



Energy graph-based visualisation for FDI on a thermal power plant with heat storage capabilities

HK Jacobie



[orcid.org/ 0000-0002-7550-8234](https://orcid.org/0000-0002-7550-8234)

Dissertation accepted in fulfilment of the requirements for the degree
Master of Engineering in Mechanical Engineering at the
North-West University

Supervisor: Prof G van Schoor

Co-supervisor: Prof KR Uren

Assistant-supervisor: Prof M van Eldik

Graduation: June 2025

Acknowledgements

”Have I not commanded you? Be strong and courageous. Do not be afraid; do not be discouraged, for the Lord your God will be with you wherever you go.” – Joshua 1:9

I would like to express my gratitude to the following individuals and institutions, in no specific order, for their support, contributions, and efforts towards the completion of this dissertation:

- To my mother and father: Thank you for your unwavering love and support. You are a safe harbour for any ship struggling to maintain its course on a rough sea. I am grateful for your emotional, financial, and spiritual support. You have taught me that life begins at the end of one’s comfort zone and that I, like everyone, am capable of more than I may think.
- To my wife Imke: Thank you for being there for me and for motivating me to be the best person I can be. I truly appreciate and love you.
- To North-West University: Thank you for all the opportunities you have provided. You have given me the platform to become a better person in all aspects of life.
- Thank you, Prof. George and Prof. Kenny, for your guidance and mentorship throughout this work. Your time, patience, and wisdom will always inspire me in everything I do.
- To the University of Applied Sciences Zittau/Görlitz, Institut für Prozesstechnik, Prozessautomatisierung und Messtechnik (IPM): Thank you for your hospitality during my time in Germany, as well as for granting me access to your facilities and providing knowledgeable insights into this work.

“This work is based on the research supported in part by the National Research Foundation of South Africa (Grant Numbers 138606 and 147669)” The Grantholder acknowledges that opinions, findings and conclusions or recommendations expressed in any publication generated by the NRF supported research is that of the author(s) and that the NRF accepts no liability whatsoever in this regard.

Abstract

This study explores the application of energy graph-based visualisation as a fault detection and isolation tool on the thermal energy storage test facility (THERESA), a complex water-steam system with integrated heat storage capabilities. Energy graph-based visualisation employs energy as a unifying parameter across physical domains to characterise a system using graph theory to maintain the system's structural integrity. Key components are represented as nodes, with energy interactions between them represented as links, forming a system-attributed graph.

The analysis involves computing energy and exergy values to characterise system behaviour, creating node signature matrices from time series data. A sensitivity analysis is conducted to compare two energy graph-based visualisation analysis methods: distance parameter approach and the sample-by-sample singular value decomposition approach. The sensitivity analysis is implemented through the use of matrix perturbations the results of which indicates that the sample-by-sample singular value decomposition method is more sensitive to changes in the node signature matrices.

This study also demonstrates the possibility for fault detection and isolation from energy graph-based visualisation on the THERESA system and also implements the distance parameter method and the sample-by-sample singular value decomposition method as energy graph-based visualisation analysis methods. For fault detection and isolation (FDI), thresholds are developed from two normal operational data sets to evaluate a fault data set with two known faults. The results of the FDI performance indicate that the distance parameter approach is superior in fault detection with both faults detected with a true alarm rate of 99.34% with the sample-by-sample singular value decomposition approach utilising three singular values achieving a true alarm rate of 44.46% with singular value 1.

The findings indicate that while the sample-by-sample singular value decomposition approach is more sensitive to changes in the system node signature matrix, the distance parameter approach provides better overall FDI performance, making it more suitable for fault detection in complex systems like THERESA.

Keywords: Energy graph-based visualisation; fault detection and isolation; power plant; thermal energy storage

Contents

1	Introduction	1
1.1	Background	1
1.2	Research question	3
1.3	Objectives	3
1.3.1	Overview and the THERESA system	4
1.3.2	Implementation of EGBV on the THERESA system	4
1.3.3	Verification of EGBV application	4
1.3.4	EGBV analysis	4
1.3.5	Sensitivity analysis	4
1.3.6	Possibility of FDI from EGBV for the THERESA system	4
1.4	Methodology	5
1.4.1	Overview and the THERESA system	5
1.4.2	Implementation of EGBV on the THERESA system	5
1.4.3	Verification of EGBV application	5
1.4.4	EGBV analysis	5
1.4.5	Sensitivity analysis	6
1.4.6	The application of EGBV for FDI purposes	6
1.5	Dissertation outline	6
2	Literature	8

2.1 Literature survey	8
2.1.1 Background on FDI	8
2.1.2 Energy-based methods and development of EGBV for FDI	10
2.1.3 Research after the establishment of EGBV for FDI	14
2.1.4 Case studies	16
2.1.5 Critical evaluation of literature	19
2.2 Literature study	20
2.2.1 EGBV by diagnostic system spaces	20
2.2.2 System measurement data and information	21
2.2.3 System graph setup	22
2.2.4 Energy and exergy analysis	26
2.2.5 Attributed graph and node signature matrix	31
2.2.6 Cost matrix by matrix comparison	32
2.2.7 EGBV analysis methods for FDI	33
2.3 Conclusion	35
3 THERESA system	36
3.1 Introduction	36
3.2 THERESA background	36
3.3 System layout and components	38
3.3.1 Data acquisition components	38
3.3.2 Preheaters	38
3.3.3 Steam generator and super-heater	39
3.3.4 Heat sink: Discharge vessel with cooling system	40
3.3.5 Universal interface for component integration	40
3.3.6 Constant pressure displacement storage with mixing preheater	40
3.3.7 Other components	41

3.4	System operation	41
3.4.1	System experimental configuration	42
3.4.2	Experiment goal	42
3.4.3	Identifying system modes	44
3.5	System process monitoring	46
3.6	Conclusion	46
4	EGBV on the THERESA system	47
4.1	Introduction	47
4.2	EGBV methodology outline	47
4.3	System graph setup	49
4.3.1	Process knowledge	49
4.3.2	System structural information	51
4.3.3	Mass and energy balance equations	52
4.3.4	Node selection and link determination	53
4.4	System-attributed graph setup	57
4.4.1	Defining node energy interactions	58
4.4.2	Energy and exergy attribute determination	61
4.4.3	Assumptions underlying the energy characterisation	61
4.4.4	Calculation of unknown process variables	62
4.4.5	THERESA system-attributed graph	63
4.4.6	THERESA system node signature matrix	64
4.4.7	THERESA system node signature matrices	64
4.5	Conclusion	65
5	Verification of energy characterisation	66
5.1	Introduction	66

5.2	EGBV method verification criteria	66
5.2.1	Process knowledge	68
5.2.2	Sensor placement	68
5.2.3	Ideal heat transfer and power transfer assumptions	68
5.3	Assessment of underlying assumptions	69
5.3.1	Verification by mass and energy balance	69
5.3.2	Data availability and process knowledge limitations	72
5.3.3	Data handling	73
5.4	Conclusion	73
6	EGBV analysis	74
6.1	Introduction	74
6.2	EGBV analysis methodology	74
6.3	Data utilised	75
6.3.1	Intervals for modes in data set A and data set B	76
6.4	EGBV analysis using fixed reference DP method	77
6.4.1	Reference node signature matrix determination	78
6.4.2	Cost matrix calculation - DP method	79
6.4.3	Fixed reference DP method visualisation	80
6.4.4	Characteristics of the DP method	81
6.5	EGBV analysis using the SBS SVD method	83
6.5.1	Cost matrix calculation - SBS SVD method	84
6.5.2	Singular value decomposition - SBS SVD method	84
6.5.3	SBS SVD method visualisation	87
6.5.4	Characteristics of SBS SVD method	88
6.6	Sensitivity analysis by node-signature perturbation	92
6.6.1	Sensitivity analysis methodology	93

6.6.2	Sensitivity analysis using the DP method	94
6.6.3	Sensitivity analysis using the SBS SVD method	95
6.7	Conclusion	96
7	Results	98
7.1	Introduction	98
7.2	EGBV application on the THERESA system	98
7.3	EGBV analysis results	99
7.3.1	Results of EGBV analysis using the DP method	99
7.3.2	Results of EGBV analysis using SBS methodology	100
7.4	Sensitivity analysis results	100
7.5	Conclusion	102
8	The application of EGBV for FDI purposes	103
8.1	Introduction	103
8.2	FDI on the THERESA system using the DP method	105
8.3	FDI on the THERESA system using the SBS SVD method	109
8.4	Conclusion	119
9	Conclusion and recommendations	121
9.1	Introduction	121
9.2	Reflection on research objectives	121
9.2.1	THERESA system overview and data acquisition	121
9.2.2	Application of EGBV on the THERESA system	122
9.2.3	Verification of EGBV applied on the THERESA system	122
9.2.4	EGBV analysis	122
9.2.5	Sensitivity analysis	122
9.2.6	The application of EGBV for FDI purposes	122

9.3	Recommendations for future work	123
9.3.1	Increase information within the node-signature matrix	123
9.3.2	Utilisation of structural information	123
9.3.3	EGBV generalisation	124
9.3.4	Reference generation in mode varying systems	124
9.4	Further research on the THERESA system	124
9.5	Closure	125

List of Figures

2.1	Illustration of the h-s graph FD approach from [1]	12
2.2	(a) Process diagram of the heated two tank system. (b) Graph representation of the heated two tank system. From [2]	16
2.3	Steam turbine system simplified system schematic representation from [3]	17
2.4	Stages of fault signature development in EGBV for FDI from [4]	20
2.5	(a) simplified representation of a coal-fired power station. (b) Attributed graph of a coal-fired power station. From [5]	23
2.6	Energy balance across a generic node	27
2.7	Exergy balance across a generic node	28
2.8	Heat exchanger represented as a single node	29
2.9	Heat exchanger represented as two nodes(primary side and secondary side)	30
3.1	Simplified THERESA system diagram	37
3.2	THERESA system schematic used in Chapter 4 for determining the system graph.	37
3.3	Theresa system 3D layout from [6]	38
3.4	Preheater from [6]	39
3.5	Steam Generator and superheater from [6]	39
3.6	Heat sink and cooling circuit from [6]	40
3.7	Constanet pressure displacement storage component from [6]	41
3.8	CPDS charge hold and discharge modes from [7]	43
3.9	CPDS control valves on flow paths in and out of CPDS	44

3.10 CPDS temperature during operation (data set A)	45
3.11 CPDS temperature during operation (data set B)	46
4.1 System flow for attributed graph configuration	48
4.2 THERESA system schematic	51
4.3 Possible system nodes in THERESA system highlighted in red	54
4.4 Full system graph with all possible THERESA system nodes included	55
4.5 Full system graph before graph reduction, with reduction areas highlighted with red boundary	57
4.6 Full system after reduction	57
4.7 Generic node with four IOs	58
4.8 Generic representation of link definitions	59
4.9 Horizontally orientated tank with torispherical head from [8]	63
4.10 THERESA system-attributed graph	64
6.1 DP Approach - distance parameter of data set A visualised over all intervals with fixed reference in holding mode	80
6.2 DP Approach - distance parameter of data set B visualised over all intervals with fixed reference in holding mode	81
6.3 Mean DP value for data sets A and B over all modes	82
6.4 Cumulative variance percentage vs number of singular values for data sets A, B and C	85
6.5 Normalised average sum of absolute differences vs window length (L) for data sets A, B and C	86
6.6 SBS Approach - first three singular values of data set A visualised over all intervals, with modes indicated	87
6.7 SBS Approach - first three singular values of data set B visualised over all intervals, with modes indicated	88
6.8 Mean of each singular value over all system modes for data set A	89
6.9 Mean of each singular value over all system modes for data set B	89

6.10 Mean value of SV-1 for data sets A and B over all modes	90
6.11 Mean value of SV-2 for data sets A and B over all modes	90
6.12 Mean value of SV-3 for data sets A and B over all modes	91
6.13 Reference thresholds for DP method sensitivity analysis of dataset A	95
6.14 Reference thresholds for SBS SVD method sensitivity analysis of dataset A	
- First singular value	96
6.15 Reference thresholds for SBS SVD method sensitivity analysis of dataset A	
- Second singular value	97
6.16 Reference thresholds for SBS SVD method sensitivity analysis of dataset A	
- Third singular value	97
7.1 Distance parameter method and SBS method alarm rate with percentage	
perturbation	101
8.1 Data set A plotted with NOC thresholds - DP approach	107
8.2 Data set B plotted with NOC thresholds - DP approach	107
8.3 Data set C plotted with NOC thresholds - DP approach	108
8.4 Data set C plotted with NOC thresholds and known fault intervals high-	
lighted - DP approach	108
8.5 Data set A plotted with combined NOC thresholds - SBS SVD approach	
for SV1	112
8.6 Data set A plotted with NOC thresholds - SBS SVD approach for SV2	112
8.7 Data set A plotted with NOC thresholds - SBS SVD approach for SV3	113
8.8 Data set B plotted with NOC thresholds - SBS SVD approach for SV1	113
8.9 Data set A plotted with NOC thresholds - SBS SVD approach for SV2	114
8.10 Data set A plotted with NOC thresholds - SBS SVD approach for SV3	114
8.11 Data set C plotted with NOC thresholds - SBS SVD approach for SV1	115
8.12 Data set C plotted with NOC thresholds and known fault intervals high-	
lighted - SBS SVD approach for SV1	115
8.13 Data set C plotted with NOC thresholds - SBS SVD approach for SV2	116

8.14 -Data set C plotted with NOC thresholds and known fault intervals high-	
lighted - SBS SVD approach for SV2	117
8.15 Data set C plotted with NOC thresholds - SBS SVD approach for SV3 . . .	117
8.16 Data set C plotted with NOC thresholds and known fault intervals high-	
lighted - SBS SVD approach for SV3	118

List of Tables

2.1	Condition monitoring terminology and definitions table	9
2.2	Applications of the EGBV method adapted from [9]	11
4.1	THERESA system process variable measurements and units	50
4.2	Description of final selected system graph nodes	56
4.3	System graph energy interactions	60
4.4	Table of underlying assumptions necessary for energy characterisation	62
5.1	Energy balance for data set A	71
6.1	Operational parameters for data sets A and B during experimental data acquisition	76
6.2	Data-set A and B approximate mode starting positions	76
6.3	Reference comparison by distance parameter method data set A	78
6.4	Reference comparison by distance parameter method data set B	79
6.5	Data set A - Mean DP value	82
6.6	Data set B - Mean DP value	83
6.7	Average singular value of first 3 singular values over mode intervals for data set A	91
6.8	Average singular value of first 3 singular values over mode intervals for data set B	92
6.9	Reference threshold values for singular value sensitivity analysis	95
7.1	Table of perturbation factors and distance parameter percentage AR	100

7.2 Table of perturbation factors and singular value percentage AR	101
8.1 Data-set C approximate mode starting positions	105
8.2 Data set C - Fault descriptions and intervals	105
8.3 Individual upper and lower NOC threshold values for each mode for data set A and B - DP approach	106
8.4 Combined upper and lower NOC threshold values for each mode - DP ap- proach	106
8.5 FDI performance metrics for data set C evaluated with NOC thresholds - DP approach	109
8.6 Upper and lower threshold values for each mode of data set A and B - SBS approach for SV1	109
8.7 Upper and lower threshold values for each mode of data set A and B - SBS approach for SV2	110
8.8 Upper and lower threshold values for each mode of data set A and B - SBS approach for SV3	110
8.9 Combined upper and lower threshold values for each mode combined across NOC- SBS approach for SV1	110
8.10 Combined upper and lower threshold values for each mode combined across NOC- SBS approach for SV2	111
8.11 Combined upper and lower threshold values for each mode combined across NOC- SBS approach for SV3	111
8.12 FDI performance metrics for data set C evaluated with NOC thresholds - SBS SVD approach - SV1	118
8.13 FDI performance metrics for data set C evaluated with NOC thresholds - SBS SVD approach - SV2	119
8.14 FDI performance metrics for data set C evaluated with NOC thresholds - SBS SVD approach - SV3	119

Table of abbreviations

A table containing a list of abbreviations that will be used throughout text.

