the fault size or location was changes. Approach 3 outperformed approached 2 in terms of isolation of FT FT2_1SPAT. The results are, however, inconsistent when the fault size and location were varied. This might indicate either corrupt data or an underlying infliction point that is found in the energy attributes when analysed according to the procedure of approach 3. Overall approach 1 had the most inaccuracy in terms of fault isolation. With these result in mind in can be concluded that approach 2 would be the overall method of choice for both detection and isolation of faults in STS's.

Table 6.4: Comparison of detection results for the 3 Approaches

	Approach 1	Approach 2	Approach 3
rFP	0%	0%	0%
rFN	0%	0%	28%
rTP	**100%**	**100%**	**72%**
Accuracy	**100%**	**100%**	**72%**

Table 6.5: Comparison of isolation results for the 3 approaches

	Approach 1				Approach 2				Approach 3			
	r_{FP}	r_{FN}	r_{TP}	Accuracy	r_{FP}	r_{FN}	r_{TP}	Accuracy	r_{FP}	r_{FN}	r_{TP}	Accuracy
FT1_ HPT3%	0.00%	0.00%	**100.00%**	**100.00%**	0.00%	0.00%	**100.00%**	**100.00%**	41.00%	41.00%	**59.00%**	**41.84%**
FT1_ MPT3%	91.00%	91.00%	**9.00%**	**4.71%**	29.00%	29.00%	**71.00%**	**55.04%**	90.00%	90.00%	**10.00%**	**5.26%**
FT1_ MPT6%	88.00%	88.00%	**12.00%**	**6.38%**	32.00%	32.00%	**68.00%**	**51.52%**	90.00%	90.00%	**10.00%**	**5.26%**
FT2_ LPT1	85.00%	85.00%	**15.00%**	**8.11%**	0.03%	0.03%	**99.97%**	**99.94%**	90.00%	90.00%	**10.00%**	**5.26%**
FT2_ LPT2	90.00%	90.00%	**10.00%**	**5.26%**	68.00%	68.00%	**32.00%**	**19.05%**	90.00%	90.00%	**10.00%**	**5.26%**
FT2_ LPT3	70.00%	70.00%	**30.00%**	**17.65%**	66.00%	66.00%	**34.00%**	**20.48%**	90.00%	90.00%	**10.00%**	**5.26%**
FT2_0,5SPAT	90.00%	90.00%	**10.00%**	**5.26%**	64.00%	64.00%	**36.00%**	**21.95%**	90.00%	90.00%	**10.00%**	**5.26%**
FT2_ 1SPAT	88.00%	88.00%	**12.00%**	**6.38%**	64.00%	64.00%	**36.00%**	**21.95%**	23.00%	23.00%	**77.00%**	**62.60%**
FT2_ 2SPAT	81.00%	81.00%	**19.00%**	**10.50%**	52.00%	52.00%	**48.00%**	**31.58%**	59.00%	59.00%	**41.00%**	**25.79%**
FT2_ 5SPAT	81.00%	81.00%	**19.00%**	**10.50%**	56.00%	56.00%	**44.00%**	**28.21%**	59.00%	59.00%	**41.00%**	**25.79%**
FT3_ LPT3%	94.00%	94.00%	**6.00%**	**3.09%**	0.00%	0.00%	**100.00%**	**100.00%**	23.00%	23.00%	**77.00%**	**62.60%**
FT3_ LPT4%	66.00%	66.00%	**34.00%**	**20.48%**	0.00%	0.00%	**100.00%**	**100.00%**	17.00%	17.00%	**83.00%**	**70.94%**
FT3_ LPT6%	84.00%	84.00%	**16.00%**	**8.70%**	5.00%	5.00%	**95.00%**	**90.48%**	60.00%	60.00%	**40.00%**	**25.00%**
FT3_ MPt3%	84.00%	84.00%	**16.00%**	**8.70%**	0.00%	0.00%	**100.00%**	**100.00%**	79.00%	79.00%	**21.00%**	**11.73%**
FT3_ MPT4%	83.00%	83.00%	**17.00%**	**9.29%**	1.00%	1.00%	**99.00%**	**98.02%**	68.00%	68.00%	**32.00%**	**19.05%**
FT4_6%NPSH	71.00%	71.00%	**29.00%**	**16.96%**	1.00%	1.00%	**99.00%**	**98.02%**	62.00%	62.00%	**38.00%**	**23.46%**
FT4_9%NPSH	91.00%	91.00%	**9.00%**	**4.71%**	0.00%	0.00%	**100.00%**	**100.00%**	72.00%	72.00%	**28.00%**	**16.28%**

CHAPTER 7

Conclusion

This chapter concludes the study. A discussion on the findings in respect to the objectives and methodology outlined in chapter 1 is given.

The chapter consists of a brief review of the objectives to reflect on whether that they were met. The consequences of the energy characterisation outlined in chapter 4 are discussed with reference to the effects of the obtained graph as seen in chapter 5. The findings in regards to FDI as described in chapter 6 are elaborated on with reference to the differences in the 3 approaches.

The chapter gives recommendations on possible improvements that can be made with reference to some problems found during the application of the EGBV approach for FDI purposes. Possible areas for further research that aligns with the aim of this study is also identified. Lastly the chapter is concluded with final remarks.

7.1 Reflection on the research objectives

In chapter 1, section 1.2 the problem statement was defined as to evaluate the EGBV for FDI applications. Four objectives were identified as being critical in successfully solving of the problem. This section focusses on how these objectives were met.

7.1.1 STS modelling

A static thermodynamic model was developed using energy and mass balance equations. Three main thermodynamic parameters were obtained from the model namely: work, exergy and heat flow. The system was simulated in the software environment EES. In terms of accuracy, the model proved to work within acceptable limits near the 100% operating-point. However, for a larger increase or decrease in the operating-point the model's accuracy was unexceptional with a deviation from the actual results of more than 1%. The reason for this is attributed to two aspects. The first pertains to the assumptions made regarding losses in the system and the second is with regards to the reduction in the system size and complexity. It was thus decided to rather use an existing model developed in the software environment Ebsilon. This model was implemented based on the knowledge acquired from developing the EES model and was used for the simulation of all fault types.

7.1.2 Energy characterisation

The energy characterisation of the system entailed the composition of an attributed graph based on five guidelines. The resulting graph consisted of 20 nodes, of which 19 represented components and 1 represented the environment. A total of 35 links connected the various nodes to form a complete representation of the system. The number of nodes and links can mostly be attributed to questions 2, 3 and 4 in chapter 4, section 4.2. Since large numbers of energy in- and outflows were present in various components while the components also had relatively large energy values, these components were included as nodes. If only question 1 and 5 in chapter 4 are considered, the number of nodes and links might be reduced to only include the components at which faults are being detected and isolated. This affects the holistic availability of system information but might be of value for FDI applications.

7.1.3 Experimental design

Measuring of normal operational temperature, mass flow and pressure at specified locations within the physical system was done with the help of staff from the plant in Jeanschwalde as described in chapter 5. The locations of the measurements were determined based on the attributed graph composed for the system in accordance to the requirements for energy based FDI. The data were utilized to extract normal fluctuations and noise within the system. Through combination of the noise and simulated data from the Ebsilon model, time series were composed for each energy variable obtained from the simulations. The four fault states and variations in the size and locations of the faults were simulated along with a normal operation.

7.1.4 Energy-based fault detection and isolation

A total of 3 different approaches for FDI were evaluated. Each of these approaches was used to detect and isolate faults based on all of the data from the time series. The composition of a cost matrix utilizing the HEOM function pertained to the application of approaches 1 and 2. A distance parameter was calculated as the indicating parameter for fault detection and isolation of approach 1. This distance parameter was obtained as a result of analysing the diagonal of the cost matrix. As a continuation to this, the eigenvalues of the cost matrix were calculated as the indicative values of detecting and isolating a fault for approach 2. For approach 3 a radical change was made by calculating a residual matrix rather than a cost matrix. The residual matrix was analysed both qualitatively and quantitatively. The qualitative analysis was used as the final analysis to indicate the false detection and isolation rate and to evaluate the robustness and sensitivity of the approach.

7.2 Results and findings of the research

Three approaches to energy-based FDI were evaluated. The approaches were evaluated in terms of both fault detection and isolation accuracy. A total of four faults in different locations and of various sizes within the system were detected and isolated. The robustness and sensitivity of each of the 3 approaches were also evaluated. The 3 approaches utilised different mathematical analysis methods of the attributed graph.

Since the 3 approaches had different methods of comparing the node signature matrices, different methods of normalising the data were utilised. In order to be able to compare the approaches, similar methods of composing the final results were followed. Due to the inherent differences, however, there were some variations in the composition of the final results. For example the results of approach 3 show some of the fault measurements being unknown when isolation of a fault was

established. This is due to the threshold implemented during the qualitative analysis.

In terms of the findings of the study it can be concluded that approach 2 had the most accurate results for both detection and isolation of faults. The approach show robustness to changes in the fault types as well as sensitivity to the change in the size of the faults. For FT 2 the sensitivity of approach 2 was inadequate, resulting in ineffective differentiation between the variations of FT 2 with isolation accuracies of 31.5% and lower. In terms of faults detection all three approaches had acceptable results with only approach 3 having a less than perfect detection rate (72%). Both approach 1 and 2, however, did not perform well in terms of isolation. Approach 1 did not achieve isolation of any fault or variation thereof with an accuracy of more than 20.48% with the exception of FT1_HPT3%, which was isolated with 100% accuracy. Approach 3, however, did manage to isolate FT1_HPT3%, FT2_1SPAT, FT3_LPT3% and FT3_LPT4% with an isolation accuracy rate higher than 41.84%. However, the results were inconsistent, showing no clear pattern when the fault size or location were changed. It is theorized that this might be due to some underlying aspect of the energy values that caused this for specific fault size due to the method of analysis used. It was also noted that the increase in information given by the eigenvalues had positive effects whereas the increase in information obtained by the residuals in approach 3 did not add value to the isolation effectiveness of approach 3.

Overall the energy-based analysis shows promise in terms of FDI due to it's reductive capabilities and holistic application regardless of the system type (i.e. mechanical, electrical etc.) The energy-based graph visualisation is also of value in terms of the conservation of structural information of the physical system. The method and techniques, however, need to be formalised for general energy-based applications in order to establish this as a valid FDI approach.

7.3 Recommendations

The accuracy of approach 3 may be increased if more precise analysis of the fault signatures can be followed. The method used in this study utilized a threshold that inevitably neglected to isolate some of the faults resulting in the "unknown" column in Figure 6.25. The implementation of some pattern recognition algorithm to analyse the bar graph signatures might prove to be of value.

The mathematical analytic tools required for FDI while utilizing the attributed graphs can also be simplified if the complexity of the attributed graph can be reduced. Neglecting some of the components in the system will result in a reduction of the number of nodes that might prove useful. This will however be subject to the requirement that the links between the nodes in terms of energy attributes can still be established. One method of accomplishing this might be to focus on faults occurring at a specific location in the system (for example the boiler or turbine). Composition of an attributed graph that encapsulates more information of the component at the specific location, while neglecting most nodes that would represent components in the rest of the system, might reduce the number of nodes and hence reduce the complexity of analysis. To analyse faults in multiple locations of the larger system this method might be applied by composing two or more graphs for the same system. Each of these graphs can then be composed to represent a specific location within the system rather than the whole system. This will reduce the size of each attributed graph and consistently also the analysis of the node signature matrices.

A large reduction of analysis to prove the effectiveness of the EGBV for FDI applications might be accomplished if the analysis approaches used to analyse the attributed graphs can be generalized. In this study each approach was physically implemented using a large amount of data. If it can

be shown that the approach can be applied in a general sense it might also allow for increased performance of the method. This can be thought of as similar to proving that $x + x = 2x$ rather than proving this for all natural numbers.

7.4 Further research

This section provides some suggestions regarding research of future topics that might be of interest to this study. The suggestions are made with reference to the objectives of this study as well as some shortcomings and discrepancies that were observed.

7.4.1 Fault detection and isolation comparison

In terms of establishing the FDI technique it might be of value to thoroughly compare the EGBV method for FDI to some established FDI techniques like principle component analysis or partial least square analysis. This will allow for the establishment of reference cases for which the technique might be more reliable than the conventional approaches. This might also indicate some scenarios where this method might be augmented by conventional approaches. One application of such augmentation that this study already showed, would be to use a machine learning approach to analyse the node signature matrices. This would especially be effective for analysis of approach 3.

7.4.2 Transient analysis of faults

One of the shortcomings of this study is the fact that faults were only detected and isolated from data obtained under non-transient operational conditions. There is therefore no evaluation of faults occurring during a transient response of the system. Since the energy within the system undergoes large changes during transient operations, the fault signatures based on these energy data will also be different under transient conditions than it was for this study.

7.4.3 Composing the attributed graph

During this study 5 guidelines were established for constructing the attributed graph. These guideline were deducted from literature on the EGBV technique, however, there is no standardised method for constructing the attributed graphs. It would be of interest for further research to study the various considerations of energy based graph construction and how different graphs would affect the detection and isolation performance of the energy graph-based approach. As discussed in the previous subsection this would be of value in terms of reducing the mathematical analysis complexity.

7.4.4 Standardise the node signature analysis methods

In terms of accomplishing the actual detection and isolation of the faults, this study focusses on 3 main approaches, each based on specific mathematical analyses. A standardised approach to this will be of value in order to establish the EGBV method as a FDI technique. To accomplish this, however, it will be necessary to standardize the approaches that are followed to analyse the node signature matrices. This requires more research regarding effective matching of matrices and presenting the information encapsulated within the matrices in a concise and efficient way. Methods other that the HEOM function for graph matching may be of value. It might also be interesting to use covariance matrices and statistical analysis to find the correlation between two node signature matrices rather than calculating a cost matrix. To consider alternative matrix

comparison methods will allow for the information in the node signature matrices to be represented with a statistical basis that can be beneficial in terms of effectively presenting the information.

7.5 Final remarks

In this study an energy-based approach to fault detection and fault isolation in a steam turbine system using graph matching is described and utilized. The approach was evaluated in terms of detection and isolation accuracy, robustness and sensitivity. It is concluded that the approach is useful in terms of encapsulating information of multi-domain systems within a simple structure pertaining to the system layout. There is still a need for improvement in terms of constructing the attributed graphs and applying the fault detection and isolation analysis methods. Specific focus on improving upon the accurate fault isolation is suggested. Due to the scope of this study it was not possible to evaluate other analysis methods that could improve fault isolation, however, it can still be concluded that the energy-based method for FDI can be effectively applied to a STS and shows value for industrial applications.

Bibliography

[1] H. C. John, *Cooling Tower Fundamentals*, 2nd ed., John C. Hensley, Ed. Kansas: SPX Cooling Technologies, Inc. Overland Park, Kansas USA, 2003, vol. 2.

[2] C. P. Du Rand and G. Van Schoor, "Fault diagnosis of generation IV nuclear HTGR components - Part II: The area error enthalpy-entropy graph approach," *Annals of Nuclear Energy*, vol. 41, pp. 79–86, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.anucene.2011.11.009

[3] K. R. Uren, G. V. Schoor, and L. Auret, "An energy-attributed graph approach for the purposes of FDI in a heated two-tank system," *IFAC-PapersOnLine*, vol. 1, no. 1, 2019.

[4] C. Karlsson, J. Arriagada, and M. Genrup, "Detection and interactive isolation of faults in steam turbines to support maintenance decisions," *Simulation Modelling Practice and Theory*, vol. 16, no. 10, pp. 1689–1703, 2008. [Online]. Available: http://dx.doi.org/10.1016/j.simpat.2008.08.013

[5] I. J. Karassik, J. P. Messina, and W. H. Fraser, *Pump Handbook*, 3rd ed., McGRAW-HILL, Ed. New York: R. R. Donnelley & Sons Company, 2000.

[6] S. Greyling, H. Marais, G. van Schoor, and K. R. Uren, "Application of exergy-based fault detection in a gas-to-liquids process plant," *Entropy*, vol. 21, no. 6, pp. 1–19, 2019.

[7] R. Beebe, "Condition monitoring of steam turbines by performance analysis," *Journal of Quality in Maintenance Engineering*, vol. 9, no. 2, pp. 102–112, 2003.

[8] H. Wentao, Y. Jun, Z. Xuezeng, and L. Xiaojun, "Fault diagnosis for steam turbine based on flow graphs and naïve Bayesian classifier," *2014 IEEE International Conference on Mechatronics and Automation, IEEE ICMA 2014*, pp. 396–401, 2014.

[9] A. T. Sufian, B. M. Abdullah, M. Ateeq, R. Wah, and D. Clements, "A roadmap towards the smart factory," in *Proceedings - International Conference on Developments in eSystems Engineering*, vol. 12. Kazan, Russia, 7-10 Oct. 2019: IEEE, 2019, pp. 978–983.

[10] M.Mansouri, *Data-Driven and Model-Based Methods for Fault Detection and Diagnosis*.

[11] Y. F. Ding and B. Y. Sheng, "Study on steam turbine fault diagnosis and maintenance service grid system," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 224, no. 3, pp. 517–530, 2010.

[12] R. Isermann, *Combustion Engine Diagnosis*, 1st ed. Heidelberger Platz 3, 14197 Berlin, Germany: Springer Vieweg, Berlin, Heidelberg, 2017. [Online]. Available: https://doi.org/10.1007/978-3-662-49467-7

[13] C. P. Du Rand and G. Van Schoor, "Fault diagnosis of generation IV nuclear HTGR components - Part I: The error enthalpy-entropy graph approach," *Annals of Nuclear Energy*, vol. 40, no. 1, pp. 14–24, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.anucene.2011.09.013

[14] T. Kotas, *The exergy method of thermal plant analysis*, 1st ed., U. o. L. Department of Mechanical Engineering, Queen Mary College, Ed. London: Anchor Brendon Ltd, Tiptree, Essex, 1985, vol. 20, no. 5. [Online]. Available: https://doi.org/10.1016/C2013-0-00894-8

[15] G. V. Schoor and K. R. Uren, "A vision of energy-based visualisation of large scale industrial systems for the purposes of condition monitoring," *31st Conference on Condition Monitoring and Diagnostic Engineering Management*, pp. 337–346, 2018.

[16] I. H. Aljundi, "Energy and exergy analysis of a steam power plant in Jordan," *Applied Thermal Engineering*, vol. 29, no. 2-3, pp. 324–328, 2009. [Online]. Available: http://dx.doi.org/10.1016/j.applthermaleng.2008.02.029

[17] K. R. Uren and G. Van Schoor, "Energy-based visualisation of a counter-flow heat exchanger for the purpose of fault identification," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 19–24, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ifacol.2016.07.210

[18] D. Randall Wilson and T. R. Martinez, "Improved Heterogeneous distance functions," *Journal of Artificial Intelligence Research*, vol. 6, no. 1, pp. 1–34, 1997.

[19] H. Marais, G. van Schoor, and K. R. Uren, "The merits of exergy-based fault detection in petrochemical processes," *Journal of Process Control*, vol. 74, pp. 110–119, 2019. [Online]. Available: https://doi.org/10.1016/j.jprocont.2017.11.005

[20] T. Escobet, A. Bregon, B. Pulido, and V. Puig, *Fault Diagnosis of Dynamic Systems*, S. Teresa Escobet Research Center for Supervision, Safety and Automatic Control (CS2AC) Universitat Politècnica de Catalunya (UPC) Terrassa, Ed. Gewerbestrasse 11, 6330 Cham, Switzerland: Springer Nature Switzerland AG 2019, 2019. [Online]. Available: https://doi.org/10.1007/978-3-030-17728-7

[21] R. Koningsveld, *Thermodynamics and Engineering Needs.*, 8th ed. New York: McGraw-Hill Education, 2 Penn Plaza, New York, 1986.

[22] L. Wald, *CONTROL SYSTEMS, ROBOTICS AND AUTOMATION*, 12th ed., H. Unbehauen, Ed. Shanghai: EOLSS Publications, 2015, no. 12.

[23] T. Tanuma, *Advances in Steam Turbines for Modern Power Plants*, L. Lawrence, Ed. Duxford UK: Joe Hayton Woodhead Publishing Elsrevier, 2016.

[24] Alexander S. Leyzerovich, *Steam Turbines for Modern Fossil Fule Power Plants*, 1st ed., Taylor & Francis Ltd., Ed., 2008.

[25] P. J. SHEPHERD, *Fundamentals of Thermodynamics*, 8th ed., U. of Michigan, Ed. Michigan: John Wiley & Sons, 2013.

[26] F. Alobaid, *Numerical Simulation for Next Generation Thermal Power Plants*, 1st ed. Switzerland: Springer Tracts in Mechanical Engineering, 2018. [Online]. Available: http://link.springer.com/10.1007/978-3-319-76234-0https://doi.org/10.1007/978-3-319-76234-0Library

[27] R. J. Wilson, *Introduction to graph theory*, 4th ed. Edinburgh Gate, Harlow, Essex CM20 2JE, England: Addison Wesley Longman Limited, 1996. [Online]. Available: https://www.bibsonomy.org/bibtex/23db0cc078e64a35a307dac0534c562cd/ytyoun

[28] C. Griffin, "Graph Theory: Penn State Math 485," *Lecture Notes*, 2017. [Online]. Available: http://www.personal.psu.edu/cxg286/Math485.pdf

[29] A. Schenker, "Graph-theoretic techniques for web content mining," Ph.D. dissertation, University of South Florida, 2003. [Online]. Available: http://scholarcommons.usf.edu/etd/1467

[30] M. S. Spencer, S. C. Bates Prins, and M. S. Beckom, "Heterogeneous distance measures and nearest-neighbor classification in an ecological setting," *Missouri Journal of Mathematical Sciences*, vol. 22, no. 2, pp. 108–123, 2010.

[31] G. Van Schoor, K. R. Uren, M. A. Van Wyk, P. A. Van Vuuren, and C. P. Du Rand, "An energy perspective on modelling, supervision, and control of large-scale industrial systems: Survey and framework," in *19th IFAC World Congress, Cape Town, South Africa,*, vol. 19, no. 3. IFAC, 2014, pp. 6692–6703. [Online]. Available: http://dx.doi.org/10.3182/20140824-6-ZA-1003.02190

[32] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, "A review of process fault detection and diagnosis," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 327–346, 2003.

[33] L. Liao and F. Köttig, "A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction," *Applied Soft Computing Journal*, vol. 44, pp. 191–199, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.asoc.2016.03.013

[34] F. Nadir, H. Elias, and B. Messaoud, "Diagnosis of defects by principal component analysis of a gas turbine," *SN Applied Sciences*, vol. 2:980, no. April, 2020. [Online]. Available: https://doi.org/10.1007/s42452-020-2796-y

[35] G. Tsatsaronis, "Comments on the paper 'a brief commented history of exergy from the beginnings to 2004'," *International Journal of Thermodynamics*, vol. 10, no. 4, pp. 187–190, 2007.

[36] T. A. Reddy, "Formulation of a generic methodology for assessing FDD methods and its specific adoption to large chillers," *ASHRAE Transactions*, vol. 113 PART 2, pp. 334–342, 2007.

[37] S. Frank, G. Lin, X. Jin, R. Singla, A. Farthing, L. Zhang, and J. Granderson, "Metrics and Methods to Assess Building Fault Detection and Diagnosis Tools." National Renewable Energy Laboratory, Michigan, Tech. Rep. March, 2019. [Online]. Available: https://www.nrel.gov/docs/fy19osti/72801.pdf.{%}0ANREL

[38] H. J. Marais, G. Van Schoor, and K. R. Uren, "An energy-based approach to condition monitoring of industrial processes," in *Part of special issue: 9th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS*, Elsevier, Ed., vol. 28, no. 21. Paris, 2–4 September 2015: Elsevier B.V., pp. 772–777. [Online]. Available: http://dx.doi.org/10.1016/j.ifacol.2015.09.620

[39] H. J. Marais, G. van Schoor, and K. R. Uren, "Energy-based Fault Detection for an Autothermal Reformer," in *11th IFAC Symposium on Dynamics and Control of Process SystemsIncluding Biosystems DYCOPS-CAB*, Elsevier, Ed., vol. 49, no. 7. Trondheim, Norway, 6—8 June 2016: Elsevier B.V., 2016, pp. 353–358. [Online]. Available: http://dx.doi.org/10.1016/j.ifacol.2016.07.325

[40] M. Elhelw, "Utilizing exergy analysis in studying the performance of steam power plant at two different operation mode," *Applied Thermal Engineering*, vol. 150, no. March 2018, pp. 285–293, 2019. [Online]. Available: https://doi.org/10.1016/j.applthermaleng.2019.01.003

[41] B. Gallagher, "Matching structure and semantics: A survey on graph-based pattern matching," *AAAI Fall Symposium - Technical Report*, vol. FS-06-02, pp. 45–53, 2006. [Online]. Available: https://www.aaai.org/Papers/Symposia/Fall/2006/FS-06-02/FS06-02-007.pdf

[42] T. Washio and H. Motoda, "State of the art of graph-based data mining," *ACM SIGKDD Explorations Newsletter*, vol. 5, no. 1, p. 59, 2003.

[43] SHINJI UMEYAMA, "An Eigendecomposition Approach to Weighted Graph Matching Problems," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE.*, vol. 10, no. 5, 1988.

[44] D. C. Karnopp, D. L. Margolis, and R. C. Rosenberg, *Design of Mechatronic Systems*, 5th ed. New Jersey: JOHN WILEY & SONS, INC, 2006.

[45] P. Rousseau and F. O. F. Engineering, "Thermal-fluid systems modelling 1," 2013. [Online]. Available: https://figshare.com/s/078faef1fec6749d2b3a

[46] P.G. Rousseau, "Thermal-fluid systems modelling ii," 2014. [Online]. Available: https://figshare.com/s/078faef1fec6749d2b3a

[47] H. P. Bloch and M. P. Singh, *Steam Turbines : Design, Applications, and Rerating*. New York: McGraw-Hill Companies, 2009, vol. 2, no. 9.

[48] M. Topel, *Steam Turbine Thermal Modeling for Improved Transient Operation*, 2014.

[49] Tadashi Tanuma, *Advances in steam turbines for modern power plants*, 1st ed., L. Lawrence, Ed. 50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States: Whoodhead Publishing, vol. 1, no. 1.

[50] A. A. Sonin, *Fundamental Laws of Motion for Particles , Material Volumes , and Control Volumes*, 1st ed. Massachusetts: Massachusetts Institute of Technology, 2001.

[51] J. J. de Bruin, K. R. Uren, G. van Schoor, and M. van Eldik, "An energy graph-based approach to fault diagnosis of a transcritical CO2 heat pump," *Energies*, vol. 13, no. 7, pp. 1–35, 2020.

[52] T. Gresh, *Compressor Performance: Aerodynamics for the User*, 2nd ed. Elsevier Science & Technology Books, 2001, no. 2.

[53] O. Khayal, "FUNDAMENTALS OF HEAT EXCHANGERS Applications and Robotics Issn 2320-7345," *INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS*, vol. 1, no. 12, pp. 1–11, 2018.

[54] R. K. Shah and D. P. Sekuli, *Selection of Heat Exchangers and Their Components*, 1st ed. Rochester, New York: JOHN WILEY & SONS, INC, 2007.

[55] T. Koroglu and O. S. Sogut, "Conventional and advanced exergy analyses of a marine steam power plant," *Energy*, vol. 163, pp. 392–403, 2018. [Online]. Available: https://doi.org/10.1016/j.energy.2018.08.119

[56] G. R. Ahmadi and D. Toghraie, "Energy and exergy analysis of Montazeri Steam Power Plant in Iran," in *Renewable and Sustainable Energy Reviews*, vol. 56. Elsevier, 2016, pp. 454–463.