AR	alarm rate
ASME	American Society of Mechanical Engineers
CO₂	carbon dioxide
CPDS	constant pressure displacement storage
CSTR	continuous stirred tank reactor
CVP	cumulative variance percentage
DP	distance parameter
EBGV	energy graph-based visualisation
EBV	energy based visualisation
EWMA	exponentially weighted moving average
FAR	false alarm rate
FD	fault detection
FDD	fault detection and diagnosis
FDI	fault detection and isolation
GTL	gas to liquids
HEOM	heterogeneous Euclidean overlap metric
HTGR	high-temperature gas reactor
HX	heat exchange
IO	input/output
IoT	Internet of Things
IPM	Institute for Process Technology, Process Automation and Measuring Technology

LTH	lower threshold
MP	mixing preheater
NOC	normal operational conditions
PCA	principal component analysis
PF	perturbation factor
SAD	sum of absolute differences
SBS	sample-by-sample
STS	steam turbine system
SV	singular value
SVD	singular value decomposition
TAR	true alarm rate
THERESA	thermal energy storage test facility
UAS	University of Applied Sciences
UTH	upper threshold

Chapter 1

Introduction

1.1 Background

Modern large-scale industrial plants operate as complex systems consisting of intricate subsystems that perform processes across multiple physical domains, including thermodynamic, electrical, chemical, and mechanical. The efficiency of these processes depends on a combination of simple and sophisticated components, each designed to function under various conditions and parameters. These systems may exhibit faulty behaviour due to a range of factors, requiring the implementation of fault detection and isolation (FDI) methods as indispensable tools for monitoring processes and ensuring effective process control [10, 11]. These methods play a crucial role in identifying and isolating faults at both the system and component levels, which in turn enhances overall system performance and reliability. Within the field of fault detection schemes, three overarching categories exist, namely model-based methods, data-driven methods, and knowledge-based methods. Contemporary FDI methodologies often incorporate aspects of these approaches in tandem, leading to the development of hybrid methods [3, 12].

One promising hybrid method for FDI is energy graph-based visualisation (EGBV), which employs energy as a unifying parameter to characterise a system and serves as a foundation for further FDI analysis through various techniques. Originally introduced as energy-based visualisation (EBV) [5], the formalisation of EGBV was later developed through the works of Wolmarans [2] and Smith [3]. This approach has demonstrated considerable potential in the analysis and potential control of large-scale complex systems, with notable results achieved through the application of FDI techniques to energy characterisations [13].

The EGBV approach addresses the challenges associated with FDI in complex industrial systems by leveraging energy as a unifying domain. By constructing an energy-attributed system graph, the approach facilitates fault detection without requiring a detailed system model or extensive sets of labelled fault data. This study explores the application of two key methods within the EGBV framework for FDI: the fixed reference distance

parameter (DP) method and the sample-by-sample singular value decomposition (SBS SVD) method. The DP method is a graph-based approach that quantifies deviations in system behaviour using a fixed reference signature, whereas the SBS SVD method employs statistical decomposition of energy signature matrices to detect anomalies in system performance. These methods allow for the detection and isolation of faults within industrial processes by quantifying deviations in system operation based on energy-based graph structures.

EGBV employs two main operations to characterise a system as an energy-attributed graph with a matrix representation, namely graph theory and energy analysis. The structural information of a system is used to construct a system graph, where key components are represented as nodes and the energy interactions between components are depicted as links. This approach preserves the relative structural information of the system. Energy analysis utilises energy equations to relate the process variables for each link to an energy attribute and to determine an exergy balance across each node. These energy and exergy equations are attributed to the system graph to form the system-attributed graph. The matrix representation of the system's attributed graph is known as the node signature matrix, which is populated with energy and exergy values derived from time-series process measurement data, with each matrix representing an instance of the process. The outputs of the EGBV method applied to a system are node signature matrices that contain energy information for the system at an instance in time. A variety of methods using graph matching or residual-based approaches are subsequently applied to these matrices to facilitate FDI.

The EGBV technique has been applied in a variety of systems for FDI, contributing to the development and validation of the method. Previous studies have demonstrated the applicability of EGBV in simulated environments such as transcritical CO₂ heat pumps by De Bruin [14] and Brayton cycle system such as the system studied by Naser [15]. Additionally, the technique has been successfully applied in practical systems such as the steam turbine system (STS) by Smith [3] and a heated two-tank system by Uren et al. [4] and Wolmarans [2]. The application of EGBV has been tested on real-world large-scale industrial systems, including steam turbine systems, demonstrating its effectiveness and potential for broader applications [3]. Despite these advancements, further research is necessary to evaluate its efficacy in complex experimental test facilities that replicate the conditions of real-world thermal power plants.

One such experimental test facility is the THERESA system, a complex research platform that simulates an operational power plant based on a water-steam cycle. This system incorporates a constant pressure displacement storage (CPDS) unit for thermal energy storage, alongside various critical components such as heat exchangers, boilers, pumps, valves, and tanks [6]. The THERESA system has been widely used in studies on the flexibility of power plants with heat storage systems. Its configurable operational parameters allow for simulated faults, making it an ideal platform for evaluating FDI methodologies. The availability of both normal and faulty operational data enables an investigation

into the extent to which EGBV can be employed for fault detection and isolation on the THERESA system.

As industrial processes become increasingly complex due to the number of subsystems and components, the operation at various operating points, and the transitions through different modes of the system, it becomes essential to critically assess the suitability of the EGBV technique in addressing these challenges. This raises a key research question regarding the extent to which the EGBV energy characterisation method can accommodate the intricacies of complex systems and whether its methodology can be applied effectively to provide abundant measurement data without relying on a simulated model.

The motivation for this research is rooted in the need for a deeper understanding of the characteristics of the EGBV technique in real-world situations where multiple operating parameters come into play, ensuring its continued relevance and effectiveness in advancing FDI methodologies. This study seeks to investigate how energy-graph-based visualisation can be utilised to analyse and characterise the THERESA thermal power plant system with integrated heat storage capabilities. Furthermore, it examines the sensitivity of the DP and SBS SVD methods using alarm rate analysis, assessing the potential of these techniques for future fault detection and isolation applications in complex thermal power plants.

1.2 Research question

How can energy-graph-based visualisation be utilised to analyse and characterise the THERESA thermal power plant system with integrated heat storage capabilities, using the fixed reference distance parameter method and a sample-by-sample singular value decomposition method, and how does the sensitivity analysis based on alarm rate inform the potential for future FDI applications?

1.3 Objectives

The primary aim of this research is to investigate the application of the EGBV method on a complex thermal power plant with heat storage capabilities. Additionally, the study examines the sensitivity of two FDI analysis methods to changes in the node signature matrices. Also, the application of the

To address the research question, the following objectives are established:

1.3.1 Overview and the THERESA system

Conduct an introduction and literature survey to define the scope of the study. Perform a literature review to develop a comprehensive understanding of EGBV principles and techniques utilised in existing research. Provide a detailed system overview of the THERESA system, including its working and relevant operational information. Conduct an in-depth analysis of the THERESA system's functioning and operations. Acquire and process operational data, including both normal and faulty data sets.

1.3.2 Implementation of EGBV on the THERESA system

Apply the EGBV methodology as outlined in the literature on the THERESA system. Configure the system graph based on the THERESA system's structure. Perform energy characterisation and develop the system-attributed graph and node signature matrix.

1.3.3 Verification of EGBV application

Verify the application of the EGBV method on the THERESA system through predefined criteria to evaluate node signature matrices, which are considered the outputs of the EGBV method applied on the THERESA system.

1.3.4 EGBV analysis

Apply the fixed reference distance parameter (DP) method for EGBV analysis. Conduct the sample-by-sample (SBS) singular value decomposition (SVD) method for EGBV analysis.

1.3.5 Sensitivity analysis

Utilise matrix perturbations of node signature matrices to assess the sensitivity of the fixed reference DP method to changes in the node signature matrix. Evaluate the sensitivity of the SBS SVD method to perturbations within the node signature matrices.

1.3.6 Possibility of FDI from EGBV for the THERESA system

Evaluate the FDI capabilities of the EGBV method using the fixed reference DP method and the SBS SVD approach implementing thresholds developed from data sets operating under normal operating conditions and evaluate the methods in terms of FDI performance metrics.

1.4 Methodology

1.4.1 Overview and the THERESA system

The THERESA thermal energy storage test facility is an experimental system designed to simulate a water-steam cycle representative of industrial power plants. It integrates a constant pressure displacement storage (CPDS) unit, heat exchangers, steam generators, and other essential components, allowing for controlled operation under normal and fault-induced conditions. The system enables the evaluation of energy-based methods by providing real-time data on temperature, pressure, and flow rates, forming the basis for implementing and assessing energy graph-based visualisation (EGBV) as a fault detection and isolation (FDI) tool.

1.4.2 Implementation of EGBV on the THERESA system

The EGBV method is applied following the most recent literature, utilising energy characterisation and graph-based representation to structure system data. Key thermodynamic properties and process parameters are used to define nodes and energy interactions as links, forming an energy-attributed system graph. The node signature matrix representation, derived from time-series measurements, provides a structured dataset for further analysis in Chapter 5. The formulation of energy and exergy balance equations across nodes ensures that the system structure aligns with the methodology outlined by and article by Venkatasubramanian et al. [12].

1.4.3 Verification of EGBV application

The validation of EGBV involves assessing the accuracy and reliability of the node signature matrices by comparing experimental data with theoretical predictions. Matrix consistency across different operational conditions is evaluated to ensure robustness, and discrepancies are analysed and discussed through the applied assumptions. The verification process aims to confirm that the generated energy signatures sufficiently represent the system's behaviour.

1.4.4 EGBV analysis

The constructed node signature matrices serve as the foundation for further analysis, where graph-based methods are employed to extract system characteristics. The fixed reference distance parameter (DP) method is applied to quantify deviations by comparing node signatures to a predefined reference, while the sample-by-sample singular value decomposition (SBS SVD) method is used to identify variations indicative of faults through the use of each sample as a reference for itself. With the goal using these methods to

analyse and provide insight into the dynamic behaviour of the THERESA system under normal and fault-induced conditions.

1.4.5 Sensitivity analysis

The performance of EGBV-based fault detection methods is assessed through a sensitivity analysis, where controlled manipulations of values in the node signature matrices called perturbations are introduced in a structured manner. As a metric for sensitivity the concept of alarm rate (AR) is introduced and used as a key indicator of method effectiveness. This analysis provides a deeper understanding of how the proposed methods respond to subtle changes in system operation.

1.4.6 The application of EGBV for FDI purposes

The final stage of the methodology involves applying EGBV as an FDI tool by evaluating its ability to detect and isolate faults within the THERESA system. The detection results are assessed based on alarm rate analysis using true alarm rate (TAR) and false alarm rate (FAR). The findings aim to demonstrate the applicability of EGBV in real-world industrial systems and its potential to enhance FDI methodologies.

1.5 Dissertation outline

Chapter 1 is the introduction chapter that states the background of the study. This chapter contextualises and states the research question. Further, objectives are outlined which address the research question with a concise methodology section stating how each objective will be accomplished to ultimately address the research question.

Chapter 2 is the literature chapter, starting with a literature review focusing on the development of EGBV for FDI, and then follows a comprehensive literature study discussing in detail the relevant literature on FD, process monitoring, EGBV, and energy characterisation. The discussed methods and concepts will form the basis for Chapter 5 as the framework for the practical application of EGBV on the THERESA system.

Chapter 3 is an overview of the THERESA system. First, a broad description and diagram are used to explain at a high level how the system operates, and then the components and functioning of the system are explained in depth for thorough understanding.

Chapter 4 is the EGBV chapter, where the EGBV method is applied to the THERESA system. Different operations associated with the application of the EGBV method are described and applied in the context of the THERESA system. The energy characterisation yields node signature matrices for the operational data that are used in Chapter 7 for the

sensitivity analysis of the energy characterisation.

Chapter 5 is the verification chapter, following the EGBV chapter. This chapter outlines the verification methodology followed to ensure confidence in the node signature data sets that are produced. Assumptions and constraints are highlighted and discussed and various factors influencing the energy characterisation are taken into account in the evaluation of the outputs produced by chapter 4.

Chapter 6 is the EGBV analysis chapter, where two approaches for the analysis of the node signature matrices are investigated and applied to the node signature matrices produced in Chapter 5. These two methods include a fixed reference DP method and a SBS SVD method. This chapter concludes with the sensitivity analysis, where the sensitivity of the two EGBV analysis methods is investigated through perturbations of a normal data set. Using a parameter of alarm rate (AR), the sensitivity of the methods to changes in the node signature matrices is evaluated compared to the unperturbed data set.

Chapter 7 is the results chapter, mainly focused on the results obtained from Chapter 6, providing discussions of the results obtained from the DP analysis, the SBS analysis, and the sensitivity analysis done using these two methods.

Chapter 8 is a chapter demonstrating the FDI possibilities from the EGBV methods applied in chapter 6 with the objective of illustrating possible application for FDI from EGBV on the THERESA system. In this chapter two methods of FDI are utilised: the fixed reference DP method and the SBS SVD analysis method. Within each method, thresholds developed from normal operational data are utilised to evaluate the FDI performance metrics of FAR and TAR on a data set containing known faults for the THERESA system.

Chapter 9 is the conclusion and recommendations chapter. It reflects all the objectives of the study to demonstrate to what extent the study has been successful and where the study is limited according to the results produced. Thus, this chapter expresses the challenges faced and provides detailed recommendations for future studies.

Chapter 2

Literature

2.1 Literature survey

Chapter 1 establishes the foundational context for further research in EGBV for FDI on a large-scale complex thermal power plant, proposing an investigation into the sensitivity of the energy characterisation provided by the EGBV method. This chapter elaborates on the objectives of Chapter 1, offering a comprehensive review of the literature on energy-based FDI. It begins with an overview of FDI, followed by a survey of the literature on the development of energy-based FDI methods. The chapter's central focus is on the application of EGBV in FDI, including case studies and concluding with a critical review of the relevant literature.

2.1.1 Background on FDI

The detection and isolation of faults within industrial processes is a field of study that has continuously evolved. It has progressed from using biological senses—such as observing changes in shape or color, listening for unusual sounds, feeling for vibrations, and detecting fumes by smell—to advanced FDI schemes supported by computer technology [16]. With the ever-increasing complexity of modern industrial systems, the need for more advanced and effective FDI techniques has become increasingly pronounced [17]. In particular, the challenge posed by the vast amounts of data associated with condition monitoring in large-scale industrial plants has necessitated innovative approaches to enhance the performance and computational efficiency of FDI methods.

Within the field of process monitoring, several terms may be used to specify the intention of a scheme or approach. For this study, Table 2.1 gives key terms and their definitions that will be used throughout the study, adopted from Isermann et al. [18].

In general, fault diagnosis methods are classified into three general categories: namely, quantitative model-based methods, qualitative model-based methods, and process history-

Table 2.1: Condition monitoring terminology and definitions table

Term	Definition
Fault detection	Determination of the faults present in a system and the time of detection
Fault isolation	Determination of the kind, location, and time of detection of a fault, follows fault detection
Fault identification	Determination of the size and time-variant behaviour of a fault. Follows fault isolation
Fault diagnosis	Determination of the kind. Size, location, and time detection of a fault. Follows fault detection. Includes fault isolation and identification
Fault	An unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable/usual/standard condition
Error	A deviation between a measured or computed value (of an output variable) and the true, specified, or theoretically correct value
Disturbance	An unknown (and uncontrolled) input acting on a system
Perturbation	An input acting on a system, which results in a temporary departure from the current state
Residual	A fault indicator, based on a deviation between measurements and model-equation-based computations

based methods [17, 19, 20]. Within these three categories, there are a variety of different search techniques that can confuse as to which technique is most applicable in a given system [12]. Model-based methods require accurate modelling of the investigated system, and in cases where dynamic modelling of complex systems is required, this can be a tedious process that may yield less than adequate results [12]. An accurate system model is essential for the accuracy and consistency of results obtained from the model and in large-scale industrial systems detailed information on the system may not be readily available to set up such a model. Thus, a qualitative model of the system utilising limited system information may be beneficial.

In the case of data-driven approaches, no model is required, and the method makes use of historical process data, which in large-scale industrial systems are available through the implementation of sensors throughout the system. However, to accurately train a data-driven method for accurate and consistent results, it is necessary to have sizeable data sets of normal and faulty behaviour. In the case of large-scale industrial systems operational data are available, however, data categorised as faulty operational data are not necessarily available.

Hybrid methods circumvent many of the challenges posed by model-based and data-driven approaches, utilising a combination of these approaches to facilitate search techniques that may be better suited than either technique used exclusively.

Due to the degree of energy transfer of large-scale industrial processes, a method utilising energy as a method of dimensional reduction is also beneficial. Taking into account the data available and the limited information available in the system, a viable hybrid diagnostic method utilising energy as a unifying parameter and limited information in the system, such as EGBV for FDI merits investigation. The method's development is explained in subsequent sections of this chapter.

2.1.2 Energy-based methods and development of EGBV for FDI

EGBV for FDI is an umbrella term for many different methodologies of applying an energy characterisation to a complex system where many forms of energy transformations occur. The EGBV method was formally introduced in an articles, first by Van Schoor et al. [5] and further elaborated on by Uren et al. [4] and is the fundamental framework of methodologies in many studies that employ the energy-based method.

Initially introduced as energy-based visualisation (EBV), the method adopted the key principles of graph theory to be known as EGBV. A thorough understanding of the EGBV method starts with the study of the EBV literature.

An overview of the development of EGBV through application in various systems is given by Uren et al. [9] and summarised in a table format. As shown in Table 2.2

Table 2.2: Applications of the EGBV method adapted from [9]

Reference	Application	Information classification
[21], [22], [1], [23]	High-temperature gas reactor (HTGR) (closed-loop Brayton thermodynamic cycle)	Flownex [®] model (dynamic)
[24]	Continuous stirred tank reactor (CSTR)	Analytical model (steady state)
[25], [26]	Counter-flow heat exchanger	MATLAB [®] (steady state)
[27]	Axial-flow compressor system	Flownex [®] (steady state)
[28], [29]	Autothermal reformer	HYSYS [®] (steady state)
[30]	Subcritical pulverised coal-fired boiler	EES [®] (steady state)
[31], [32]	Transcritical CO ₂ Heat Pump	EES [®] (steady state)
[33], [13], [34]	Gas to liquids (GTL) process	HYSYS [®] (steady state)
[4]	Heated two-tank system	MATLAB [®] Simulink [®] (dynamic with noise) with practical data
[35]	Tennessee Eastman process	MATLAB [®] Simulink [®] (dynamic with noise)
[3]	Lignite plant	Practical data with Epsilon model (steady state with noise)

The EBV finds its roots in early applications of what is known as the entropy-entropy method. The enthalpy-entropy method is introduced by Du Rand [22, 1] in an investigation of an energy-based graph method for condition monitoring in a nuclear-powered closed Brayton cycle. This method utilises measurements calculates enthalpy and entropy from measurements at each point in the system, reducing the measurement space to a single energy domain. These measurements are visualised using an enthalpy-entropy graph (h-s graph) shown in Figure 2.1. This is done for each operating condition, referred to as a signature. By determining the difference between signatures, Du Rand was able to determine the location and magnitude of faults that occur in the system, deeming this method viable for condition monitoring of this system.

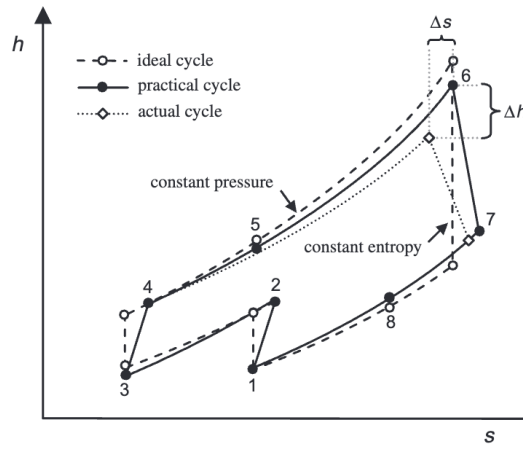


Figure 2.1: Illustration of the h-s graph FD approach from [1]

Further research by Van Schoor et al. [36] combines the work of Du Rand with the concept of representing systems as equivalent electrical circuits, which is described as analogous to conventional control theory. This work proposes a well-founded energy-based framework to use energy as a unifying domain to work on, analyse, and optimise large-scale industrial systems [36]. However, Van Schoor et al. [36] states that further investigation is required into the abstraction of relevant energy formalisms and subsequent transformations to energy signatures.

In 2016, Fouché et al. [27] proposed an energy visualisation approach for FDI within an axial-flow compressor system. The methodology involves an investigation of steady-state and transient behaviours of the system to generate residuals for FD. A nonlinear model capable of representing various compressor instabilities is developed. The methodology includes a power-energy analysis using the system's state variables, both under normal operating conditions and fault conditions. These analyses are presented visually, and the results show that the power and energy characteristics of the system can be useful for fault detection and diagnosis (FDD). Fouché et al. [27], similar to Van Schoor et al. [36], also utilises energy a unifying domain for the measurement space. However, the system is represented in terms of subsystems, each identified by a node, and power interactions between nodes are represented as links. A matrix representation is then configured from

nodes and links, where attributes of steady-state energy values are assigned to each node, and attributes of steady-state power values are assigned to each link. From the system node diagram, a matrix representation is given for power transfer, and a vector representation is given for the energy of each node. A method of visualisation implemented for the power and energy of the system is an energy vs. power graph. A power matrix and energy vector are defined for normal operating conditions and the fault conditions are compared to the normal operating conditions by residual calculation of the energy-power graph between normal and fault operations. System instability can be differentiated from faults within the system and FDI performance is deemed promising.

In addition, Marais et al. [28] investigates the application of a novel energy-based method that is adapted from the method introduced by Du Rand [22, 1] to petrochemical plants. This method referred to as the extended enthalpy-entropy method is adapted for an open-loop process because the initial method would not be viable for application on an open-loop system. The extended enthalpy-entropy method for FDI is applied on an autothermal reformer of a gas-to-liquids process. For the faults considered, the technique holds for FD; however, this work states that the technique breaks down in terms of fault isolation. Multiple reasons for the limitations of fault isolation are discussed, with a key emphasis on the nature of entropy and enthalpy that does not adequately represent faults characterised by changes in chemical composition. Marais states that the enthalpy-entropy method's reductive properties in terms of reducing the measurement space are highly advantageous and that the method's use of energy as a unifying domain holds significant value in terms of FDI. Marais recommends further investigation of whether the method can be extended for better fault isolation performance in petrochemical processes.

Uren et al. [25] utilises EBV on a counter-flow heat exchanger described as a residual model-based method of FDD. In this study a state space model is produced and faults within the system are defined. The method is similar to that applied by Fouché et al. [27]. The counter-flow heat exchanger is characterised by steady-state and transient operation and each method is evaluated for FDI using residuals. Steady-state structural residuals were able to uniquely determine fault positions and magnitudes and could identify multi-faults, whereas transient structured residuals could not uniquely classify the different faults. The energy-based approach to residual generation is deemed viable for providing rich information for FDI.

In 2017, Van Graan [37] introduces formal graph theory with EBV for FD on a counterflow heat exchanger. In essence, this work is a formal introduction to graph theory in practice similar to the method utilised by Fouché et al. [27]. That is, to split a system into subcomponents that resemble a directional graph, Van Graan [37] formally using linear graph theory with the concept of using an attributed graph for graph matching as a means of FDI. The method is derived from Jouili et al. [38], where a system, in this case, a counter-flow heat exchanger, is represented as a linear attributed graph.

The attributed graph combines the structure of the system in the form of energy interac-

tions between components, with the energy data obtained from a simulated model, referred to as energy characterisation. Van Graan [37] also uses energy as a unifying parameter; however, in this case, the attribute of choice is the flow of energy $\dot{q}$ and for components or nodes x . The nodes and links for the attributed graph are chosen with the availability of data and data redundancy in mind, and the attributed graph is determined to have six nodes and eight edges. When nodes and links have been defined and the exergy flow and energy flow attributes have been determined from the energy characterisation, a matrix representation of the graph attributed for normal operation is determined, called the original node signature N_O . This is done with the method proposed by Jouili et al. [38].

The faults under investigation are simulated in the model, and a node signature is also determined for each fault N_F . The normal and fault node signatures are then compared using graph matching, employing a comparison function to determine the similarity between two matrices. The function used is the heterogeneous Euclidean overlap metric or HEOM. Applying the HEOM function to two matrices generates a cost matrix which is analysed in terms of eigendecomposition to determine the eigenvalues of the cost matrix that may be unique to differentiate between normal and fault behaviour and even be indicative of the type and magnitude of the fault. A distance measure called the Hungarian algorithm is also employed to determine a DP from the cost matrices. Analysis of the eigenvalues and DP for normal and fault cases is done and the results show that the graph-matching approach to the energy visualisation of a counterflow heat exchanger is successful for FD and could uniquely identify if heat leakage, a fluid leak, or fault had occurred within the component.

2.1.3 Research after the establishment of EGBV for FDI

After the introduction of graph theory into energy-based FDI, the method as currently utilized is formally introduced in an article by Van Schoor and Uren [5], published in 2028, as a vision for EBV for FDI. This method has been further researched and developed in terms of formalising notation, application to different systems, node and link criteria, methods of utilising node signature matrices for FDI, and mainly the use of the rate of energy transfer as link attributes and the use of the rate of exergy change as node attributes. From the formal introduction, this method has been applied to different systems for the purpose of FDI.

Studies of applying EGBV for FDI in systems include a trans-critical CO₂ heat pump cycle by De Bruin [14] and Uren et al. [39], the Brayton cycle power conversion unit by Naser et al. [15], the gas-to-liquid process by Greyling et al. [33], an industrial STS by Smith [3], the Tennessee Eastman process by Steyger et al. [40] and Vosloo et al. [35] and a heated two-tank system by Wolmarans et al. [2, 41].

The most recent literature shows promise for the application of EGBV as a method of energy characterisation for FDI applications [3, 15, 5]. The methods stand up compared

to what is regarded as a benchmark fault detection scheme known as principle component analysis (PCA) in terms of FDI performance [2]. Performance in terms of detection, isolation, sensitivity, and robustness has been evaluated and has garnered many positive results.

Studies have shown that varying the energy characterisation analysis method for FDI can greatly impact the results with the DP, the residual method, the eigendecomposition, and spectral analysis, showing promise in terms of FDI performance [15, 42].

The latest literature on the method applied to a system is based on the application to a heated two-tank system by Wolmarans [2] and a STS by Smith [3]. Compared to one another, the STS is considered to be more complex than the heated two-tank system and perhaps the most complex system investigated yet. The STS has more components, the components are more complex, the system is larger scale, and the system is not an experimental plant. Thus, data acquisition is limited to the sensors the plant is controlled with, the plant is in operation, and access is limited. On the other hand, the heated two-tank system is a system specifically equipped with the necessary sensors for data acquisition, lower component complexity, and ease of access allowing quality research to be done with minimum hurdles to jump over [2].

From these studies, it is evident that the research done by Smith [3] provides a framework for how to apply the EGBV method to a system in terms of graph setup, and the research done by Wolmarans [2] gives a framework for a possible FDI scheme otherwise known as EGBV analysis to complement the data generated from the EGBV application. There may be merit in applying the current EGBV method for FDI on an experimental system of high complexity with the necessary information available, access to the plant outside of operational periods, and a higher degree of data acquisition instrumentation than in the case of the STS. A valuable system proposed for such an application is a thermal storage test facility known as THERESA.

Some studies have also indicated that the variation of the method applied to the characterisation of the energy has an impact on whether the operating point changes and influences the energy signatures produced [15].

One problem area in FDI schemes is the introduction of multiple faults and the instability of these individual faults. This, however, is a common challenge in most FDI schemes.

The development of the EGBV method shows that it is key to evaluate the most recent developments of the methods to develop a method that is up to date with the necessary advancements. Identifying case studies where the method has been applied to systems will provide context to how the method is practically applied to a large-scale industrial system.

2.1.4 Case studies

Case study 1: A comparative study between EGBV and PCA, by Wolmarans [2].