[57] S. R. Moosavi, J. Qajar, and M. Riazi, "A comparison of methods for denoising of well test pressure data," *Journal of Petroleum Exploration and Production Technology*, vol. 8, no. 4, pp. 1519–1534, 2018. [Online]. Available: https://doi.org/10.1007/s13202-017-0427-y

[58] S. Athichanagorn, "DEVELOPMENT OF AN INTERPRETATION METHODOLOGY FOR LONG-TERM PRESSURE DATA FROM PERMANENT DOWNHOLE GAUGES c Copyright 1999 by Suwat Athichanagorn," Doctor of philosophy, stanford university, 1999.

[59] A. Raudys and Z. Pabarskaite, "Optimising the smoothness and accuracy of moving average for stock price data," *Technological and Economic Development of Economy*, vol. 24, no. 3, pp. 984–1003, 2018.

[60] N. B. Gallagher, "Savitzky–Golay filter for smoothing and differentiation," *Signal processing*, no. January, pp. 2–6, 2013. [Online]. Available: https://doi.org/10.1016/j.sigpro.2005.02.002

[61] H. Huang, S. Hu, and Y. Sun, "A discrete curvature estimation based low-distortion adaptive savitzky–golay filter for ECG denoising," *Sensors (Switzerland)*, vol. 19, no. 7, pp. 1–18, 2019.

[62] R. W. Katz, "Autoregressive moving average processes to model meteorological time series," *American Meteorological Society*, vol. 109, no. 3, pp. 479–484, 1980. [Online]. Available: https://doi.org/10.1175/1520-0493(1981)109{%}3C0479:OTUOAM{%}3E2.0.CO;2

[63] Y. Zhang, R. Wang, S. Li, and S. Qi, "Temperature sensor denoising algorithm based on curve fitting and compound kalman filtering," *Sensors (Switzerland)*, vol. 20, no. 7, pp. 1–13, 2020.

[64] T. Cour, P. Srinivasan, and J. Shi, "Balanced graph matching," *Advances in Neural Information Processing Systems*, pp. 313–320, 2007.

[65] W. v. N. BW. Botha, M .v.Eldik, "INTRODUCTION TO THERMAL-FLUID SCIENCES NUCI 521 PEC," vol. 1, 2012. [Online]. Available: https://figshare.com/s/078faef1fec6749d2b3a

[66] Fchart, "Engineering Equation Solver Manual," pp. 608–836, 2015. [Online]. Available: http://www.fchart.com/ees

[67] "EES: Engineering Equation Solver | F-Chart Software : Engineering Software." [Online]. Available: http://fchartsoftware.com/ees/

[68] "E E S Engineering Equation Solver for Microsoft Windows Operating Systems Commercial and Professional Versions F-Chart Software," Tech. Rep., 1992. [Online]. Available: http://www.fchart.com/

[69] B. El Hefni, D. Bouskela, B. El Hefni, and D. Bouskela, *Modeling and Simulation of Thermal Power Plants*, 1st ed. Springer Nature Switzerland, 2019. [Online]. Available: https://doi.org/10.1007/978-3-030-05105-1

[70] K. Rayaprolu, *Boilers forpower and process*, 1st ed. CRC Press Taylor and Francis Group, 2009, no. 1.

[71] D. G. Kroger, *Air-cooled Heat exchangers and Cooling Towers: Thermal-Flow Performance Evaluation and Design*, 2nd ed. Tulsa,: PennWell Corporation, 2013, vol. 2, no. 9.

[72] Europa, "Janschwalde Plant spesification document." [Online]. Available: https://figshare.com/s/078faef1fec6749d2b3a

[73] C. PE, "Diagnostic monitoring of solid particle erosion in steam turbines," *IEEE Transactions on energy conversion*, vol. 3, no. 2, 1988.

[74] W. Sun and Y. Wang, "MALFUNCTIONS OF A STEAM TURBINE MECHANICAL CONTROL SYSTEM," *IOP Conference Series: Earth and Environmental Science*, vol. 354, no. 1, 2019. [Online]. Available: https://iopscience.iop.org/article/10.1088/1755-1315/354/1/012066

[75] D. C. RENNELS, G. E. C. (ret.), and HOBART M. HUDSON Aerojet General Corporation (ret.), *PIPE FLOW A Practical and Comprehensive Guide*, 1st ed. John Wiley & Sons, Inc., Hoboken, New Jersey. [Online]. Available: www.wiley.com

[76] S. Greyling, G. V. Schoor, K. Richard, and U. Henri, "Exergy graph-based fault detection and isolation of a gas-to-liquids process," 2019. [Online]. Available: https://figshare.com/s/078faef1fec6749d2b3a

[77] S. van Graan, G. van Schoor, and K. R. Uren, "Graph matching as a means to energy-visualisation of a counter-flow heat exchanger for the purpose of fault diagnosis," *20th IFAC world congress, Toulouse, France, 2017*, vol. 50, no. 1, pp. 2842–2847, 2017. [Online]. Available: https://doi.org/10.1016/j.ifacol.2017.08.637

[78] H. Neser, "Energy-based visualisation of a Brayton cycle power conversion unit for the purpose of condition monitoring," Doctor of Philosophy, North-West University, 2019. [Online]. Available: http://hdl.handle.net/10394/34078

[79] D. R. Wilson and T. R. Martinez, "Improved heterogeneous distance functions," *Journal of Artificial Intelligence Research*, vol. 6, pp. 1–34, 1997.

[80] L. W. Johnsonn, R. D. Riess, and J. T. Arnold, *Introduction to Linear Algebra*, 5th ed., Virginia Polytechnic Institute and State University, Ed. Boston: Addison Wesley Longman Limited, 2017. [Online]. Available: https://doi.org/10.4324/9780203788219

Appendices

APPENDIX A

System diagram

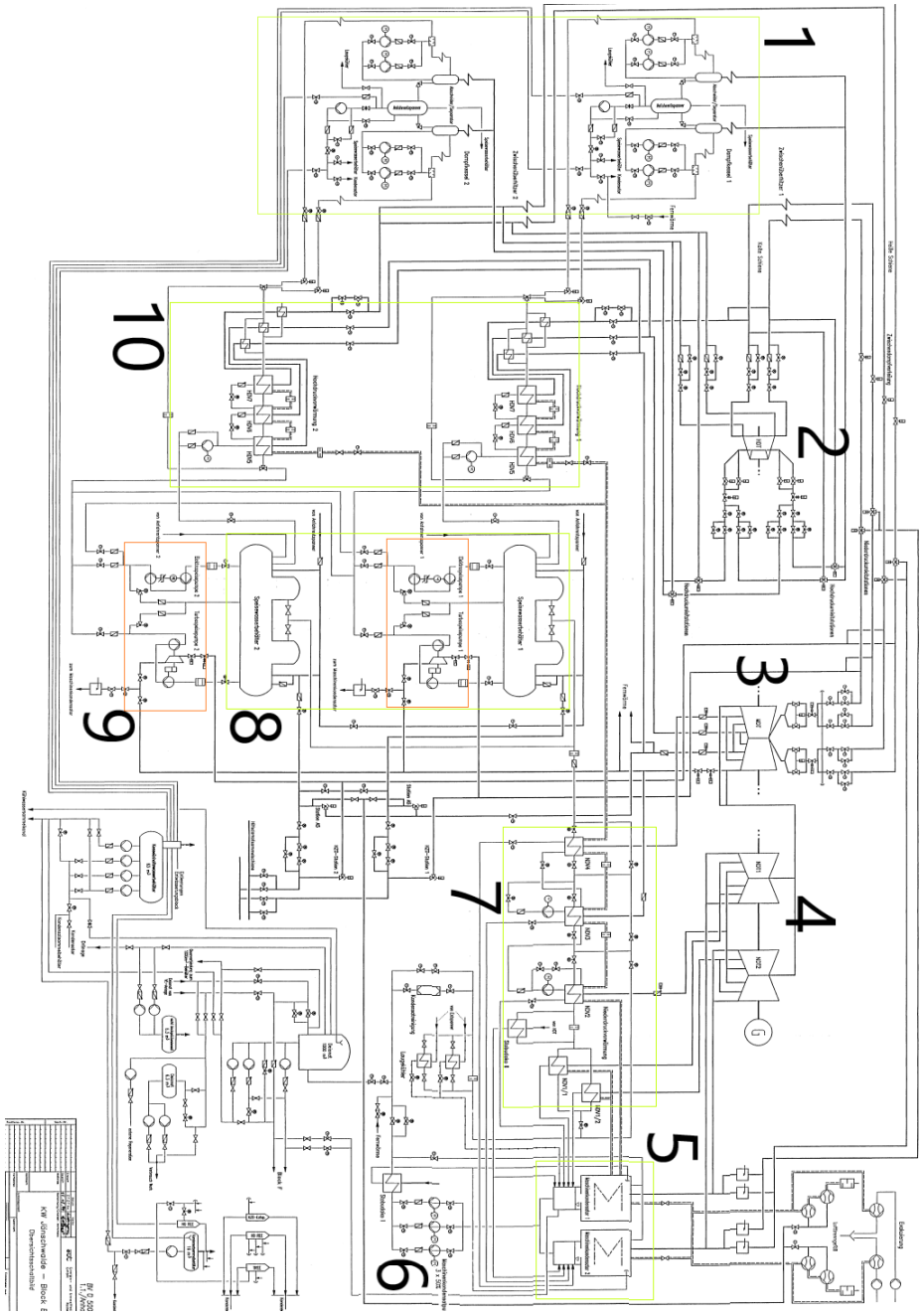


Figure A.1: Schematics of one unit/block of the physical system.

APPENDIX B

Simulations and practical data

B.1 EES Model validation

Table B.1: Result from comparing the EES and Ebsilon model at a 103% generator load requirement.

Generator load at 103%						
	Ebsilon		EES		Difference	
ID	**kPa**	**C**	**kPa**	**C**	**kPa**	**C**
1si	18777.6	258.79	19508	248	730.4	-10.79
1se	17280.9	535	16800	535	-480.9	0
2i	15805.9	528.878	15150	526.7	-655.9	-2.178
2e	4372	347.432	4116	333.7	-256	-13.732
1sri	4372	347.432	4066	333.2	-306	-14.232
1sre	3947.1	540	3853	540	-94.1	0
3.1i	3875.2	536.607	3733	685.1	-142.2	148.493
3.2i	2135.5	446.354	2027	457.9	-108.5	11.546
3.3i	1412.2	289.221	1261	395.2	-151.2	105.979
3.4i	338.2	219.668	524.4	286.6	186.2	66.932
3.4e	338.2	219.668	335.8	230.1	-2.4	10.432
4.1f	335.8	230.1	335.8	230.1	0	0
SPe	335.8	230.1	335.8	230.1	0	0
4.1i	329.3	233.339	335.8	230.1	6.5	-3.239
4.2i	106	131.211	78.4	108	-27.6	-23.211
4.3i	19.7	59.755	13.6	51.97	-6.1	-7.785
4.4e	5	33.074	4.2	29.81	-0.8	-3.264
5pe	5	33.074	4.2	29.81	-0.8	-3.264

B.2 Ebsilon simulation of fault types

This section shows the changes made tot the Ebsilon model in order to simulate the various faults. Each figure shows a specific change that either changed the faults size or location within the system.

B.2.1 Fault type one

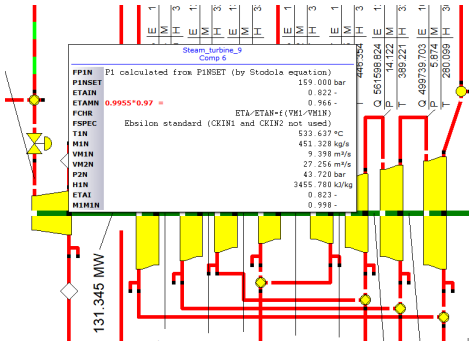


Figure B.1: Set-up of Ebsilon to simulate the effects of solid particle erosion simulated on the HPT with 3% drop in efficiency and 3% increase in mass flow rate.

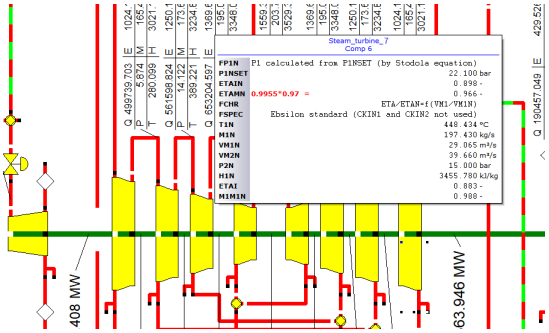


Figure B.2: Set-up of Ebsilon to simulate the effects of solid particle erosion simulated on stage 1 of the MPT with 3% drop in efficiency and 3% increase in mass flow rate.

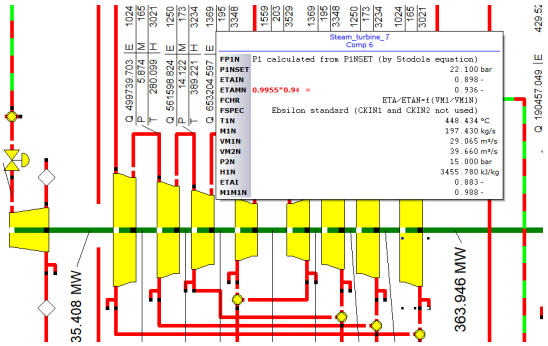


Figure B.3: Set-up of Ebsilon to simulate the effects of solid particle erosion simulated on stage 1 of the MPT with 6% drop in efficiency and 6% increase in mass flow rate.

B.2.2 Fault type two

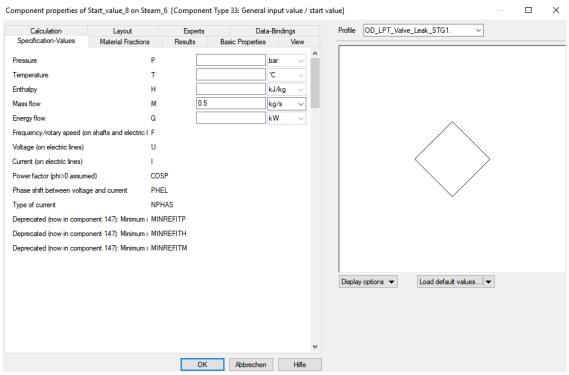


Figure B.4: Set-up of Ebsilon to simulate the effects of a 0,5kg/s leak in the overflow valve of stage 1 of the LPT.