This study compares PCA and EGBV in terms of FDI performance in a heated two-tank process. The heated two-tank process is simple enough to practically implement as a pilot plant, yet also provides enough complexity to study the effects of interdependent process variables when faults occur [2]. A illustration of the heated two tank system is shown in Figure 2.2

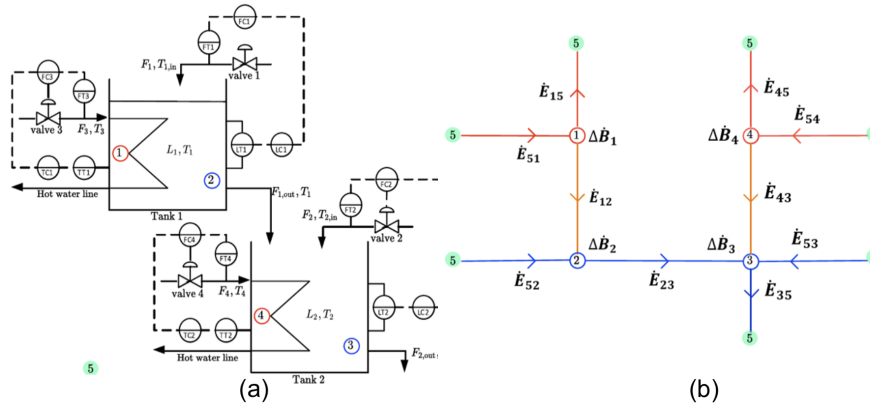


Figure 2.2: (a) Process diagram of the heated two tank system. (b) Graph representation of the heated two tank system. From [2].

The aim of the study is to investigate and verify the results of the comparison of the EGBV method for FDI with principle component analysis (PCA). For this comparison, the performance parameters utilised are sensitivity, robustness, and promptness. The results show that the EGBV method for FDI is more sensitive and more prompt, whereas the PCA method was more robust. EGBV for FDI had better FD rates and better isolatability. EGBV was found to have more ideal attributes of FDI than PCA.

Practical data of system operation are used to validate a process model set up in MATLAB[®] Simulink[®], which was done using mass balance and energy balance. The configuration of the model is based on the process parameters, taking into account control of the system, noise, and disturbance. Furthermore, valve modelling is performed to determine the valve characteristics for accurate process model derivation. The model is validated to verify all FDI results through results obtained from normal and faulty operation. The study determined that the practical process is functioning at suboptimal performance as a result of oversized valves and external control utility. These changes cause unintended changes in operating conditions that are believed to have an impact on the final FDI results. However, based on the assumption that not all industrial processes are optimally configured, the system and model are deemed to be sufficient for FDI testing.

The methodology can be briefly described starting with the measurement space that in-

cludes pressures, mass flows, and temperatures. This space is transformed into the feature space using graph theory. Within the feature space, exergy and energy flow attributes are represented on a graph. The feature space is transformed into the decision space through matrix representation of the attributed graph and graph comparison methods. Within the decision space are node signature matrices and cost matrices. These are transformed to the class space using SVD. Within the class space are SVD and graph matching.

Case study 2: EGBV for FDI in a steam turbine system, by Smith [3].

In this research Smith [3] applies EGBV to an industrial STS which can be considered a complex thermal system. The method followed by Smith is perhaps a better representation of the application of EGBV for FDI in a large-scale industrial system, where the components are much larger in scale and quantity increasing system complexity significantly as illustrated in Figure 2.3.

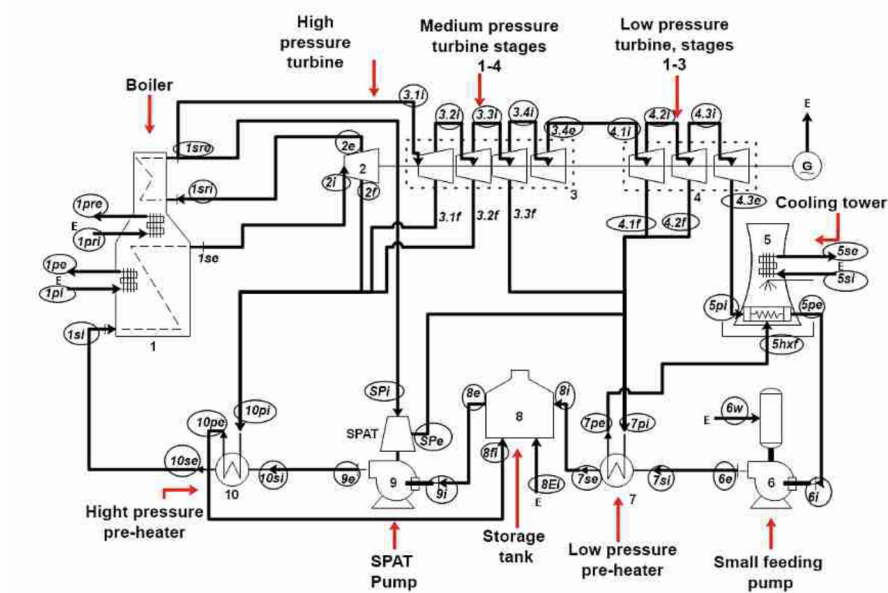


Figure 2.3: Steam turbine system simplified system schematic representation from [3].

In this case, the challenges associated with a process model configuration are highlighted, as well as the challenge of the availability of data within such a system. With this, the complexity of the system and the large amounts of components require the determination of node and link criteria for the setup of the system graph to ensure that the system graph represents the system well and does not lead to exhaustive data loads.

The analysis method used after the attributed graph and node signature matrices have been constructed is three variations of the FDI analysis to determine differences in performance characteristics. These methods are the DP approach (approach 1), the eigendecomposition approach (approach 2), and the residual approach (approach 3).

The eigendecomposition approach is based on the properties of eigenvalues and their rela-

tionship with the diagonal of the cost matrix. In this approach, two matrices are compared forming the same cost matrix as in the DP, however, from this the eigenvalues of the cost matrix are calculated and utilised qualitatively to determine the FDI performance.

The SBS SVD method had the best performance. The detection accuracy of approach 2 was 100% and the isolation accuracy ranged from 19%, for some variations of FT2, to 100% for FT4. Approach 3 had the worst detection accuracy with moderate isolation, attaining only 72% accuracy in terms of FD and an isolation accuracy ranging between 5.2% and 70%. From these methods, it is determined that the eigendecomposition of the cost matrix has the best performance compared to the DP approach that is also used by Van Graan [26] or the residual-based method.

Case Study 3: Graph-based fault detection for a gas-to-liquids process by Greyling et al. [42]

Greyling uses two energy-based methods for FDI in a gas-to-liquid process: fixed threshold exergy-based method and graph matching approach using DP and graph matching approach using eigendecomposition. Using metrics of detection, isolation, and fault pre-definition ability, the methods were evaluated, and it was found that graph-matching using the DP yielded superior results as faults could be isolated.

Case Study 4: EBV of a Brayton cycle power conversion unit by Nesar et al. [15]

The Brayton cycle power conversion unit is used to determine the suitability of EBV for FDI. Two methods are used to create energy signatures, namely, the eigendecomposition-based method and the residual method. Both methods generated unique signatures that were successfully used for FDI. Based on the results obtained, both methods could detect multiple faults but were unable to isolate individual faults. The residual method generated a unique fault signature, regardless of the magnitude of the fault. The residual method could detect changes in the operational set-point. The eigendecomposition-based method could not differentiate between different set points and was not as sensitive as the residual-based method. The study concludes that EBV on a Brayton-cycle PCU is suitable for FDI.

Case Study 5: EBV of a trans-critical CO₂ heat pump cycle by De Bruin [14]

The study applies EGBV to a trans-critical CO₂ heat pump system and reviews the FDI performance of eigendecomposition. The results show that using only eigenvalue information is not sufficient in FDI; however, utilising eigenvector and eigenvalue information is suitable for FDI.

2.1.5 Critical evaluation of literature

A critical evaluation of the literature is the basis for the research question formulated in Chapter 1. The literature shows that the EGBV method for FDI is still in the development stages. The method has been rigorously applied and tested against benchmarks for FDI methods such as PCA and the method has held up [2]. In most cases, these applications were done on systems where either an experimental plant/test bench was available or a model/partial model of the system was available. The largest and possibly the most complex system to which the method has been applied is the Tennessee Eastman process, or the industrial STS [35, 40, 3].

Reviewing case studies reveals that there are challenges that may arise when applying EGBV for FDI on a large complex system. In many cases, a practical and simulated system is utilised to overcome the limitations in terms of process variable measurements at many points of interest. In this case, the simulated model can be utilised to model such sensor data.

In most of the studies reviewed, the study evaluated the EGBV method for FDI by investigating the results of the FDI performance. In these cases, the method is declared viable if the FDI performance metrics give adequate results. Thus, it can be the case that the method can be applied to a system and energy characterisation can be evaluated through the sensitivity of energy characterisation. From the literature it is clear that the results of different studies on EGBV for FDI are conflicting with the results achieved, and it can be assumed that the method used for FDI could be system-dependent, meaning that different methods perform well for different systems.

Applying the identical concept of EGBV and using various FDI methods Smith [3] found that eigendecomposition of the cost matrix showed the best results, compared to the residual method and the eigendecomposition method. Greyling et al. [42] found that between the eigendecomposition and the DP, the DP was the superior FDI method, while Naser et al. [15] found that between the residual method and the eigendecomposition method the residual method was more sensitive and was able to detect changes in the operational set point.

Hence, the results obtained throughout multiple studies have not consistently highlighted a clear FDI method paired with the same EGBV method of energy characterisation. The question that arises from this research is the following: What if there is more to understand about the energy characterisation of EGBV and that insights from studying the characteristics of the energy characterising can give insights into the interpretation of an FDI search method performance?

The fundamental principle of EGBV is based on energy as a unifying domain to characterise a system, and the results from previous studies utilising the same core principle have received conflicting results. Perhaps, it is necessary to go deeper into the understanding of characteristics of EGBV as an energy characterisation tool before further study is done on

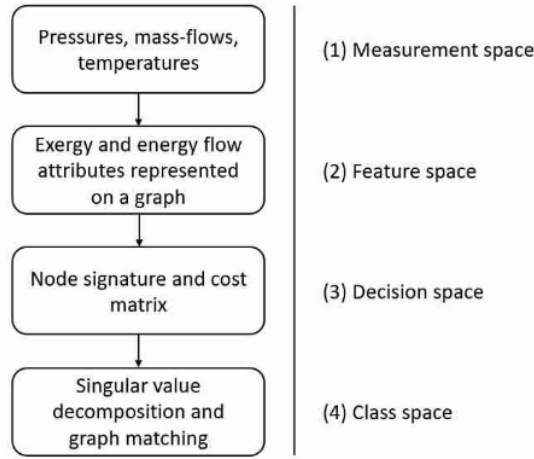


Figure 2.4: Stages of fault signature development in EGBV for FDI from [4]

the performance of FDI applied to the energy characterisation as this could give valuable insights on how the FDI results are interpreted or produced.

2.2 Literature study

The literature survey conducted in the first section of Chapter 2 gives an in-depth discussion on FDI and the development of the hybrid FDI method EGBV for FDI. The literature study to follow in this section discuss the application of EGBV for FDI in a system, with a specific focus on the application of the method to a large-scale thermal power plant with heat storage capabilities as introduced in Chapter 1.

2.2.1 EGBV by diagnostic system spaces

As discussed in Section 2.1 EGBV has been applied to several systems and components over various studies; however, the method is not generalised across all studies when referring to the application or further analysis, thus the method used in this study is devised using studies based on systems of similar complexity. As the most relevant literature on the application of EGBV to a large complex system the research of Smith [3] will be utilised, providing insight into the challenges for the application of EGBV to a complex system and the basic framework to follow. Other sources used are relevant in application or analysis also, namely, core sources for EGBV, which was utilised by Smith [3] and Wolmarans [2].

The research carried out on EGBV for FDI as a diagnostic system has applied the methodology described by Venkatasubramanian et al. [12], with processes within the EGBV method assigned to each of these spaces, as seen in Figure 2.4. The deliverables of each space are also seen as used by Wolmarans [2] and Uren [4].

The measurement space is the first space in the transformation spaces within the diagnostic

system. The measurement space contains raw process measurement data recorded from the system for the duration of defined experiments. The raw data are captured from sensors measuring pressure, temperature, flow rates, fluid level, electrical power, and valve control logic. In the measurement space, information is gathered and equations are set up so that the measurement space can be transformed to the feature space. In the measurement space, the energy and exergy balance equations and the mass balance equations are determined to be used for the reduction of the dimensionality in the feature space [2].

The feature space results from the transformation applied to the measurement space which involves the use of the *a priori* process knowledge and feature extraction or feature selection [12]. According to Venkatasubramanian et al. [12] the transformation from the measurement space to the feature space leads to a reduction in dimensionality, in the case of the EGBV method, this reduction in dimensionality is achieved by using energy and exergy as a unifying parameter for the physical domains that the different measurement data of the process occupy. Also, dimensionality is reduced by feature selection, namely, selecting only certain process measurements to wholly describe certain energy interactions in the system. In both of these operations, *a priori* knowledge is used to determine the system graph and to calculate the energy and exergy flow.

The decision space involves utilising features such as energy and exergy attributes, in combination with components of graph theory to ultimately construct a node signature matrix. In the decision space, the system-attributed graph is constructed which can be seen as the graphical representation of the matrix form which is the node signature matrix. Within a single data set representing an operational period of the system, each node signature matrix is a representative snapshot of the energy and exergy information of the system at a given time.

In the context of EGBV for FDI, this can be regarded as the EGBV output before further analysis for FDI. Further operations will involve applying some form of graph matching to form cost matrices and class space analysis methods to determine the FDI performance characteristics, called EGBV analysis. Thus, in the context of the diagnostic spaces, the class space can be seen as the FDI component in EGBV for FDI.

2.2.2 System measurement data and information

The available process data are a key driver of whether it is viable to apply a hybrid method such as EGBV for FDI to a practical system. With a complex large-scale system, a validated model has proven to be beneficial in the application of the method, as measurements at various points are available, faults can be simulated within the system, and experiments can be devised to best illustrate the goal of the research.

Some studies also found that the necessary data are available to apply the method without access to a process model; however, certain limiting factors have led to a simulated system being necessary to conduct an experimental design. These cases are where a system cannot

be experimented with and faulty operational data are not available for the system.

The research question defined in Chapter 1 of this study does not intend to produce a process model. Instead, the method is applied without access to a process model.

2.2.3 System graph setup

The system graph in the context of EGBV includes the use of graph theory to define the theoretical definition of a graph. In addition, the system graph follows the definition of a graph to configure a system graph for a specified system which will be specific to the system structure and components, where nodes represent components and links represent energy interactions between components.

Graph theory

A simple graph G is defined as an ordered pair $G(N, L)$, where $N(G)$ is a finite set of vertices, called nodes, and $L(G)$ is a set of edges or links [43, 4]. A node denoted by n_i for $i = 1$ to k where k is the total number of nodes. A link $l_{a,b}$ is said to join the nodes a and b . This simple graph is also called the system graph.

The system-attributed graph goes further to define the weights to the edges of the system graph. For the system-attributed graph, let $G_A = (N, L, A)$ where N is a finite non-empty set of nodes or vertices, L is a finite set of links or edges, and A is the attribute set [4]. The set of attributes can be defined as $A = \{a_{n_i}, a_{l_{i,j}}\}$ where a_{n_i} is the attribute of n_i , $a_{l_{i,j}}$ are the attribute of the link incident to n_i .