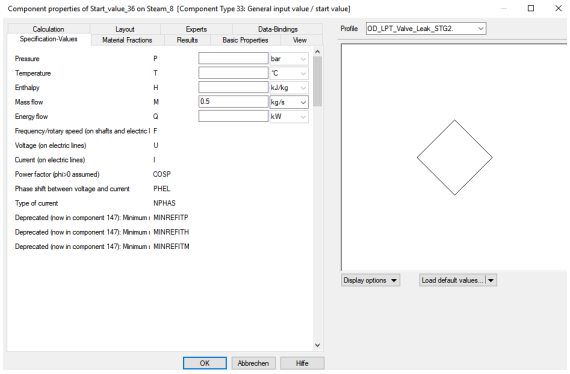


Figure B.5: Set-up of Ebsilon to simulate the effects of a 0,5kg/s leak in the overflow valve of stage 2 of the LPT.

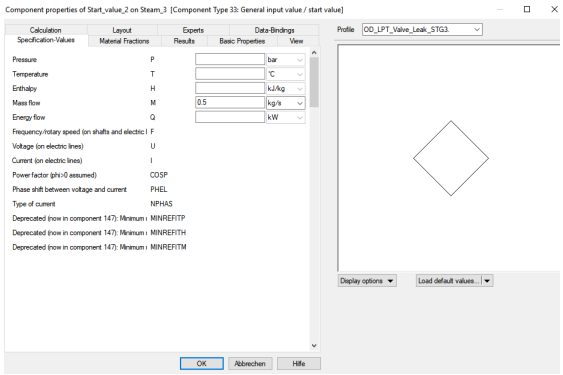


Figure B.6: Set-up of Ebsilon to simulate the effects of a 0,5kg/s leak in the overflow valve of stage 3 of the LPT.

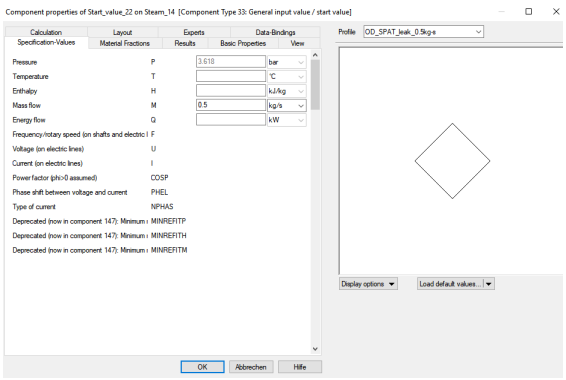


Figure B.7: Set-up of Ebsilon to simulate the effects of a 0,5kg/s leak in the overflow valve of the SPAT.

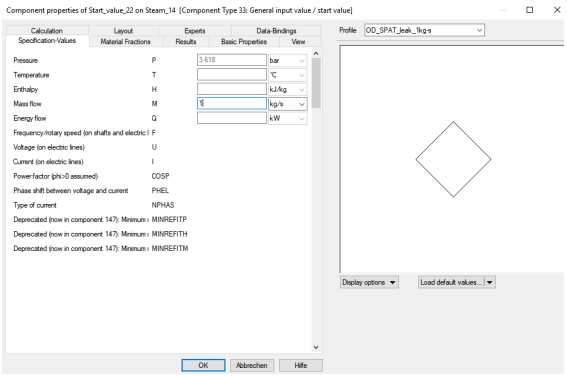


Figure B.8: Set-up of Ebsilon to simulate the effects of a 1kg/s leak in the overflow valve of the SPAT.

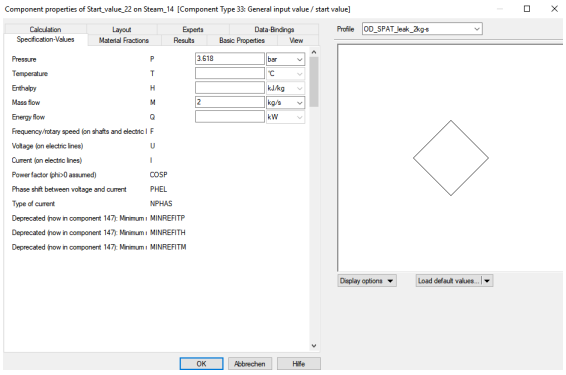


Figure B.9: Set-up of Ebsilon to simulate the effects of a 2kg/s leak in the overflow valve of the SPAT.

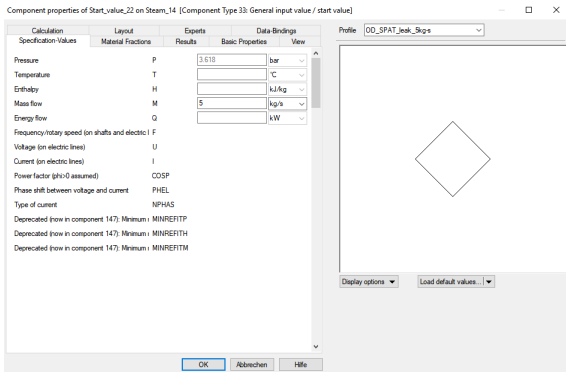


Figure B.10: Set-up of Ebsilon to simulate the effects of a 5kg/s leak in the overflow valve of the SPAT.

B.2.3 Fault type three

Stage One of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET	3.500	bar
ETAIN	0.849	-
ETAMN	0.9955*0.97 = 0.966	-
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N	235.579	°C
M1N	90.368	kg/s
VM1N	59.794	m³/s
VM2N	148.221	m³/s
P2N	1.130	bar
H1N	3455.780	kJ/kg
ETAI	0.848	-
M1M1N	0.943	-

Stage Two of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET	1.130	bar
ETAIN	0.849	-
ETAMN	0.9955*0.97 = 0.966	-
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N	133.175	°C
M1N	84.310	kg/s
VM1N	138.285	m³/s
VM2N	590.209	m³/s
P2N	0.210	bar
H1N	3455.780	kJ/kg
ETAI	0.848	-
M1M1N	0.940	-

Stage Three of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET	0.210	bar
ETAIN	0.849	-
ETAMN	0.9955*0.97 = 0.966	-
FCHR	ETA/ETAN=f(M1/M1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N	61.145	°C
M1N	80.825	kg/s
VM1N	565.812	m³/s
VM2N	1969.372	m³/s
P2N	0.053	bar
H1N	3455.780	kJ/kg
ETAI	0.848	-
M1M1N	0.939	-

Figure B.11: Set-up of Ebsilon to simulate the effects of overall wear and ageing causing a 3% drop in mechanical performance of the LPT.

Stage One of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		3.500 bar
ETAIN		0.849 -
ETAMN	0.9955*0.96 =	0.966 -
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		235.579 °C
M1N		90.368 kg/s
VM1N		59.794 m³/s
VM2N		148.221 m³/s
P2N		1.130 bar
H1N		3455.780 kJ/kg
ETAI		0.848 -
M1M1N		0.943 -

Stage Two of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		1.130 bar
ETAIN		0.849 -
ETAMN	0.9955*0.96 =	0.966 -
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		133.175 °C
M1N		84.310 kg/s
VM1N		138.285 m³/s
VM2N		590.209 m³/s
P2N		0.210 bar
H1N		3455.780 kJ/kg
ETAI		0.848 -
M1M1N		0.940 -

Stage Three of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		0.210 bar
ETAIN		0.849 -
ETAMN	0.9955*0.96 =	0.966 -
FCHR	ETA/ETAN=f(M1/M1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		61.145 °C
M1N		80.825 kg/s
VM1N		565.812 m³/s
VM2N		1969.372 m³/s
P2N		0.053 bar
H1N		3455.780 kJ/kg
ETAI		0.848 -
M1M1N		0.939 -

Figure B.12: Set-up of Ebsilon to simulate the effects of overall wear and ageing causing a 4% drop in mechanical performance of the LPT.

Stage One of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		3.500 bar
ETAIN		0.849 -
ETAMN	0.9955*0.94 =	0.966 -
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		235.579 °C
M1N		90.368 kg/s
VM1N		59.794 m³/s
VM2N		148.221 m³/s
P2N		1.130 bar
H1N		3455.780 kJ/kg
ETAI		0.848 -
M1M1N		0.943 -

Stage Two of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		1.130 bar
ETAIN		0.849 -
ETAMN	0.9955*0.94 =	0.966 -
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		133.175 °C
M1N		84.310 kg/s
VM1N		138.285 m³/s
VM2N		590.209 m³/s
P2N		0.210 bar
H1N		3455.780 kJ/kg
ETAI		0.848 -
M1M1N		0.940 -

Stage Three of the LPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		0.210 bar
ETAIN		0.849 -
ETAMN	0.9955*0.94 =	0.966 -
FCHR	ETA/ETAN=f(M1/M1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		61.145 °C
M1N		80.825 kg/s
VM1N		565.812 m³/s
VM2N		1969.372 m³/s
P2N		0.053 bar
H1N		3455.780 kJ/kg
ETAI		0.848 -
M1M1N		0.939 -

Figure B.13: Set-up of Ebsilon to simulate the effects of overall wear and ageing causing a 6% drop in mechanical performance of the LPT.

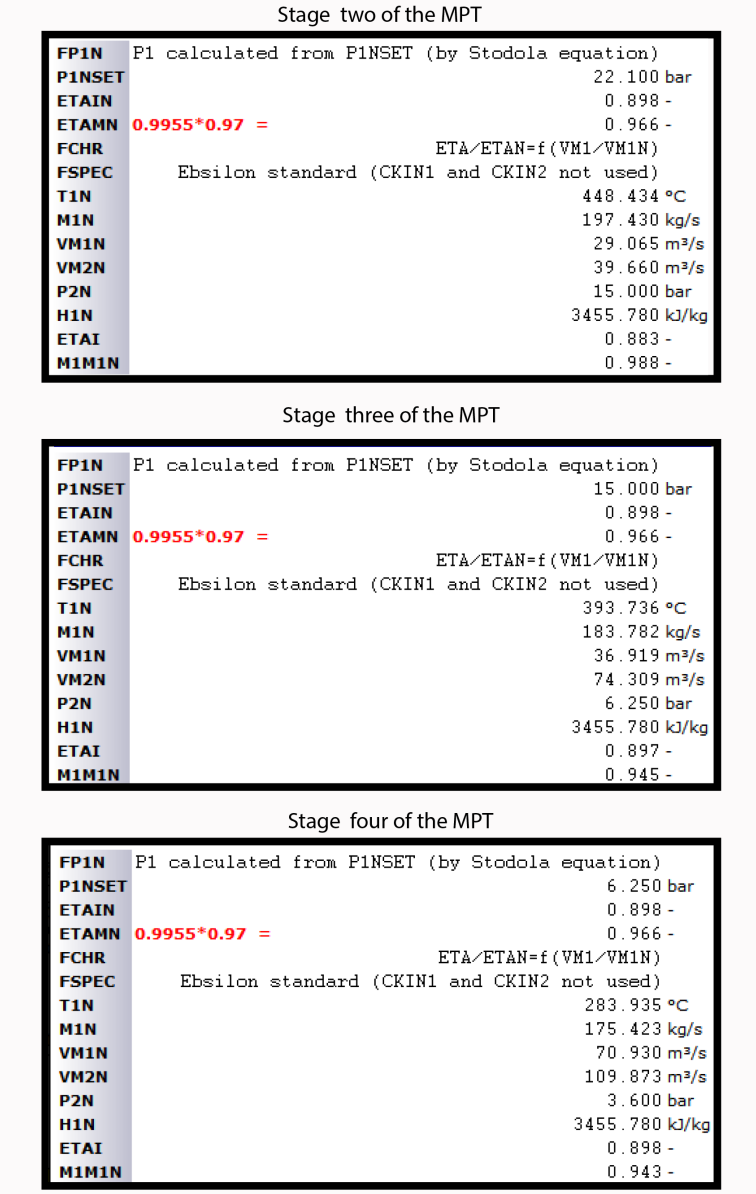


Figure B.14: Set-up of Ebsilon to simulate the effects of overall wear and ageing causing a 3% drop in mechanical performance of the MPT.

Stage two of the MPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		22.100 bar
ETAIN		0.898 -
ETAMN	0.9955*0.96 =	0.966 -
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		448.434 °C
M1N		197.430 kg/s
VM1N		29.065 m³/s
VM2N		39.660 m³/s
P2N		15.000 bar
H1N		3455.780 kJ/kg
ETAI		0.883 -
M1M1N		0.988 -

Stage three of the MPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		15.000 bar
ETAIN		0.898 -
ETAMN	0.9955*0.96 =	0.966 -
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		393.736 °C
M1N		183.782 kg/s
VM1N		36.919 m³/s
VM2N		74.309 m³/s
P2N		6.250 bar
H1N		3455.780 kJ/kg
ETAI		0.897 -
M1M1N		0.945 -

Stage four of the MPT

FP1N	P1 calculated from P1NSET (by Stodola equation)	
P1NSET		6.250 bar
ETAIN		0.898 -
ETAMN	0.9955*0.96 =	0.966 -
FCHR	ETA/ETAN=f(VM1/VM1N)	
FSPEC	Ebsilon standard (CKIN1 and CKIN2 not used)	
T1N		283.935 °C
M1N		175.423 kg/s
VM1N		70.930 m³/s
VM2N		109.873 m³/s
P2N		3.600 bar
H1N		3455.780 kJ/kg
ETAI		0.898 -
M1M1N		0.943 -

Figure B.15: Set-up of Ebsilon to simulate the effects of overall wear and ageing causing a 4% drop in mechanical performance of the MPT.