From the theoretical definition of a system graph, it is necessary to determine nodes and links that connect nodes. There are multiple approaches for the configuration of a system graph and an energy-attributed graph. The system components should represent nodes and the links should represent a connection between the system components [5]. With a complex system with many components and flow paths, the determination of nodes and links should be thoroughly considered. With the ultimate goal of developing a system-attributed graph and the matrix representation of that attributed graph, there can be different variations for such a graph developed from a specific system.

Figure 2.5 shows the transformation of a system from the schematic representation of the system to a representation of the system attributed graph, which is a system graph with attributed energy and exergy to links and nodes respectively. Energy interactions between the system and the environment is also represented as links, these can be seen in Figure 2.5 as interactions between various components and the node representing the environment as node 8.

Research done on the development of system-attributed graphs and graph reduction tech-

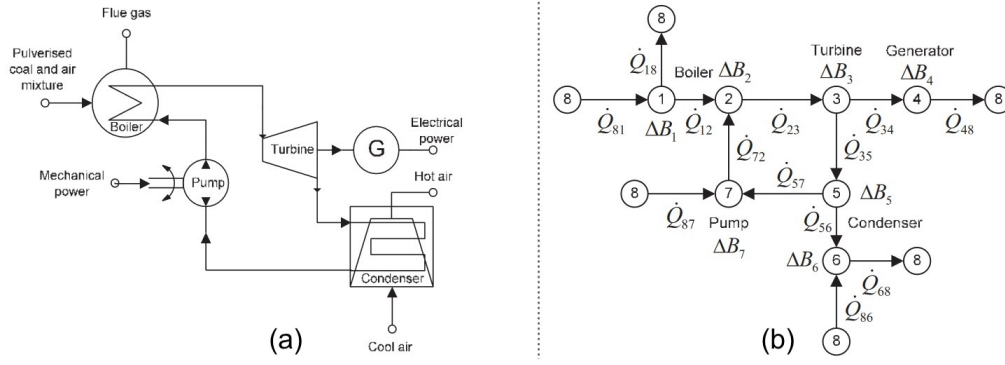


Figure 2.5: (a) simplified representation of a coal-fired power station. (b) Attributed graph of a coal-fired power station. From [5]

niques has been investigated by Steyger [40], showing that different reduction techniques may have an impact on the overall performance of FDI techniques applied to said graphs.

When a system is considered, all possible nodes are selected, and all possible links between nodes are defined, the system graph will be highly complex; also, the amount of data necessary for attributes will be of a high order. This will lead to a complex graph that will require high computational power to utilise. As a framework of how a system is evaluated to determine nodes, links, and attributes, it can be implicitly and explicitly derived from previous research. Thus, it is important to devise a system graph that represents the key nodes and links that facilitate efficient calculations, and that is still representative of the system.

The system graph setup is an iterative process in which the system links and nodes can be configured based on the structure of the system, the components of the system, the criteria of the nodes and links of the system, the operational information of the system, and the data available for the defined nodes and links.

The first step in this process is identifying all components that form part of the system. These components may include: Valves, tanks, heat exchangers, inlets and outlets of fluid or power, pumps, and other components considered part of the system. Also part of this step is identifying the system boundary and components interfacing with the environment. The inlets and outlets of the fluid or the power supplied to a component are determined to be environmental nodes and define the boundary of the system.

To reduce the graph, specific component types such as valves can be neglected as nodes to reduce graph complexity and enhance computational efficiency in further operations.

Only nodes of high priority are chosen; this is based on a priori knowledge from system operators.

Node criteria

Thorough criteria for node selection are discussed by Smith [3] and can be adopted as a guideline for choosing nodes for a large-scale complex system. The criteria include five rules:

- Where should faults be detected?
- Which components have many inflows or outflows in terms of energy?
- Which components will affect the system if a small change in the energy occurs at/in these components?
- At which components do large amounts of energy flow in/out?
- How can the energy characterisation be done to compose the least complex graph possible?

These criteria can be utilised, along with the system information and process knowledge, to iteratively determine the system-attributed graph.

Specific node criteria can be established on the basis of the system and its components. An important part of determining the nodes is to determine a system boundary.

The boundary of the system can be irregular and adaptable through different iterations of the system graph. This is where the system interacts with the environment. One such example is the electrical power delivered to a pump. This power delivered can be described as a link between two nodes, the pump is represented as a node, and the environment can be represented as a node. Thus, the environmental node is defined as the interaction of the system with the environment. The environmental node can be represented by multiple nodes, with the same subscript (the highest numerical subscript of the system nodes), to illustrate the interaction with the environment across the defined system boundary. The links to the environmental node are defined in the same way as the links in the rest of the system.

Valves within a system do not significantly change energy or exergy values within a system and are considered to be a control component, but valves are not deemed essential points of FD. One can argue that if two nodes are defined (that are not valves) with a valve controlling a flow stream between these two nodes, the valve will inherently be included within the graph but will form part of a link. This reasoning allows the system graph to be reduced in terms of the number of nodes, as these components will not be considered nodes.

Pumps with power input should be represented as nodes. These pumps contribute to the system in terms of large energy flows and are critical points for FD.

Boilers also receive power and transfer the power to the flow stream as heat. These components significantly contribute to the energy within a system and are critical components for FD.

Heat exchangers transfer heat between two fluid streams; they may have a large number of inflows and outflows and a large amount of energy transfer; also, heat exchangers are prone to faults like fouling and are critical for FD. Heat exchangers can be represented as two nodes. One for the primary side and one for the secondary side.

Vessels or accumulators may receive multiple inflows and may distribute flows to multiple points within the system. Such components with few parts that may fail and the high energy flow rates in and out of a vessel constitute a node by the defined criteria.

In some cases, it might seem rational to include all components as nodes, but in later operations where the energy attributes are associated with nodes and links, problems may arise in the availability of data or computational cost with diminishing returns in terms of the performance of the method. Thus, node selection should be iterative and specific.

Special cases of determining a node can be in the case where a three-way intersection of a pipeline occurs. This point can be termed a temporary node.

Link criteria

In the configuration of the system, graph links can be determined by analysing flow paths of energy to and from the selected nodes. In any case, where two components interact in terms of energy flow, a link can be defined. To determine links between nodes, *a priori* process knowledge (or expert knowledge about system process operation) is also implemented.

Links connect nodes and are an indication of the energy flow within the system. Links can indicate the flow of fluid, power, heat, or other forms of energy.

Based on the operational information of the system, it is necessary not only to determine which flow paths are used during operation but also if the practical system has the necessary data available to determine the energy flow of these flow paths. In systems where a process model is available, the determination of links can be simplified, as the system model should be able to give reliable approximations to values where sensors are not available in the practical system.

If paths with insufficient information are determined and missing information cannot be calculated, nodes should be reevaluated by merging nodes into a single node, possibly excluding nodes, or by merging nodes to the environmental node.

2.2.4 Energy and exergy analysis

Basic mass balance equations can be used within the measurement space for calculations. These mass balance equations are based on the assumptions that at each point steady-state flow occurs, a non-deformable control volume is defined, and height and friction effects are negligible [44, 45]. The resultant mass balance equations are given as:

$$\sum_i (\rho_i A_i V_i)_{in} = \sum_i (\rho_i A_i V_i)_{out}, \quad (2.1)$$

which reduces to the commonly used equation in engineering analysis [45]:

$$\sum_i (\dot{m}_i)_{in} = \sum_i (\dot{m}_i)_{out}. \quad (2.2)$$

Mass balance equations can be used to determine mass flow rates at necessary positions, the mass balance equations can be adapted for components as discussed further in this subsection.

In the analysis of a thermal system, the concept of exergy analysis is used to describe the quality of energy and the degradation of energy within a system [46, 47]. Energy is defined as the ability to do work; exergy can be defined as the amount of energy available to do work, compared to a reference state. The reference state or dead state is defined as the environmental state at temperature T_0 and pressure P_0 . Exergy analysis uses the principles of energy and energy conservation, combined with the second law of thermodynamics, for the analysis of energy systems [47]. Exergy analysis can be used as a method of evaluating the efficiency of energy resource use, as it allows the magnitudes, types, and locations of energy wastes to be determined [47].

The dead-state conditions are integral in the energy and exergy analysis and can be chosen as the ambient temperature and pressure, as this is a realistic representation of the temperature of the environment to which the system can be brought into equilibrium, without the expenditure of work [46]. The reference dead state can be defined as absolute zero, or as the ambient temperature and pressure that the system is experiencing. To ensure consistency with previous applications of the method, ambient temperature and pressure are used as experienced by the system under investigation [3].

The exergy associated with a stream of matter can be divided into components, which include kinetic exergy, potential exergy, physical exergy, and chemical exergy. A system with mainly thermal energy transfer, characterising the system energy interactions, physical energy, and exergy is considered. The physical exergy component is defined as the maximum amount of energy available from a fluid stream when it is brought from its initial state to the dead state by processes that involve only thermal interaction with the environment [46].

In studies where system operation is not significantly influenced by chemical changes, only physical exergy is considered. In large-scale thermal power plants, physical exergy predom-

inantly characterises the energy associated with heat transfer and electrical power generation. When employing physical exergy in analysis, the following assumptions are made: the process is considered to be at steady state, meaning that thermal properties remain constant over time at any given point in the system. Additionally, while heat exchange with the surroundings is inevitable in real-world applications, certain components—such as well-insulated piping, turbines operating under ideal conditions, and selected heat exchangers—can be reasonably approximated as adiabatic. This assumption simplifies the thermodynamic analysis by neglecting minor heat losses to the environment, allowing for a focused evaluation of system performance.

To demonstrate the use of energy and exergy analysis within the EGBV method, a generic node can be defined, the node being considered the control volume. Assumptions in thermodynamic analysis include that the control volume is a finite non-deforming control volume. Figure 2.6 shows an energy balance across a generic node. From the first law of

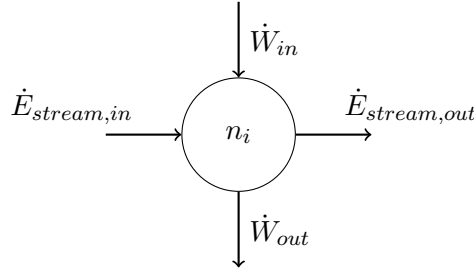


Figure 2.6: Energy balance across a generic node

thermodynamics, the rate of energy flow into and out of a control volume can be given shown in Figure 2.6:

$$\sum_i (\dot{E}_{i,in} + \dot{W}_{i,in}) = \sum_i (\dot{E}_{i,out} + \dot{W}_{i,out}). \quad (2.3)$$

The rate of energy flow into the defined control volume is equal to the rate of energy flow out.

The rate of energy transfer by a fluid stream ($\dot{E}_{stream}$) is calculated from the fluid mass flow ($\dot{m}_{stream}$) and the specific enthalpy (h_{stream}) of the fluid stream. For links of electrical power supplied to the components, $\dot{W}_{stream}$ is equal to the power supplied to the component.

In terms of exergy analysis of the same generic component represented as a node and under the same assumptions as for the energy analysis, Figure 2.7 shows the node considered where B is used to denote exergy.

Within the EGBV method exergy analysis is used to determine a change in exergy or $\Delta \dot{B}$ across a control volume, discussed by Van Schoor et al. [5]. The exergy analysis can be represented mathematically as:

$$\Delta \dot{B} = \Delta \dot{B}_p + \Delta \dot{B}_c = \sum \dot{W} + \dot{B}_{destroyed}. \quad (2.4)$$

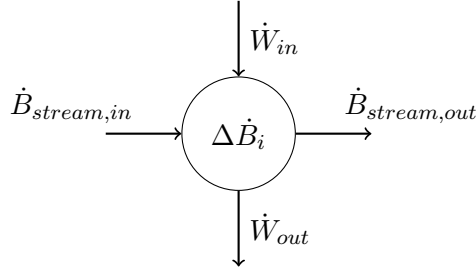


Figure 2.7: Exergy balance across a generic node

If chemical exergy is regarded as negligible and only physical exergy is considered then,

$$\Delta \dot{B} = \Delta \dot{B}_p. \quad (2.5)$$

In (2.4), taking into account specific entropy s , the term for exergy destroyed is given as,

$$\dot{B}_{destroyed} = -\dot{m}T_0(s_{in} - s_{out}). \quad (2.6)$$

The term $\sum \dot{W}$ includes all forms of energy transfer rates throughout the control volume considered; thus the term can be expanded to $\sum \dot{W} = \sum \dot{E}_{stream} + \sum \dot{W}_{electrical}$. Electrical power in a component is considered to be lossless transfer, and all work is done in the fluid stream.

Using (2.4), the exergy flow rate into the generic component is calculated as:

$$\dot{B}_{in} = \dot{m}_{stream}(h_{stream,in} - h_0) - \dot{m}_{stream}T_0(s_{stream,in} - s_0) + \dot{W}_{in}, \quad (2.7)$$

where $\dot{m}_{stream}$ is the mass flow rate of the stream, h_{stream} is the specific enthalpy of the stream, s_{stream} is the specific entropy of the stream, T_0 is defined as the reference temperature, h_0 is the reference enthalpy, and s_0 is the reference entropy. The same equation can be rewritten for the fluid flow stream out of the component, where the component does not perform the work:

$$\dot{B}_{out} = \dot{m}_{stream}(h_{stream,out} - h_0) - \dot{m}_{stream}T_0(s_{stream,out} - s_0). \quad (2.8)$$

Thus, the equation for describing the change in exergy can be written as:

$$\Delta \dot{B} = \dot{m}(h_{in} - h_{out}) - \dot{m}T_0(s_{in} - s_{out}) + \dot{W}_{electrical}. \quad (2.9)$$

In cases where heat transfer occurs the following equation is utilised to calculate the exergy of such transfer [48]:

$$\dot{B}_{hx} = (1 - \frac{T_0}{T})(\dot{E}_{hx}). \quad (2.10)$$

Where T is the temperature at which the heat transfer takes place in the case of heat exchangers this can be rewritten as [49]:

$$\dot{B}_{hx} = T_0 \left(\frac{1}{T_{source,in}} - \frac{1}{T_{source,out}} \right) (\dot{E}_{hx}). \quad (2.11)$$

Pumps

Pumps are considered as single nodes and it is relevant to obtain the power transferred to the pump as a node as this will ensure that the pump is well described as a node interacting with the environment, this can be described as:

$$\dot{W}_{in} = \dot{m}(h_{in} - h_{out}), \quad (2.12)$$

where heat addition is considered negligible.