B.2.4 Fault type four

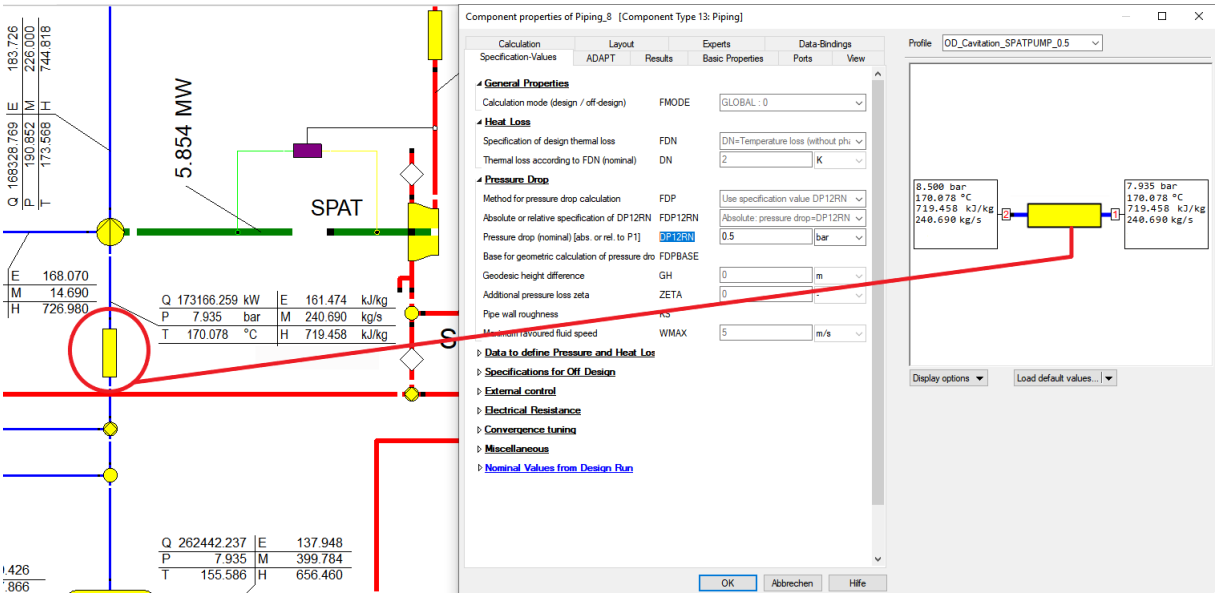


Figure B.16: Set-up of Ebsilon to simulate a state during which cavitation will happen with the fluid pressure +- 6% below the vapour pressure.

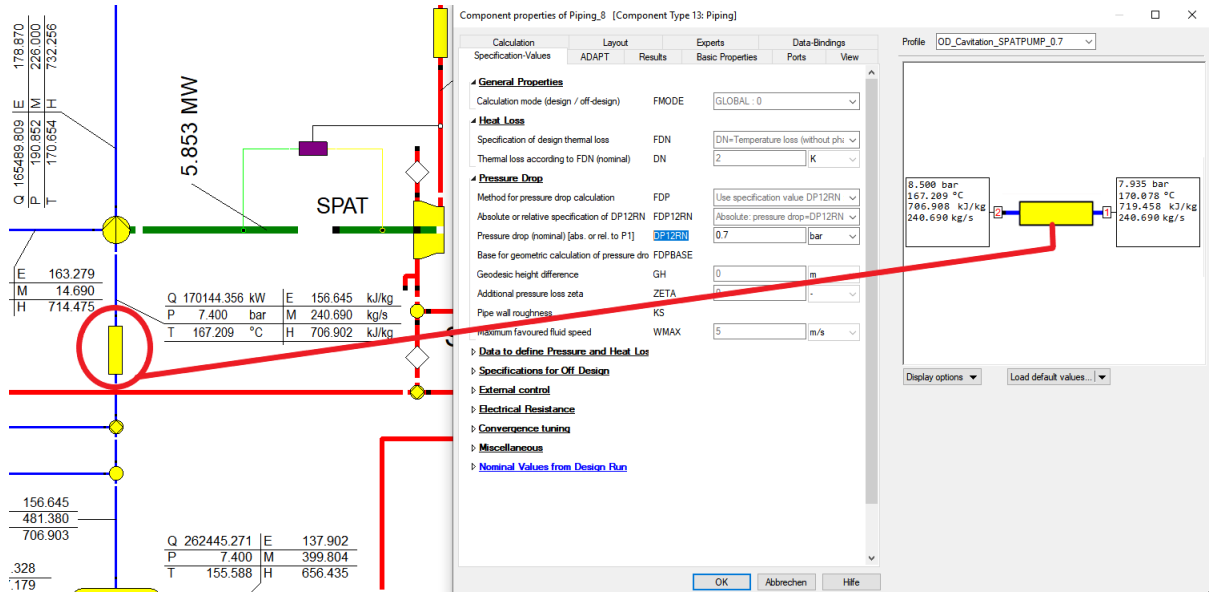


Figure B.17: Set-up of Ebsilon to simulate a state during which cavitation will happen with the fluid pressure +- 9% below the vapour pressure.

B.3 Physical system data

The data measured in the physical system can be obtained by following this link:
Physical system data

B.4 Ebsilon Data

The simulation results obtained for each fault with the Ebsilon model can be obtained by following this link:
Ebsilon fault simulations

B.5 Timeseries data

The time series data that was generated can be obtained by following this link:
Timeseries data

B.6 Algorithms for FDI approaches

B.6.1 Approach 1

The variable $\mathbf{C}_{i,j}$ is the cost matrices calculated for each FT compared to each reference graph. The distance parameter as discussed in (6.1) is represented by $\mathbf{DC}_{i,j}$ for each FT compared to each reference graph.

Algorithm 1: Distance parameter approach. Calculate for $i \in [1,2,...,n]$ and $j \in [1,2,...2880]$ with n the number of variations on the fault states and the normal state (i.e 18).

Input: Operational and reference node signature matrices $\mathbf{G}_{op},(i,j)$ and $\mathbf{G}_{ref}(i,j)$.

$\mathbf{C}_{i,j} \leftarrow (\mathbf{G}_{op},(i,j),\mathbf{G}_{ref}(i,j))$
$\mathbf{DC}_{i,j} \leftarrow (\mathbf{C}_{i,j})$

B.6.2 Approach 2

The variable $\mathbf{Z}_{i,j}$ is defined in (6.2). The final output $\mathbf{Q}_{i,j}$ is the qualitative fault signature.

Algorithm 2: Eigenvalue approach. Calculate for $i \in [1,2,...,n]$ and $j \in [1,2,...2880]$ with n the number of variations on the fault states and the normal state (i.e 18). Final output $\mathbf{Q}_{i,j}$ is the qualitative fault signature.

Input: Operational and reference node signature matrices $\mathbf{G}_{op},(i,j)$ and $\mathbf{G}_{ref}(i,j)$.

$\mathbf{C}_{i,j} \leftarrow (\mathbf{G}_{op},(i,j),\mathbf{G}_{ref}(i,j))$
$\lambda_{i,j} \leftarrow (\mathbf{C}_{i,j})$
$\mathbf{Z}_{i,j} \leftarrow (\lambda_{i,j})$
$\mathbf{Q}_{i,j} \leftarrow (\mathbf{Z}_{i,j})$

B.6.3 Approach 3

The variable $\mathbf{G_{res}}$ is defined in (6.3). The variable $\mathbf{GN_{res}}$ is the normalised residuals as defined in (6.5). The final output is the frequency vector as discussed in Figure 6.24.

Algorithm 3: Residual approach. Calculate for $i \in [1,2,...,n]$ and $j \in [1,2,...2880]$ with n the number of variations on the fault states and the normal state (i.e 18).

Input: Operational and max reference node signature matrices $\mathbf{G}_{op},(i,j)$ and $\mathbf{G}_{max}(i,j)$.

$\mathbf{G_{res}}(i,j) \leftarrow (\mathbf{G}_{op},(i,j),\mathbf{G}_{max}(i,j))$
$\mathbf{GN_{res}}(i,j) \leftarrow (\mathbf{G_{res}}_{i,j})$
$\mathbf{Vf}_{i,j} \leftarrow (\mathbf{GN_{res}}(i,j))$

APPENDIX C

Results

C.1 Fault detection and isolation results

C.1.1 Approach 1: Distance parameter

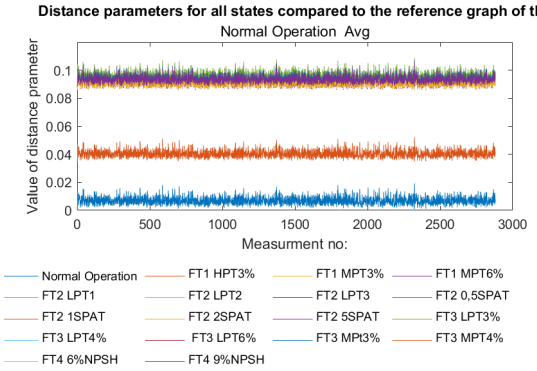


Figure C.1: Detection results in terms of the distance parameter plots for all the fault types.

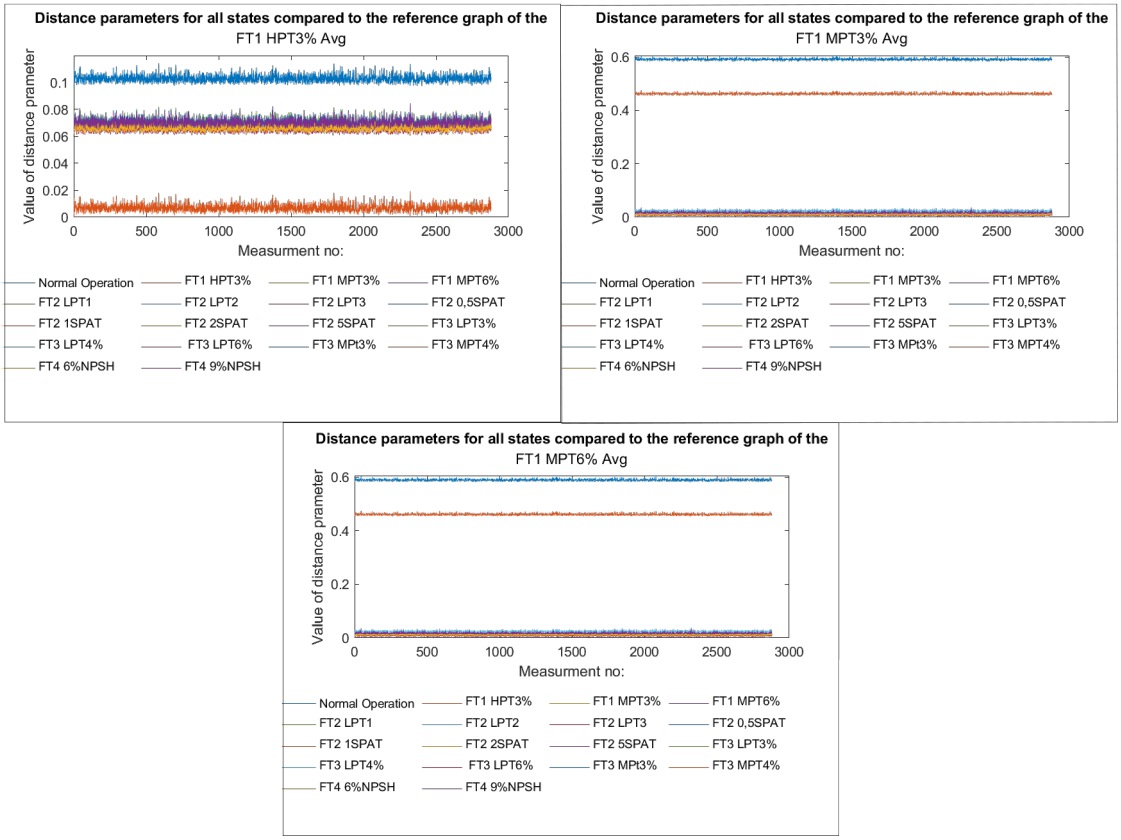


Figure C.2: Isolation results in terms of the distance parameter plots for FT 1.

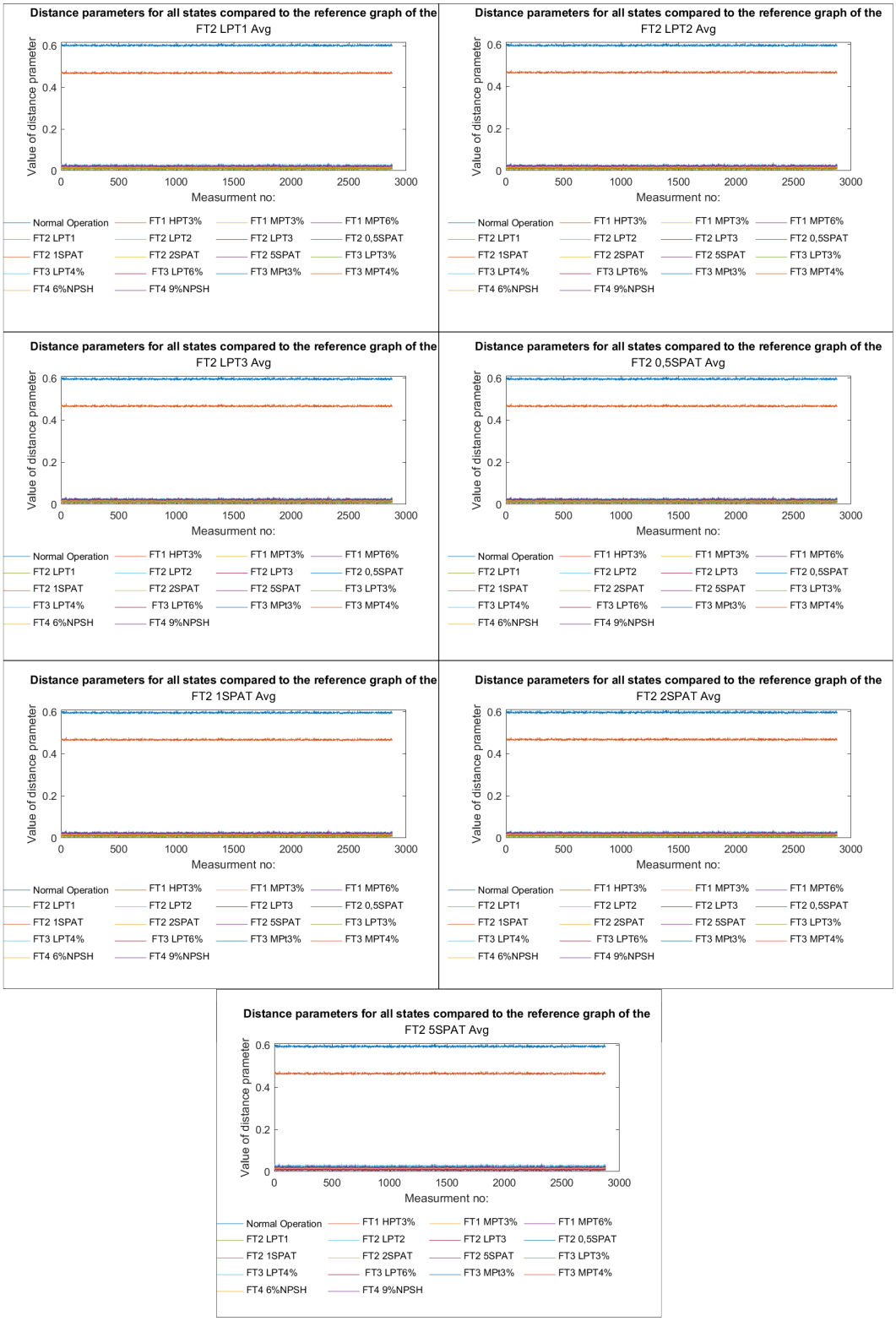


Figure C.3: Isolation results in terms of the distance parameter plots for FT 2.

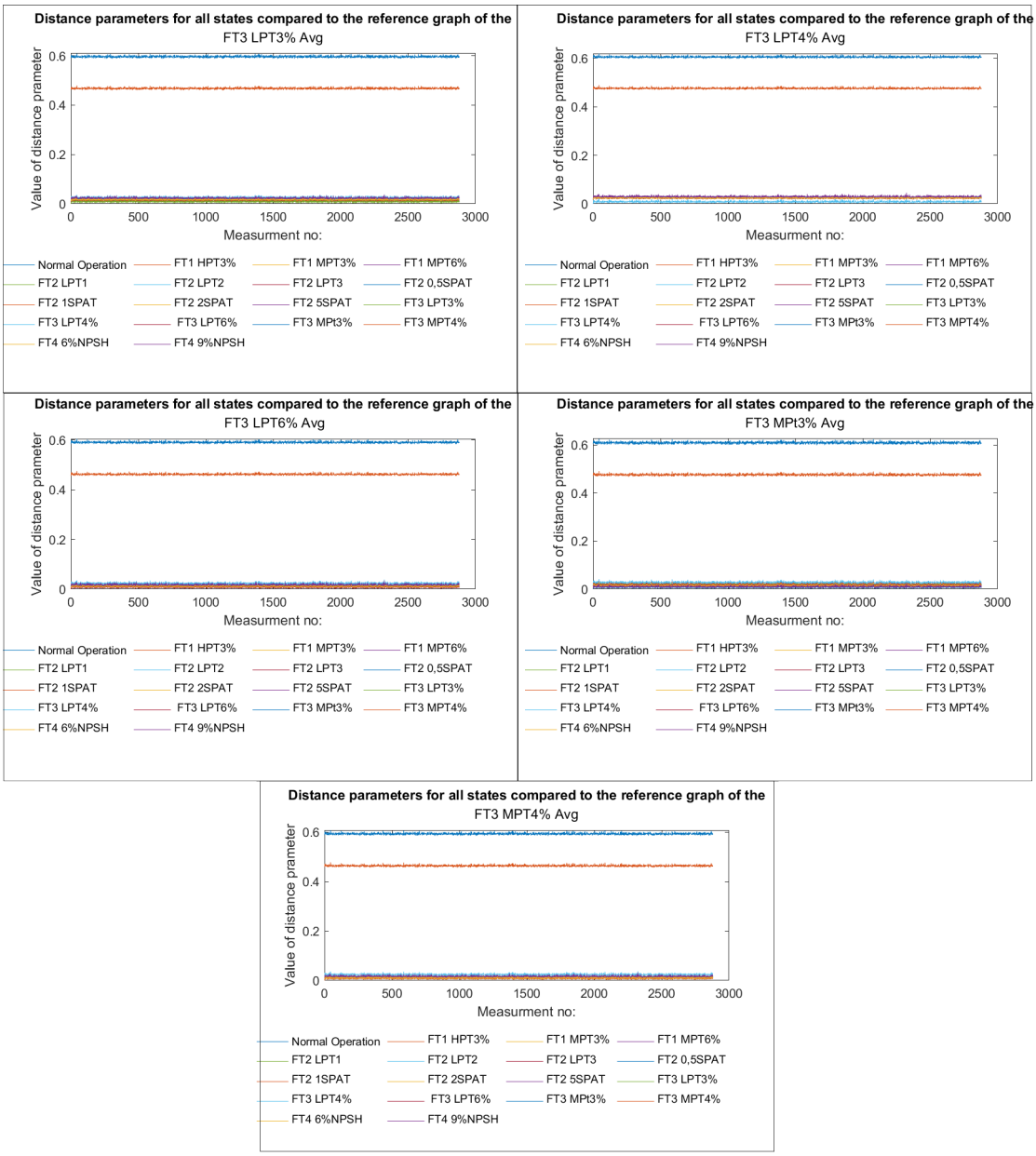


Figure C.4: Isolation results in terms of the distance parameter plots for FT 3.

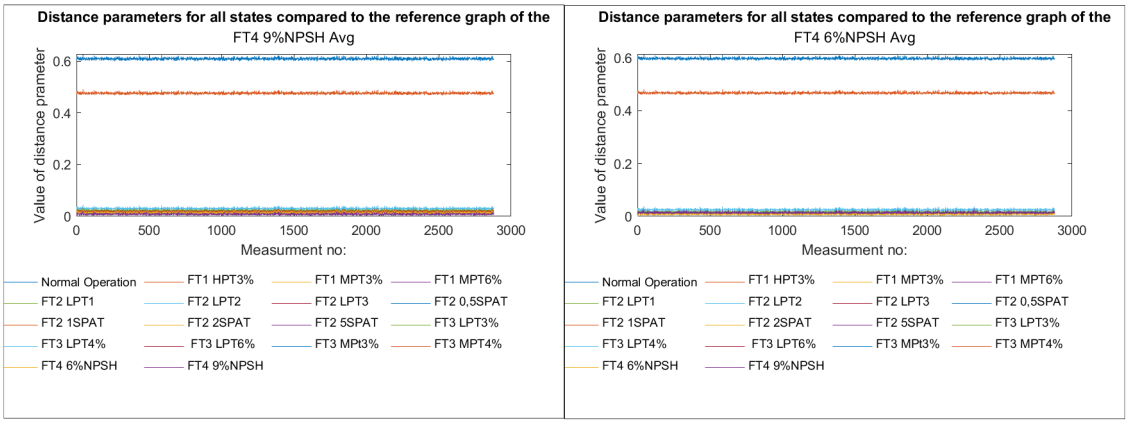


Figure C.5: Isolation results in terms of the distance parameter plots for FT 4.

C.1.2 Approach 3: Residual graphs

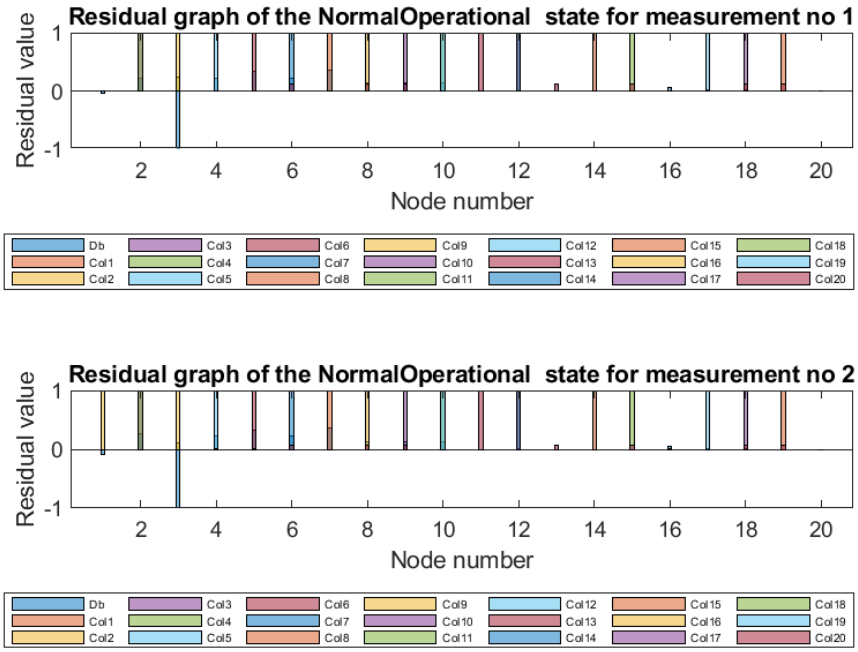


Figure C.6: Residual fault signatures obtained for 2 of the normal graphs.

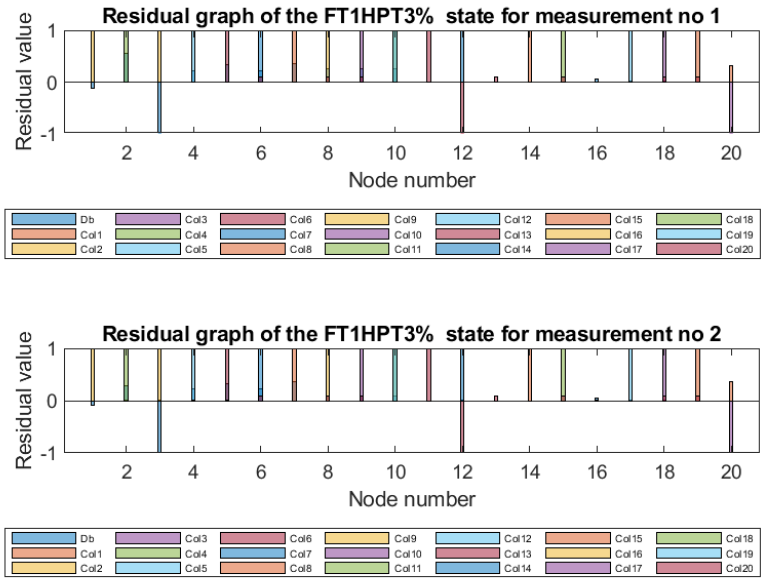


Figure C.7: Residual fault signatures obtained for 2 of the FT1_HPT3% states.

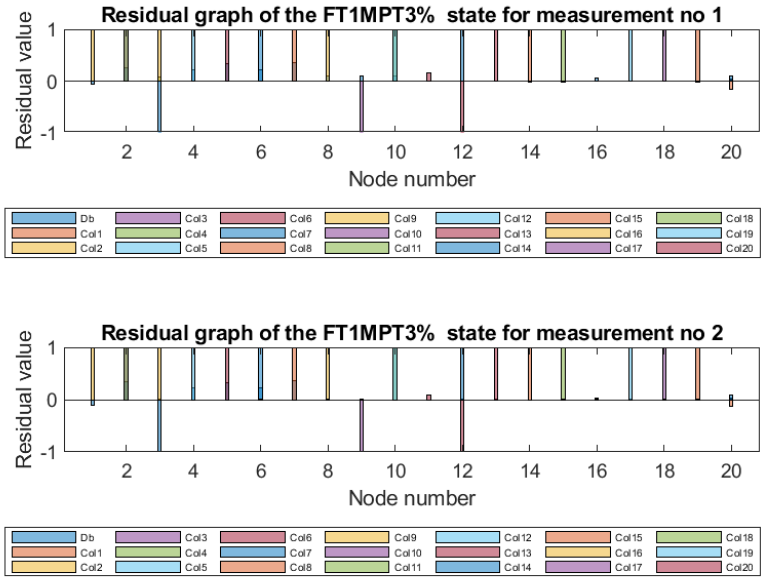


Figure C.8: Residual fault signatures obtained for 2 of the FT1_MPT3% graphs.

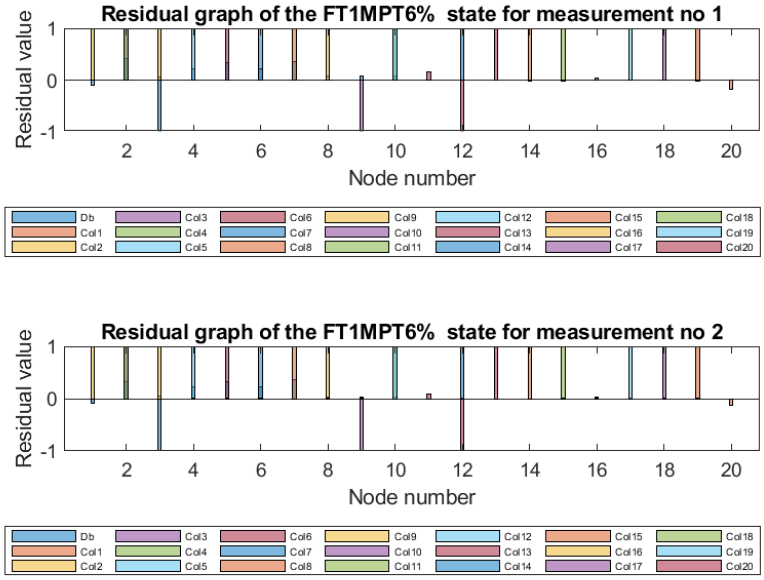


Figure C.9: Residual fault signatures obtained for 2 of the TF1_MPT6% graphs.