Heat exchangers

The rate of energy transfer for a heat exchanger is given by,

$$\dot{E}_{transfer} = \dot{m}_h(h_{hi} - h_{ho}) - \dot{m}_c(h_{ci} - h_{co}), \quad (2.13)$$

where $\dot{E}$ is the heat transfer rate in watts, $\dot{m}_h$ and $\dot{m}_c$ are the mass flow rates of the fluid on the hot and cold sides respectively. Where h_{hi} , h_{ho} , h_{ci} , and h_{co} are the specific enthalpy of the fluid streams entering the hot side, exiting the hot side, entering the cold side, and exiting the cold side, respectively. This flow of energy is represented in Figure 2.8.

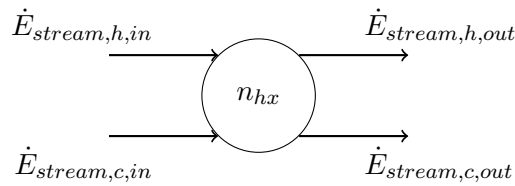


Figure 2.8: Heat exchanger represented as a single node

A shell-in-tube heat exchanger has two sides, a primary side and a secondary side. The equation given previously calculates the exchange of heat between these two sides, and by definition, a flow of energy can be represented as a link. The energy transfer between the primary and secondary sides of the heat exchanger has been represented as a link by defining the heat exchanger as two separate nodes, one for the primary side and one for the secondary side [5]. This is represented in Figure 2.9.

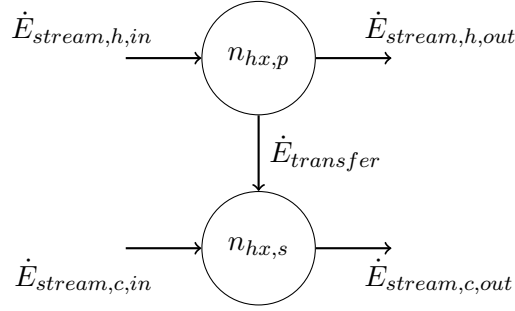


Figure 2.9: Heat exchanger represented as two nodes(primary side and secondary side

Tanks

In a water-steam system many components accumulate fluid and for accurate representation the determination of mass balance equations for vessels are required.

The mass balance equation for a vessel or tank is expressed as:

$$\dot{m}_{in} - \dot{m}_{out} = \frac{\rho A dh}{dt}, \quad (2.14)$$

where $\dot{m}_{in}$ is the mass flow rate into the tank, $\dot{m}_{out}$ is the mass flow rate out of the tank, ρ is the fluid density, A is the effective area of the tank, h is the height of the fluid in the tank (in meters), and t is time (in seconds).

Modelling the fluid flow out of tanks of different geometries and orientations is relatively straightforward if the tank remains full, however, it becomes increasingly complex to determine the volume of fluid within a tank if the fluid level within the tank is fluctuating and the tank is only partially full [50]. These equations are further developed in Subsection 4.4.4.

The energy balance across a tank that accumulates fluid is given as:

$$\dot{E}_{in} - \dot{E}_{out} - \dot{E}_{stored} = 0. \quad (2.15)$$

For dynamic operation where transient operation occurs, under steady-state conditions, this equation is simplified to:

$$\dot{E}_{in} - \dot{E}_{out} = 0. \quad (2.16)$$

The exergy equation for the transient case includes the term $\dot{E}_{stored}$:

$$\Delta \dot{B} = \dot{m}(h_{in} - h_{out}) - \dot{m}T_0(s_{in} - s_{out}) + \dot{E}_{stored} \quad (2.17)$$

2.2.5 Attributed graph and node signature matrix

An attribute matrix, also known as the node signature matrix, is derived from the system-attributed graph. It encodes the energy and exergy values associated with each node and link in the graph. The system-attributed graph follows from the system graph, which is evident in Figure 2.5 where the system structure is encoded by the system graph and the energy and exergy attributes are attributed to the system graph to form the system attributed graph.

Node attributes and link attributes

Attributes assigned to nodes in the system-attributed graph, referred to as node attributes are the rate of exergy change attributes across a node. This is the rate of exergy change, rate of exergy stored, or rate of exergy lost in a node. The notation for assigning node attributes follows that $\Delta\dot{B}_a$ indicates the change in exergy across node a .

Attributes assigned to links in the system-attributed graph, referred to as link attributes are energy flow rate attributes, where the power flowing between components is considered as well as the direction of the power flow. $\dot{E}_{(a,b)}$ where a denotes the node from which the energy flows and b denotes the node to which the energy flows.

Node signature matrix

The matrix representation of the system-attributed graph is shown in (2.18). The number of rows in $\mathbf{N}_s$ equals the number of nodes in the system-attributed graph (N) and the number of columns in $\mathbf{N}_s$ equals $N + 1$. Here, the row n describes all attributes associated with the n^{th} node. The first entry in a row describes the rate of exergy change ($\Delta\dot{B}$) across the node with the corresponding row number. In the case of row n , the entry $\mathbf{N}_s(n, 1) = \Delta\dot{B}_1$. where $\mathbf{N}_s(n, 2) = \dot{E}_{(n,1)}$, describes the link attribute from node n to node 1, then $\mathbf{N}_s(n, 3) = \dot{E}_{(n,2)}$ shows the rate of energy flow from node n to node 2. This is repeated until $\mathbf{N}_s(n, N + 1) = \dot{E}_{(n,N)}$.

$$\mathbf{N}_s = \left[\begin{array}{c|cccc} \Delta\dot{B}_1 & \dot{E}_{11} & \dot{E}_{12} & \dots & \dot{E}_{1N} \\ \Delta\dot{B}_2 & \dot{E}_{21} & \dot{E}_{22} & \dots & \dot{E}_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta\dot{B}_N & \dot{E}_{N1} & \dot{E}_{N2} & \dots & \dot{E}_{NN} \end{array} \right] \quad (2.18)$$

Criteria for the determination of the node signature matrix development can be deduced from Wolmarans [2] and Smith [3]:

- The diagonal of $\mathbf{N}_s$ will be assigned values of 0 as there is no energy interaction between a node and itself.

- In cases where there is no interaction described between nodes, a value of 0 will be attributed to an entry within the attributed graph.
- If a link attribute in the attributed graph describes the interaction between two nodes from node a to node b , the negative will be assigned to the inverse of that interaction, for example, $(\dot{E}_{a,b} = -\dot{E}_{b,a})$
- The exergy change for the environmental node will also be assigned 0 as this is the dead state.

2.2.6 Cost matrix by matrix comparison

Cost matrices are direct outputs from comparing two matrices by using a difference metric. In EGBV the choice of distance function implemented is the HEOM function, the input matrices are n by m , which is characteristic of the node signature matrices.

Heterogeneous Euclidean overlap metric

The HEOM function is a mathematical measure used to compare matrices, serving as a valuable tool in the assessment of similarity or dissimilarity between matrices [38, 51]. The HEOM function can be used to determine a cost matrix $\mathbf{C}_k \in \mathbb{R}^{n \times n}$, which contains the difference between elements in two matrices being compared [5, 3, 2].

Given matrices $\mathbf{N}_{s_n}$ and $\mathbf{N}_{s_f}$ with dimensions $n \times n + 1$ (where n is the number of rows and $n + 1$ is the number of features), the cost matrix entry between the i -th row of $\mathbf{N}_{s_{ref}}$ and the j -th row of $\mathbf{N}_{s,k}$ is given by:

$$c_{ij,k} = \sqrt{\sum_{a=1}^{n+1} \left(\frac{|\mathbf{N}_{s_{ref}}(i, a) - \mathbf{N}_{s,k}(j, a)|}{\{\text{range}\}_a} \right)^2} \quad (2.19)$$

where $\mathbf{N}_{s_{ref}}(i, a)$ and $\mathbf{N}_{s,k}(j, a)$ are the elements in the a -th column of the i -th and j -th rows respectively, and $\{\text{range}\}_a$ is the range (maximum-minimum) of the a -th column in $\mathbf{N}_{s_{ref}}$. If the range is 0, it is set to 1 to avoid division by zero.

Using the HEOM function graph matching of two matrices of $n \times n + 1$ yields an output cost matrix that is a square matrix of $n \times n$. These matrices contain the distances from the entries in the operational matrix to the reference.

Various ways of determining the reference node signature matrix exist and merit further investigation.

Reference node signature matrix

The node signature matrix time series set of normal data can be used to determine a reference node signature matrix. This represents a measure to compare operational data to extract data from subsequent operations to facilitate FDI.

Fixed reference

In systems that operate under steady-state conditions, an average of the normal data of the process can be used as the reference node signature for the set. Operational data will then be compared to this single matrix to determine from subsequent operations whether a fault was present in the system's operational data.

Dynamic reference

More recently the method of SBS analysis has been used to construct cost matrices. This method uses the process data and chooses each point as its reference, thus when compared, a cost matrix is produced from the same node signature matrix compared to itself. The advantage of this method is that within a dynamic process, the reference node signature matrix changes with each time step without making use of process knowledge about how and where to define a fixed reference node signature matrix.

2.2.7 EGBV analysis methods for FDI

As discussed in Chapter 2, many methods have been followed in the development of energy-based methods that have played a role in the development of the EGBV framework for FDI. Applicable EGBV analysis techniques are discussed. These methods represent the prominent methods that have been compared and evaluated in studies.

FDI search techniques include various techniques such as the implementation of moving average filters, threshold functions assessing AR, or qualitative analysis of trends and states for evaluating faults and FDI performance.

Residual method

With the residual approach, cost matrices are not constructed, the comparison between a normal reference node signature matrix and operational matrix is done through the use of a residual matrix. The residual matrix is obtained by evaluating the difference between the expected (reference) value and the actual value and normalising this by using the reference value as shown in (2.20).

$$\mathbf{G}_{res}(i, j) = \frac{g_{ref}(i, j) - g_{op}(i, j)}{g_{ref}(i, j)}, \quad (2.20)$$

with, $i \in \{1, 2, \dots, n\}$ and $j \in \{1, 2, \dots, m\}$ for all the graphs representing a $n \times m$ matrix. This method is outlined by Naser et al. [15], however, the data become exponentially more complex and computationally exhaustive as the nodes are increased to the point of 20 nodes such as in the work done by Smith [3].

In addition, this method relies on a process model to generate the expected normal reference data to be compared to operational data.

Distance parameter

The DP is a simple way of evaluating the distance between two matrices by utilising the average value of the diagonal entries of the cost matrix formed when two matrices are compared. This is a result of the cost matrix diagonal holding important information, which is reduced to a single value through summation of the diagonal matrix as per (2.21):

$$DP = \frac{\sum_{k=1}^n C_{kk}}{n}, \quad (2.21)$$

where kk refers to the relevant diagonal elements in the cost matrix. n represents the number of diagonal elements in the square cost matrix.

Eigen-decomposition

The eigen-decomposition method is a method that uses the characteristics of the cost matrix, namely the eigenvalues and eigenvectors, to determine the characteristics of the process through time. It finds extensive application in various mathematical, statistical, and computational contexts.

When applied to a square matrix $\mathbf{A}$ (where $\mathbf{A} \in \mathbb{R}^{n \times n}$ and n represent the dimension of the matrix), the eigendecomposition can be represented as

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}. \quad (2.22)$$

In (2.22), $\mathbf{V}$ is a matrix whose columns are the eigenvectors of, $\mathbf{\Lambda}$ is a diagonal matrix that contains the eigenvalues of $\mathbf{A}$ and $\mathbf{V}^{-1}$ denotes the inverse of the matrix $\mathbf{V}$.

The eigenvalues in $\mathbf{\Lambda}$ provide information on the scaling factor in the direction of their corresponding eigenvectors in $\mathbf{V}$. Eigendecomposition is particularly useful in understanding the characteristics of linear transformations, including their stability and behaviour under various operations. In the context of FDI methods, eigenvalues and eigenvectors can be employed to identify system characteristics and detect anomalies or faults in system

behaviour. This method is applied qualitatively and quantitatively in the literature on EGBV for FDI.

Singular value decomposition analysis

SVD is a crucial matrix factorisation method in linear algebra that is similar to eigendecomposition. When applied to a square matrix $\mathbf{A}$ where $\mathbf{A} \in \mathbb{R}^{n \times n}$ and n represent the dimension of the matrix, SVD can be represented as:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \quad (2.23)$$

with $\mathbf{U}$ the orthogonal matrix ($\mathbf{U}^T\mathbf{U} = \mathbf{I}$), $\mathbf{\Sigma}$ is a diagonal matrix containing the singular values of $\mathbf{A}$ in non-increasing order and $\mathbf{V}^T$ denotes the transposition of the orthogonal matrix $\mathbf{V}$.

The singular values in $\mathbf{\Sigma}$ are non-negative and are typically arranged in descending order. The singular values indicate the variability of the process and can be utilised for FDI search methods.

Within the SVD methods, a normalisation method is used and a threshold determination is used to detect anomalies.

2.3 Conclusion

Chapter 2 provides the foundational theoretical knowledge necessary to understand and apply the EGBV method to a large-scale industrial process. This chapter will serve as a reference for the subsequent chapters on EGBV application and analysis. The established theoretical framework will be consulted to ensure that the method is applied according to current best practices. In cases where the literature does not offer specific guidance for operations related to the analysed system, it will be used to support decision-making processes. With the theory established and the method understood, the next step in addressing the research question is to gain a comprehensive understanding of the system. This will enable the effective application of the EGBV methodology in the following chapters.

Chapter 3

THERESA System

3.1 Introduction

In this chapter, the complex thermal power plant with heat storage capabilities, as introduced in Chapter [1](#), is discussed in more detail by providing background on the system, discussing the layout of the system and elaborating on the components and functioning of the system. The system is a thermal energy storage test facility, referred to in further detail in this study by the acronym: THERESA.

3.2 THERESA background

Situated in Zittau, Germany, the test facility was developed and built by the Institute of Process Technology, Process Automation and Measuring Technology (IPM) at the Zittau/Görlitz University of Applied Sciences (UAS). THERESA consists of a reconstructed thermal water-steam process and is conveniently located next to the Zittau Stadtwerke where steam is produced for various applications, including its use by the THERESA system through inlets connected to various points in the system [\[52\]](#). The system includes a thermal energy storage (TES) component. This technology allows the accumulation of thermal energy by heating water to use the stored energy at a later time step for heating and cooling applications [\[53\]](#). THERESA is capable of emulating real power plant parameters and has been used for research on the flexibilisation of thermal power plants with the primary focus on TES, in an experimental environment [\[52\]](#).

The system is large and complex, but the basic functioning of the system can be explained at a high level using a block diagram as in Figure [3.1](#). The figure shows that the system has external inputs of water and steam which are used by components to deliver the desired outcome during operation; valves can be opened or closed to determine flow paths and the distribution of water or steam within the system, such as delivering hot water to the storage component or steam to the heat-sink component.