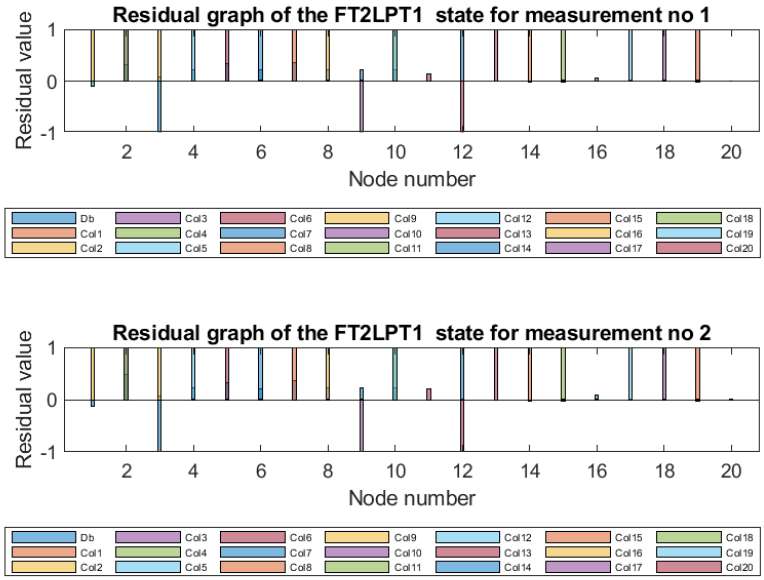


Figure C.10: Residual fault signatures obtained for 2 of the FT2_LPT1 graphs.

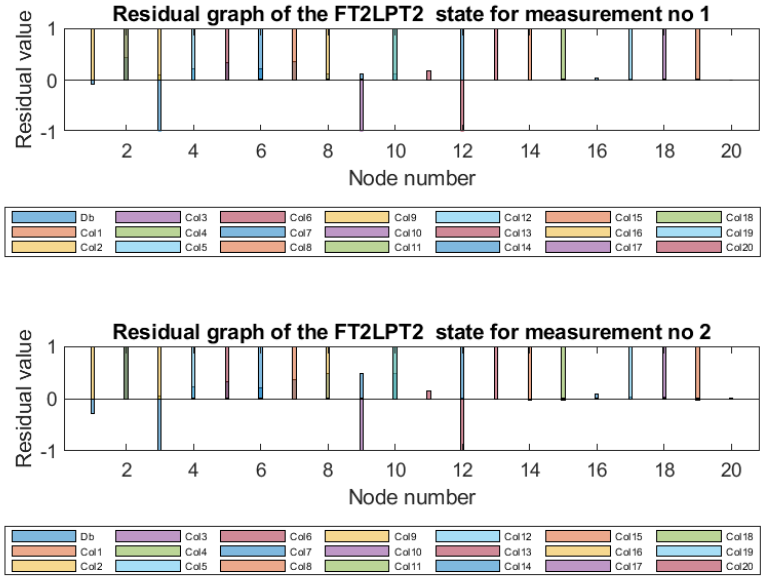


Figure C.11: Residual fault signatures obtained for 2 of the FT2_LPT2 graphs.

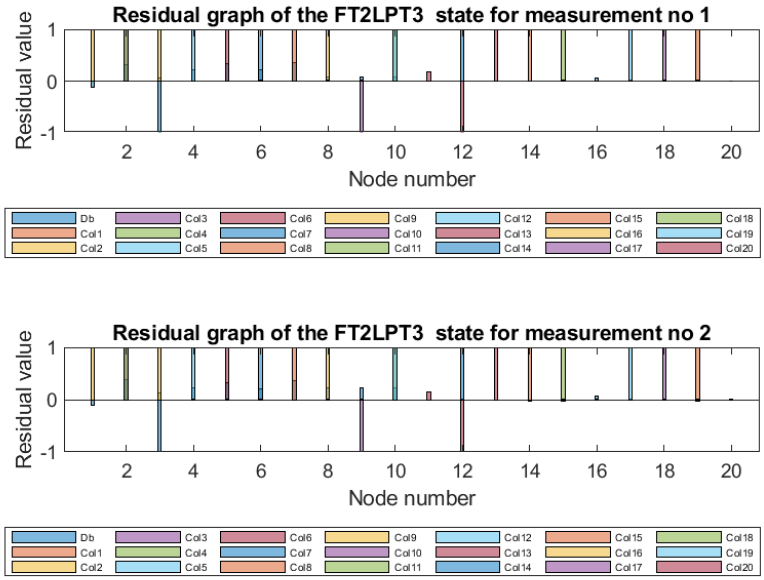


Figure C.12: Residual fault signatures obtained for 2 of the FT2_LPT3 graphs.

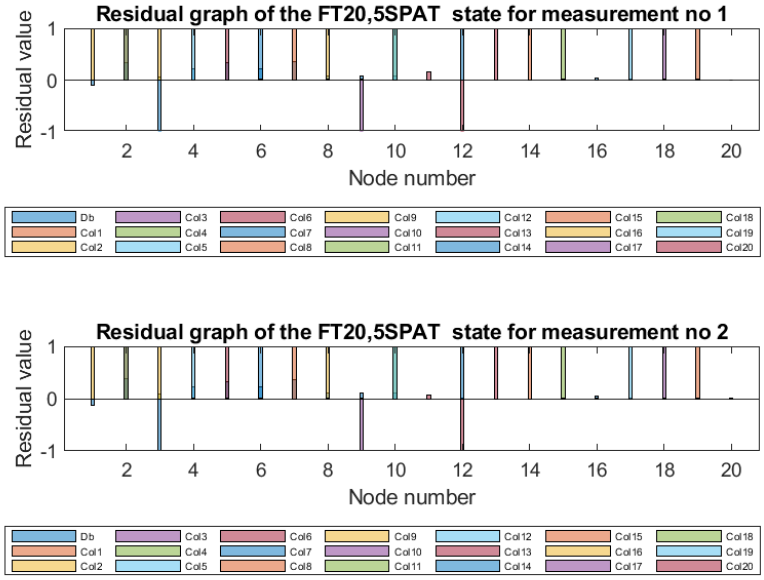


Figure C.13: Residual fault signatures obtained for 2 of the FT2_0.5SPAT graphs.

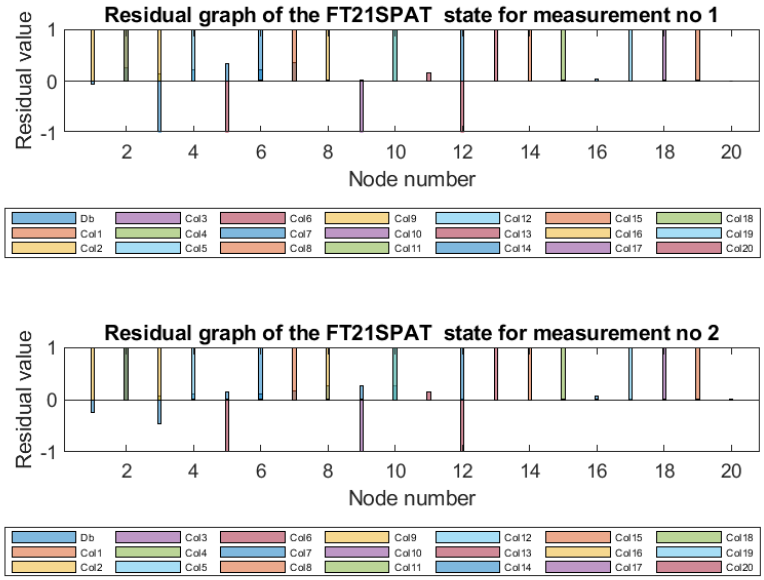


Figure C.14: Residual fault signatures obtained for 2 of the FT2_1SPAT graphs.

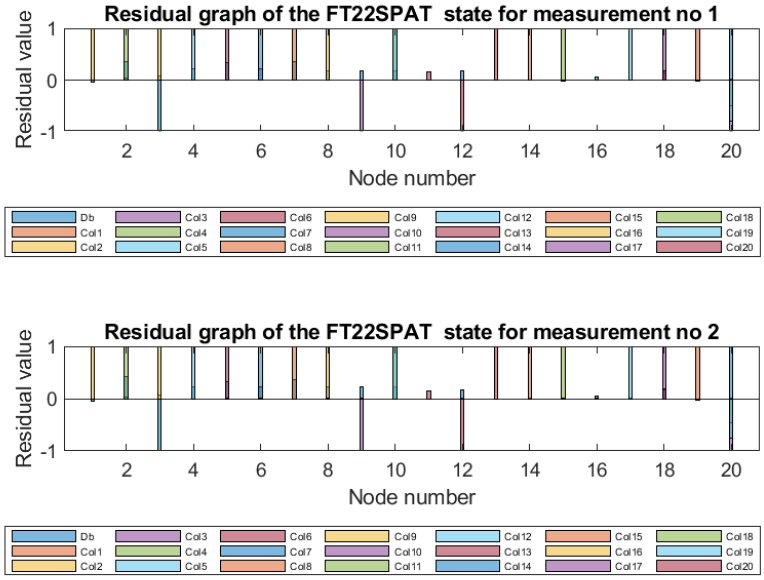


Figure C.15: Residual fault signatures obtained for 2 of the FT2_2SPAT graphs.

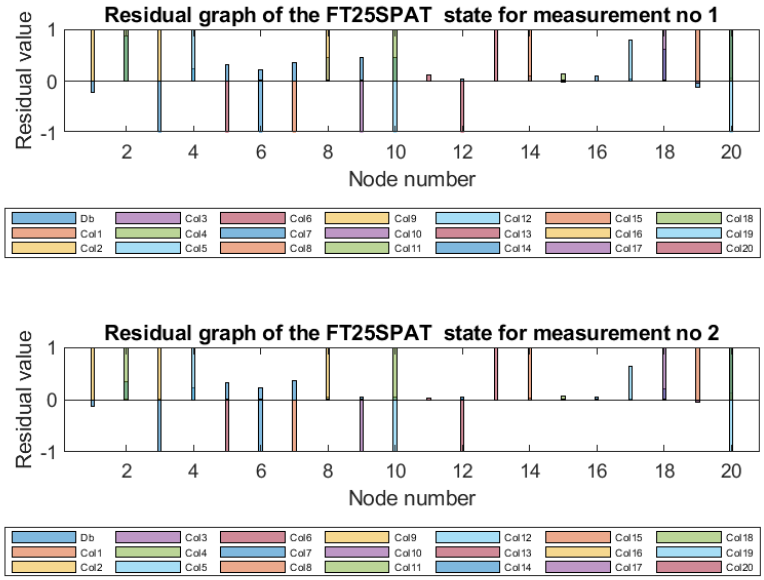


Figure C.16: Residual fault signatures obtained for 2 of the FT2_5SPAT graphs.

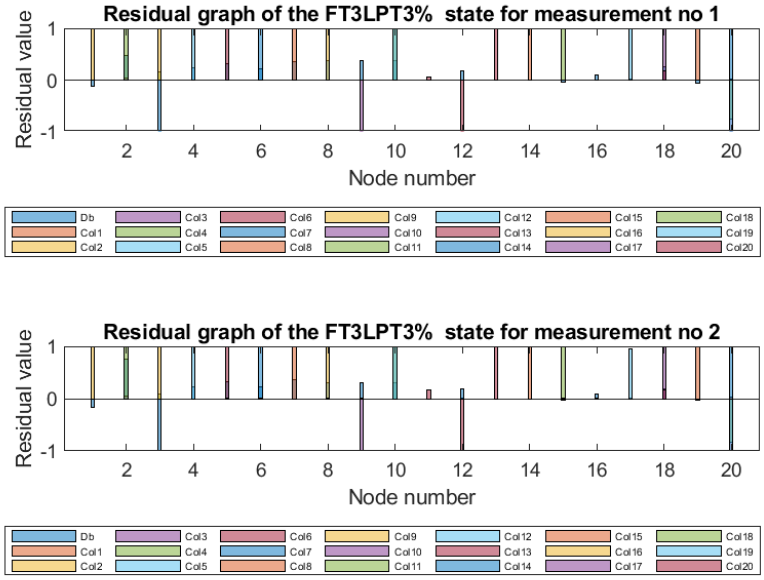


Figure C.17: Residual fault signatures obtained for 2 of the FT3_LPT3 graphs.

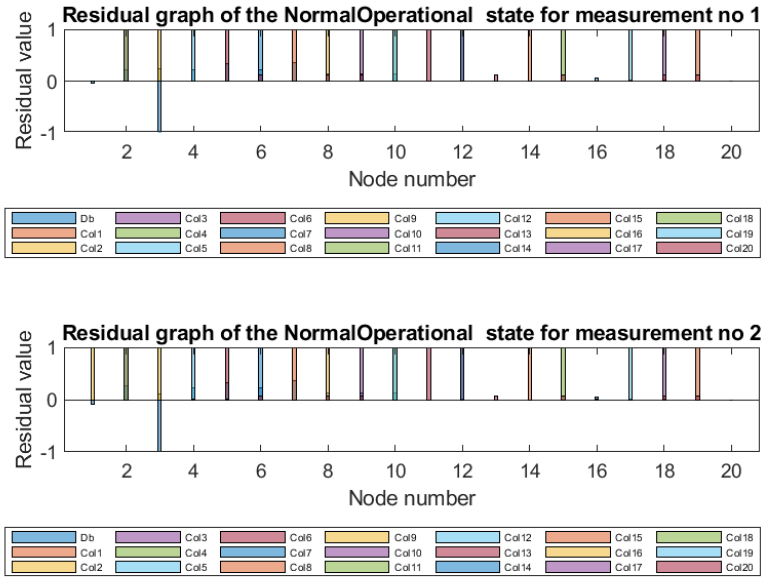


Figure C.18: Residual fault signatures obtained for 2 of the normal graphs.