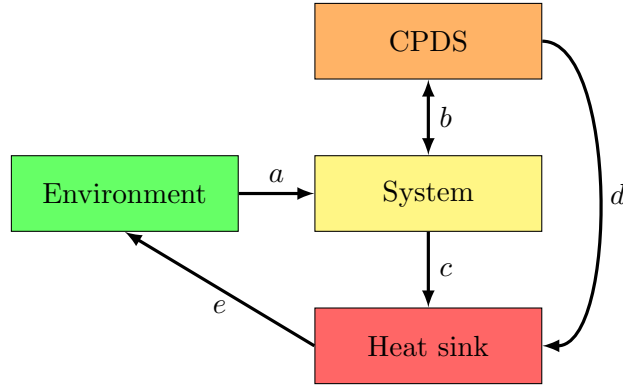


Figure 3.1: Simplified THERESA system diagram

For more detailed operation and understanding of the flow loops, a reduced system schematic is shown in Figure 3.2. There are a total of 20 components illustrated in this schematic, with the addition of a three-way-valve between the steam generator and the first preheater included. Water is introduced into the system with a pump from the deionized water tanks, the first preheater uses external steam to heat the water and this is transferred to the steam generator where electrical power is used to produce steam. This steam can be superheated through a similar component labelled the superheater. From here steam can be distributed through the system, either to the second preheater, the constant pressure displacement storage (CPDS) component, the universal interface or to the cooling circuit. In subsequent parts of this chapter, the different operating modes of the system is discussed.

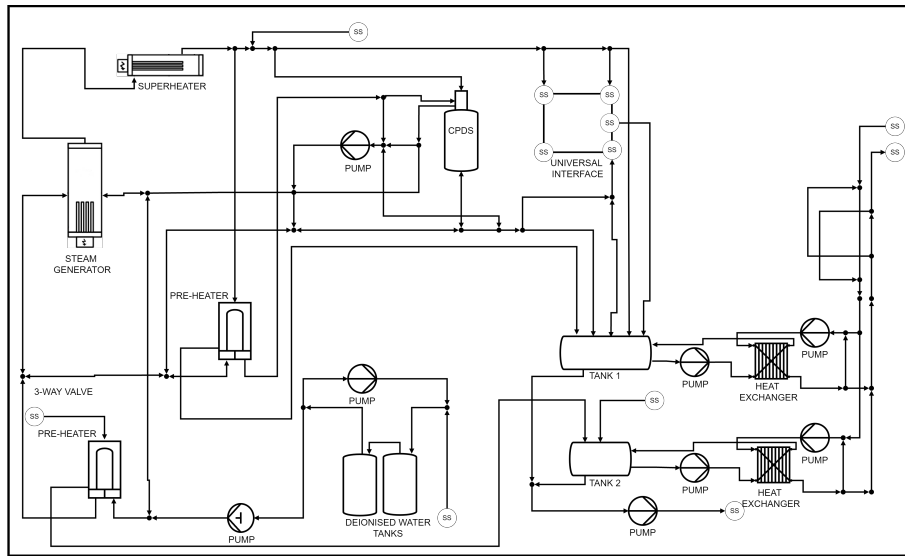


Figure 3.2: THERESA system schematic used in Chapter 4 for determining the system graph.

3.3 System layout and components

THERESA facilitates research on various aspects of power plant processes, including component analysis, dynamic operating behaviour, system integration, system optimisation, and safety-related investigations. In this section, components are described, and the system layout is discussed for context on the system operation.

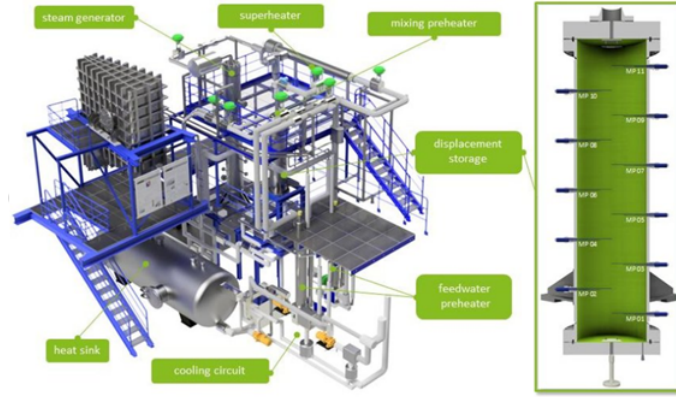


Figure 3.3: Theresa system 3D layout from [6]

3.3.1 Data acquisition components

The THERESA system has an integrated comprehensive sensor network and data acquisition systems, allowing real-time data collection of system parameters within components and flow paths. The prospect of rich data sets is valuable for researchers, as this allows thorough system analysis [52].

This sophisticated data capture network includes sensors for properties of water and steam, as well as system operation, and at various points, such as temperature, pressure, fluid flow, fluid level, and the state of valves as open or closed are also available. These measurements are described in detail in Chapter 4.

3.3.2 Preheaters

Preheaters are responsible for raising the temperature of the feed water to a desired level. Two upright U-tube preheaters are employed for feed-water preheating. These preheaters utilise a U-tube design, where the cold feed water flows through the tubes and is heated by the flow of steam in the shell space.

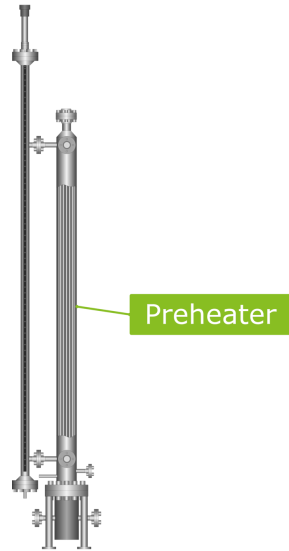


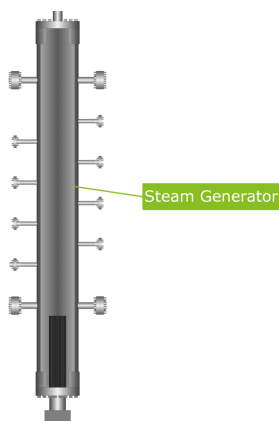
Figure 3.4: Preheater from [6]

3.3.3 Steam generator and super-heater

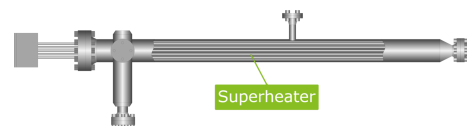
The steam generator is a power-to-heat steam generation unit. It is designed for the production and storage of saturated steam and is valuable for maintaining system pressure. It is equipped with an electric immersion heater with a heating capacity of 200 kW for steam generation.

A controllable electric superheater is connected downstream of the steam generator and enables the provision of super-heated steam with precise temperature control.

The steam generator and superheater are integral components of the THERESA system, which facilitate steam production and control.



(a) Steam generator.



(b) Steam Generator

Figure 3.5: Steam Generator and superheater from [6]

3.3.4 Heat sink: Discharge vessel with cooling system

The heat sink is a vessel with two compartments and multiple inlets and outlets connecting to various flow paths in the system. The heat sink's function is to condense steam and accumulate hot water. One of the compartments is coupled to an active cooling circuit in an arrangement similar to a shell and tube heat exchanger, where a cooling circuit transfers heat from the heat sink to the environment.

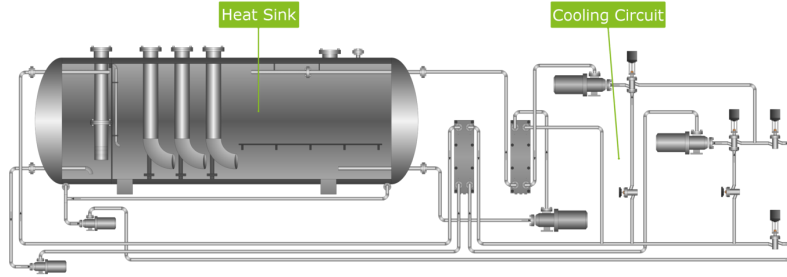


Figure 3.6: Heat sink and cooling circuit from [6]

3.3.5 Universal interface for component integration

The universal interface allows for the connection of further components and is utilised in various research projects. It enables the analysis of the integration of thermal energy storage systems in thermal processes of different nominal capacities and facilitates the investigation and testing of specific components such as heat engines or heat pumps.

3.3.6 Constant pressure displacement storage with mixing preheater

The CPDS operates at a thermal power capacity of $240 \text{ kW}_{\text{th}}$ and has an energy storage capacity of $60 \text{ kWh}_{\text{th}}$. Designed to withstand high operating conditions, the unit functions at a pressure of 60 bar and a maximum temperature of 265°C , with a total storage volume of 0.6 m^3 . The accumulator is charged and discharged using high-temperature pressurised water. The CPDS can be regarded as sensible heat storage that utilises hot water from the boilers and preheaters to store thermal energy for defined periods to be discharged at a later stage [54]. The CPDS serves as a high-pressure saturated water reservoir, enabling thermal buffering by storing and discharging steam energy as required. Its function is critical in load-following applications, particularly in systems with high variability in thermal demand.

The CPDS is integrated within a broader thermal high-performance storage system (TH-PSS), which comprises the CPDS and a mixing preheater (MP). The CPDS enables the storage of saturated water with minimal thermal losses, allowing for high-efficiency energy retention and controlled discharge. The mixing preheater, functioning as a direct-contact heat exchanger, facilitates high-capacity thermal exchange through the direct condensa-

tion of steam into water. This setup allows for efficient thermal energy transfer and rapid response to system demand fluctuations.

To ensure precise monitoring and performance evaluation, the system incorporates multiple axial measuring planes, each containing radially distributed temperature sensors. These allow for accurate thermal stratification analysis within the storage system through all operational modes of the CPDS, these modes will be discussed in more detail in Section [3.4](#).

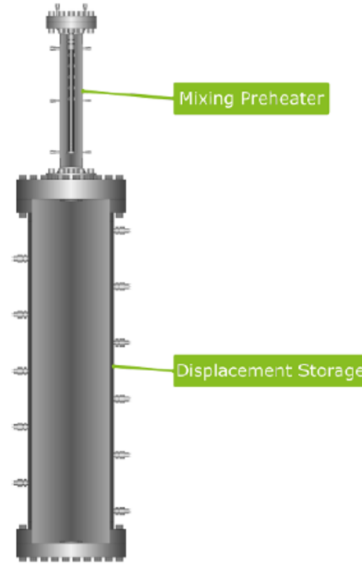


Figure 3.7: Constanet pressure displacement storage component from [\[6\]](#)

3.3.7 Other components

Other components within the system include pipes for the transport of fluid, feedwater tanks for storage of deionised water, driers for removing condensate from steam, and valves for control of fluid in terms of flow rate and flow paths.

3.4 System operation

The system operation has been simplified in Figure [3.1](#), which is only for a high-level explanation of the typical operation of the system. The system can mainly be described as four process blocks and five links between the blocks. The process blocks are environment, system, CPDS, and heat sink, where the links *a*, *b*, *c*, *d* and *e* represent flow paths of energy. The energy interactions can be observed as directional, either in one direction or bidirectional. In the context of operation, the main function of the system is to heat water to a certain temperature and pressure, this is referred to as the warm-up mode. After the working fluid is heated, the system then distributes the working fluid to the CPDS at a specific flow rate into the CPDS to enable charging of the CPDS, also referred to as the

charging mode. The system will then reroute the working fluid to circulate within the system or bypass the CPDS to distribute the working fluid into the heat sink, working as an open-loop system. This is referred to as the holding phase of the CPDS. After a set period has passed, the CPDS moves into the discharging phase, where it discharges the working fluid to either the heat sink or back into the system. The heat sink also functions as an open-loop system whereby hot working fluid is received from the system or the CPDS, and energy is transferred to the environment through either heat exchange or direct fluid expulsion to the environment.

3.4.1 System experimental configuration

The THERESA system can operate with various active or inactive components. For example, when the CPDS component is used during an experiment, the universal interface for component integration can be inactive for an entire experiment. Thus, the system experimental configuration will be used to describe the active components within the system for a given experiment, as this is highly dependent on the experiment goal described in Subsection [3.4.2](#).

3.4.2 Experiment goal

For this study, the system experimental configuration includes the following key components as active components: steam generator, preheaters, CPDS, heat sink with cooling circuit, feed-water tanks, and various pumps. The large components that are inactive during this study are the superheater and the universal interface.

The operational cycle of the CPDS consists of charging, holding, and discharging phases, controlled through a network of valves and pipelines that regulate energy flow within the system. During the charging phase, high-temperature steam enters the CPDS, where it condenses and transfers its thermal energy to the stored water. The holding phase maintains the stored energy with minimal losses, utilizing an insulated design to preserve thermal stratification. In the discharging phase, thermal energy is released in a controlled manner by drawing heated water from the storage tank, which is then converted back to steam or used directly in process heating applications.

The system's operational goal during all data sets acquired is to warm up, charge, hold, and discharge. That is, to warm up the system from a cold start by bypassing the CPDS, then charge the CPDS with water at a specific high temperature and with a specific flow rate, then hold the fluid within the CPDS for an undefined period, and discharge the fluid from the CPDS. In the sub-sections to follow, the charge, hold, and discharge modes are discussed in detail using Figure [3.8](#).

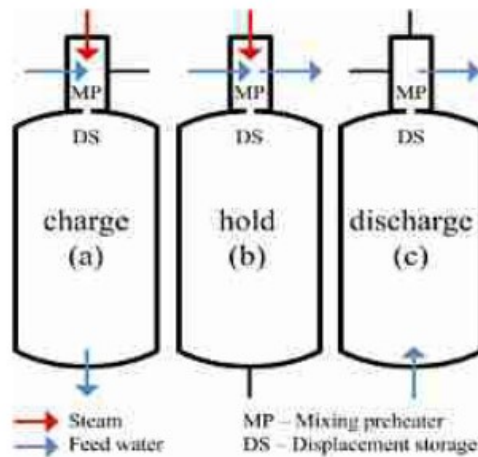


Figure 3.8: CPDS charge hold and discharge modes from [7]

Startup, warm-up and shutdown system modes

In the startup mode, the system is started from a cold state and enters the warm-up mode as the system is prepared for operation by gradually warming up components. Gradually, the cold working fluid is heated within the various heating components to generate hot working fluid, and the hot working fluid is circulated within the system to warm certain components to a desired temperature. The system pressure is also gradually increased simultaneously to the desired pressure. During the warm-up mode, the CPDS can be charged and discharged as necessary to reach the desired system and component temperature levels before proceeding into the main charging phase. The charge, hold, and discharge events within the warm-up mode are considered as warm-up modes and not as charging, holding, and discharging modes as in the subsequent modes.

Charge CPDS mode

During the CPDS charging phase, valves are opened within the system to route high-temperature pressurised fluid to the CPDS inlet. As the fluid passes through the inlet of the CPDS, the low-temperature fluid is displaced by the hot fluid, forming a thermal boundary layer between the cold fluid and hot fluid. The charging process is complete once the inlet and outlet valves are closed.

Hold CPDS mode

The holding mode of the CPDS stores thermal energy by preventing the fluid from entering through the inlet or being discharged through the outlet. The pressure of the fluid prevents it from boiling, even if its temperature is above the boiling point of the fluid at atmospheric pressure.