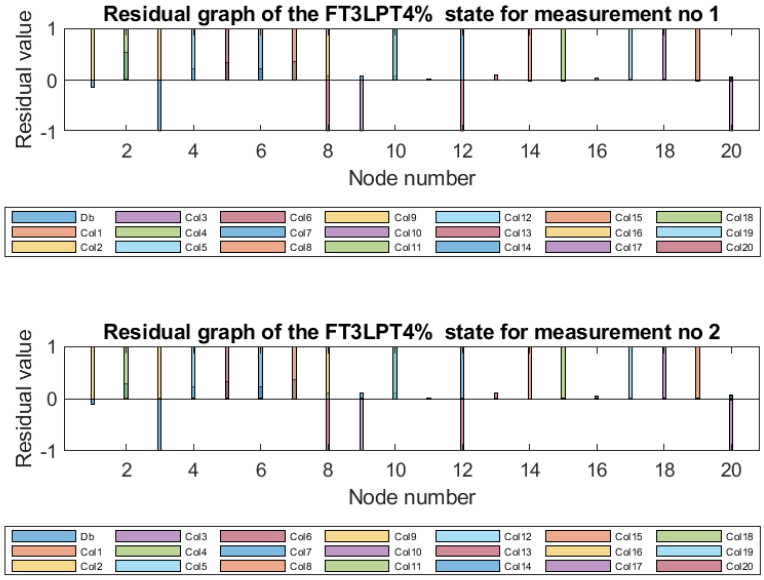


Figure C.19: Residual fault signatures obtained for 2 of the FT3_LPT4 graphs.

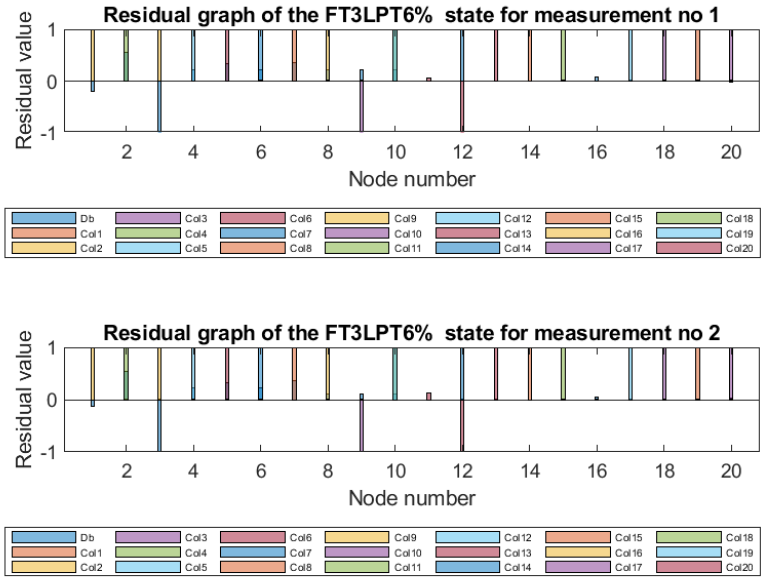


Figure C.20: Residual fault signatures obtained for 2 of the FT3_LPT6 graphs.

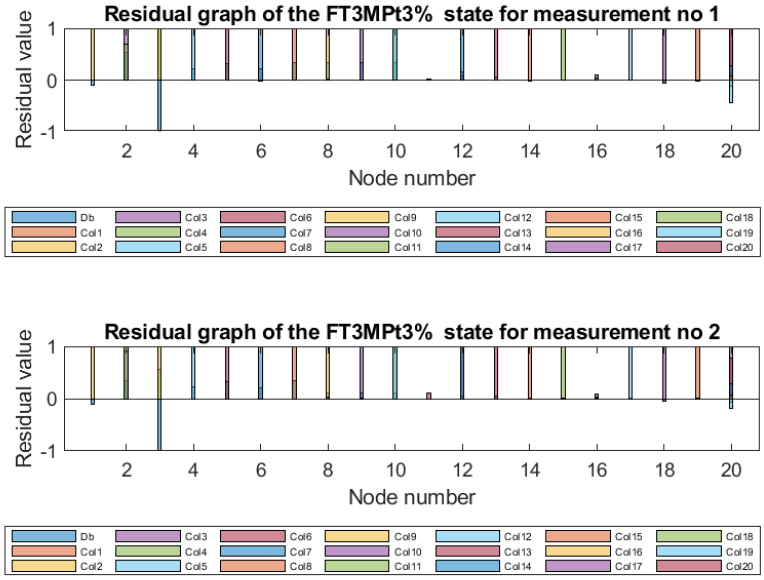


Figure C.21: Residual fault signatures obtained for 2 of the FT3_MPT3 graphs.

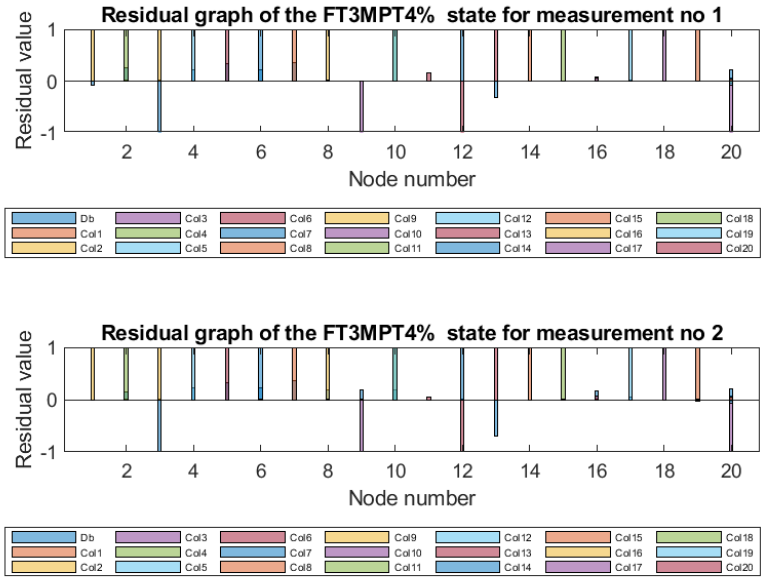


Figure C.22: Residual fault signatures obtained for 2 of the FT3_MPT4 graphs.

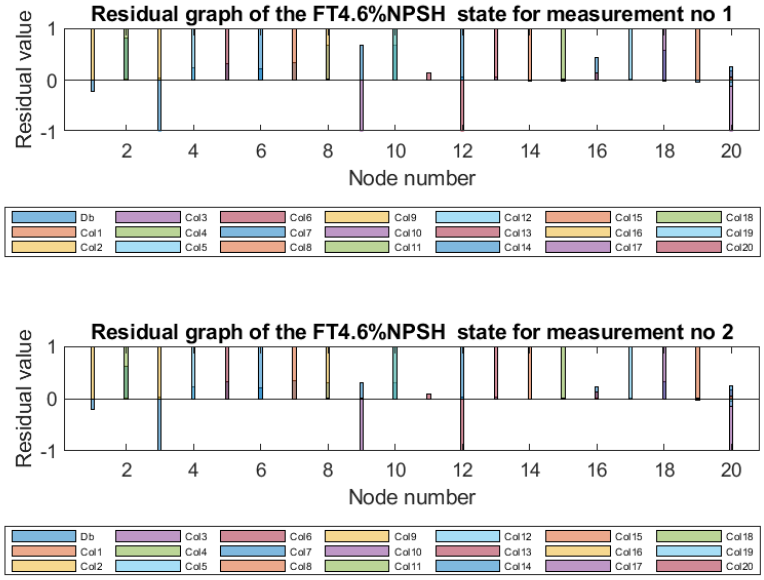


Figure C.23: Residual fault signatures obtained for 2 of the FT4_6NPSH graphs.

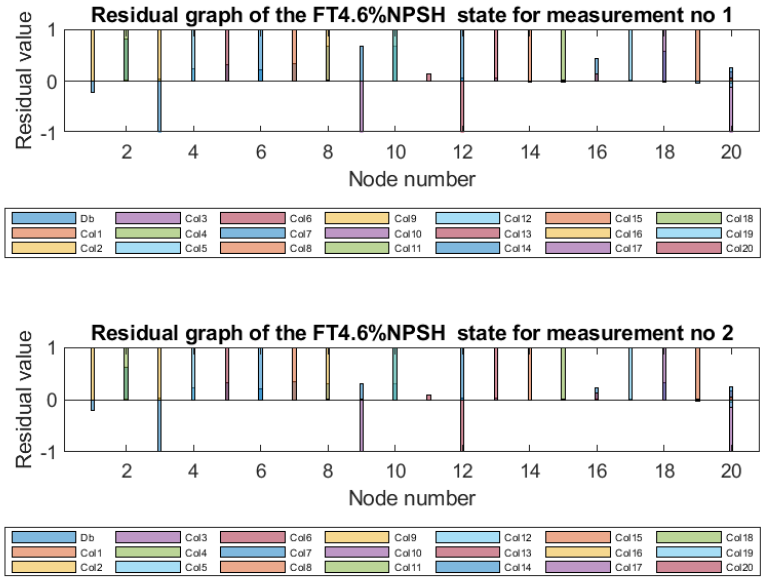


Figure C.24: Residual fault signatures obtained for 2 of the FT4_NPSH6 graphs.

APPENDIX D

Eigenvalue and vector fundamentals

To describe the fundamental theory related to eigenvalues and eigenvector for the application of this study consider the matrix in (D.1).

$$\mathbf{A} = \begin{pmatrix} a & 1 \\ 0 & b \end{pmatrix} \tag{D.1}$$

When multiplying this matrix with any vector of the correct dimensions, the matrix will transform this vector in a specific way. For most vectors this transformation will include a rotation and a change in the vector's length. For eigenvectors, however, only a change in length equal to the size of the eigenvalue will occur, while the vector after the transformation stays on the span of the vector before the transformation. This property of eigenvectors indicate an intuitive advantage for the application of FDI. For a vector,

$$\mathbf{v} = \begin{pmatrix} c \\ d \end{pmatrix}, \tag{D.2}$$

the transformation given by $\mathbf{Av}$ can be visualized as shown in Figure D.1 for the variables a and b of matrix $\mathbf{A}$, having values that resulted in $\mathbf{v}$ being a non-eigenvector and an eigenvector of $\mathbf{A}$ respectively. The plot clearly shows that the eigenvector is only elongated with the size of the eigenvalue while staying on the span of the original vector. In contras with this the non-eigenvector is both rotated and elongated.

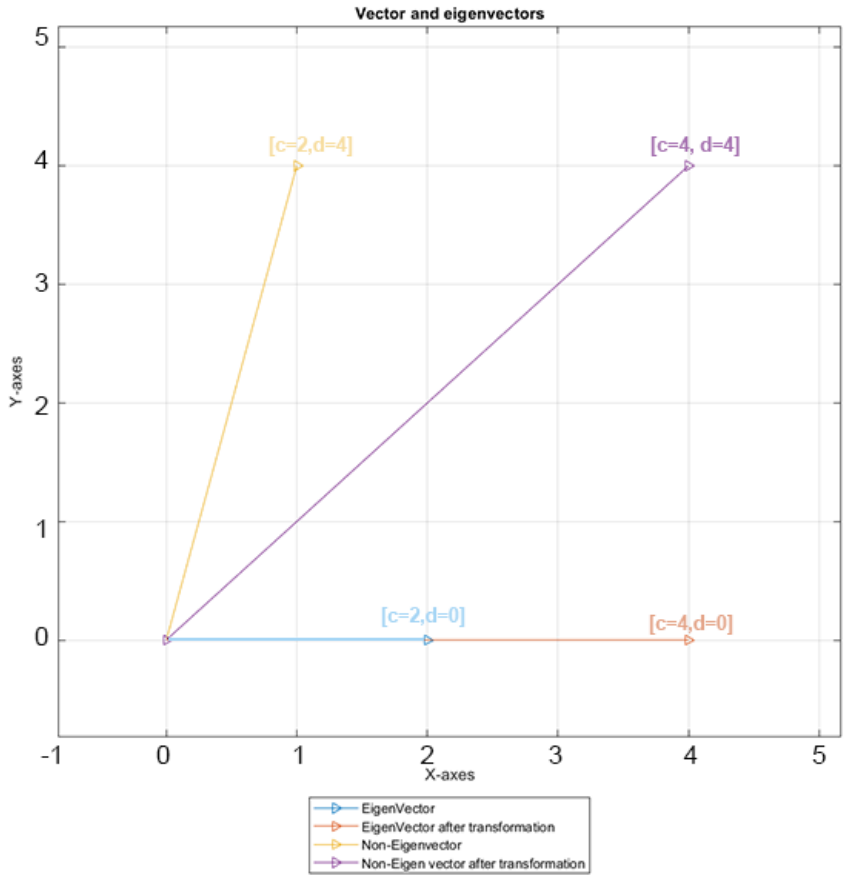


Figure D.1: Example of the effect of matrix transformation on a vector if the vector is an eigenvector or not

Since eigenvectors have this unique property the basic question that can be asked is which vector ($\mathbf{v}$) has the property, that when multiplying $\mathbf{v}$ with the matrix ($\mathbf{A}$), the resulting vector would simply be equal to some eigenvalue (Λ) multiplied with $\mathbf{v}$. Mathematically the expression that must be solved can be written as shown in (D.3),

$$\mathbf{A}\mathbf{v} = \Lambda\mathbf{v}. \tag{D.3}$$

Since (D.3) represents matrix multiplication, on the left-hand side, equal to scalar multiplication on the right-hand side, it can be rewritten to represent only matrix multiplication by changing Λ to a diagonal matrix containing all the eigenvalues on the diagonal. The resulting equation applied to matrix $\mathbf{A}$ and vector $\mathbf{v}$ of this example, is shown in (D.4),

$$\begin{pmatrix} a & 1 \\ 0 & b \end{pmatrix} \times \mathbf{v} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \times \mathbf{v} \tag{D.4}$$

Through inspection it can easily be seen that matrix $\mathbf{A}$ is linked to the eigenvalue matrix in terms of the diagonals. This implies that the eigenvalues gives an indication of the effects the diagonal of matrix $\mathbf{A}$ has on the eigenvector. This is even clearer when rewriting (D.4) as shown in (D.5)

$$\begin{pmatrix} a-\lambda_1 & 1 \\ 0 & b-\lambda_2 \end{pmatrix} \times \mathbf{v} = \vec{0} \tag{D.5}$$

Note that the non-diagonal values (0 and 1) of the matrix A are not effected by the eigenvalues.