Discharge CPDS mode

The discharge mode of the CPDS is the exact reversal of the charging mode. Here, a low-temperature fluid is input to displace the hot fluid while still maintaining the thermal boundary layer.

3.4.3 Identifying system modes

The modes of the system are governed by the valves surrounding the MP and the CPDS. These valves are 2A22, 2A26, 2A27, and 2A28. The state of these valves at a given time step is an indication of the transition areas expected within the THERESA system throughout an experiment. In addition, the temperature within the CPDS is an indication of the phase of the experiment in which certain transitions are expected or observed. Figure 3.9 shows the position of these valves on the flow paths in and out of the CPDS.

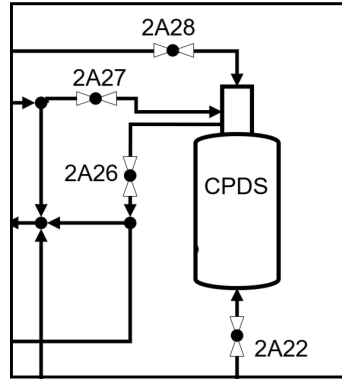


Figure 3.9: CPDS control valves on flow paths in and out of CPDS

In Figure 3.10, for data set A the temperature within the CPDS is plotted over time. The experiment starts with a warm-up phase with various points of temperature fluctuations as the system prepares for the charging phase at about $t = 15500$ seconds. The temperature within the CPDS rapidly increases (a clear indication of charging) as the temperature within the CPDS reaches a constant temperature of about 235°C , when the system is said to be in a holding phase. At about $t = 2630$ seconds, the system changes to discharge mode and a rapid decrease in temperature is observed, whereafter the temperature is constant.

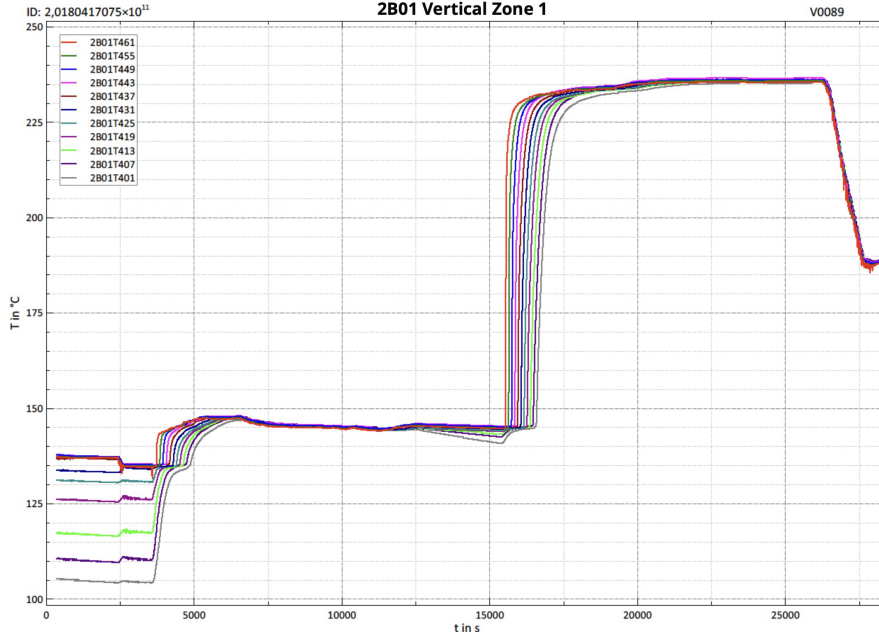


Figure 3.10: CPDS temperature during operation (data set A)

Utilising the valve states and the temperature data of the CPDS, mode changes can be linked to changes in the valves of the CPDS that allow fixed reference points for analysis. Figure 3.11 shows the plot of temperature in the CPDS over time for data set B. The key differences between data set A and B are that the parameters of the experiments differ as per Table 6.1 this includes the intervals of each mode being different causing a shift in the observed temperature plot over time for the CPDS. Similar to data set A the experiment starts with a warm-up phase with various points of temperature fluctuations as the system prepares for the charging phase; at about $t = 6500$ seconds, the temperature within the CPDS rapidly increases (a clear indication of charging) as the temperature within the CPDS reaches a constant temperature of about 235°C ; the system is said to be in a holding phase. At about $t = 20500$ seconds, the system changes to discharge mode and a rapid temperature decrease is seen, where the temperature is constant.

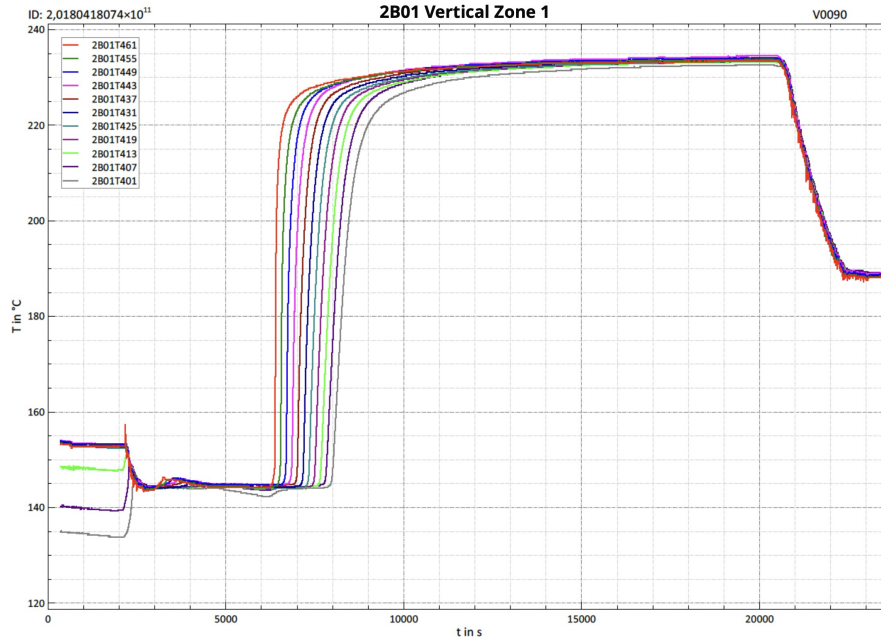


Figure 3.11: CPDS temperature during operation (data set B)

These temperature graphs are used to illustrate the expected modes within each data set over time and will be further elaborated on in Chapter 6.

3.5 System process monitoring

The THERESA system is a physical system without a completely validated process model. To the author's knowledge, a partial model was utilised during research conducted with a specific focus on the CPDS [55, 7]. Currently, the process monitoring of the THERESA system is facilitated through component control and process variable monitoring. Some components have integrated control procedures which are mainly based on thresholds for maximum and minimum process variables. If a fault occurs, the system will record at which point a threshold is crossed by a specific process variable. Using an observer-based method, THERESA plant operators may make informed decisions if the system control does not cause corrective action.

3.6 Conclusion

The THERESA system is a complex system with various favourable characteristics for research within condition monitoring. The operational, structural, and technical information of the THERESA system, conveyed in this chapter, provides the knowledge necessary for understanding the THERESA system. In Chapter 4, the information on the application of EGBV for FDI from Chapter 2 and the system information from this chapter are used in tandem to apply EGBV for FDI in the THERESA system.

Chapter 4

EGBV on the THERESA system

4.1 Introduction

The theory and concepts behind the application of the EGBV method for FDI in a large-scale industrial system have been discussed in Chapter 2. This chapter discusses how the EGBV method for FDI is applied to the thermal storage test facility under investigation, namely the THERESA system. The objective of this chapter is to utilise EGBV operations to produce a graph that is representative of the system. For a system graph to be representative of the system it should utilise the available information as thoroughly as possible and adhere to current best practices within the research field. The objective of this chapter is to develop an attributed graph representation and node signature representation for the THERESA system by following these guidelines.

The implications of different graph configurations are not investigated; however, many aspects of graph reduction are taken into account throughout the relevant operations.

4.2 EGBV methodology outline

From the literature study and the review of the literature in Chapter 2, the operations required for the application of EGBV for FDI to a complex system are a linear process with defined steps. Completion of these steps, given that the operation of a system is understood and all information is available, leads to the deliverables expected from the EGBV methodology, the system-attributed graph.

With the system operation and knowledge obtained in Chapter 3, the other aspect is system process data that are required as input to the process. Within the literature, such process data are obtained from the use of the physical system or a process model. In this study, historical process data are available from previous experiments.

With the THERESA system, it is evident that although the system is well equipped with sensors, the system does not have all process variables at all positions.

Without the availability of a process model where such variables are readily available, a methodology is devised to determine the necessary information for the application of the EGBV method on the THERESA system, taking into account the information available and allowing multiple iterations of certain processes to overcome the information deficit.

Thus, the application of the EGBV methodology can be used in an iterative process as shown in Figure 4.1.

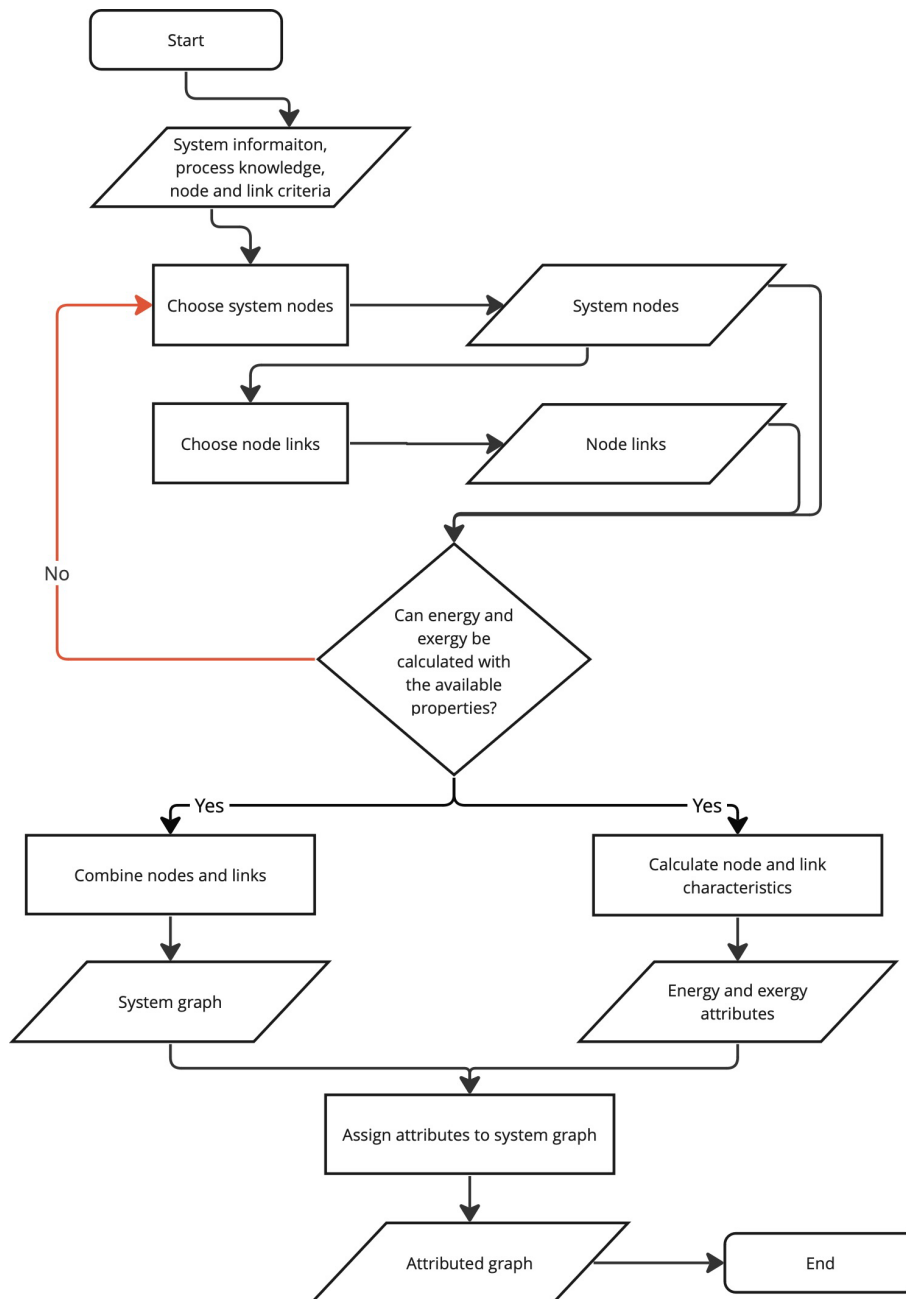


Figure 4.1: System flow for attributed graph configuration

The process is initiated with the acquisition of system information, expert knowledge of the system processes, and criteria for the selection of nodes and links, as seen in Figure 4.1. The following processes are dependent on node selection. The selection of nodes are based on the acquired information and criteria. Upon selection of nodes, links can be determined. To ensure that the system graph is viable it is necessary to evaluate the links to ensure that energy and exergy attributes can be calculated from the system graph. This evaluation may expose areas within the available data where assumptions are necessary to calculate energy or exergy attributes. In cases where insufficient information is available to determine the energy or exergy attributes, the system nodes are re-evaluated. The process is repeated until the resulting system graph is well-defined regarding information available for energy and exergy attribute calculation. Parallel processes may follow where the selected nodes and links are combined to form the final system graph, and the energy and exergy attributes for the final system graph are calculated. The attributes are allocated to the system graph to form the system-attributed graph.

The methodology is implemented in MATLAB[®] to meet the computational demands of the methodology applied to a complex system with substantial data loads. These computational demands can vary further based on various decisions made in the EGBV application process, such as the nodes chosen and the input data provided.

4.3 System graph setup

The system graph is the framework to which the energy attributes are related to form the system-attributed graph. For a complex system such as the THERESA system, it is necessary to ensure that the system graph is configured carefully and accurately, utilising the available information and tools. As seen in Figure 4.1, this process is iterative and depends on other processes.

4.3.1 Process knowledge

Throughout this chapter, *a priori process knowledge* is used and refers to the knowledge that is available about a process not strictly obtained through experience but could be obtained through various sources such as academic sources or information from experts. Process knowledge includes factual information from experts, system schematics, system history, and expected system response. This can be in the form of schematics of the system and factual documents, data sets, and other relevant information on the system. This knowledge is applied where appropriate during the various processes necessary to meet the predetermined objectives.

From preliminary discussions with staff at the THERESA system it was agreed that data are available that have distinct faults, and data are available that are defined as normal operation. This allows the process to be applied without a process model given the avail-

ability of measurements within the system in normal and fault operation. A data-informed methodology is followed to facilitate the necessary subsequent analysis.

A unique characteristic of the THERESA system being an experimental large-scale system is that it is equipped with a dense network of sensors measuring process variables. Table 4.1 shows the process variable measurement data available and the units that are associated with the process measurement data.

Table 4.1: THERESA system process variable measurements and units

Measurement type	Measurement units
Pressure	bar, mbar
Temperature	°C
Fluid level	%, cm
Fluid flow rate	m ³ /h, kg/s
Control Logic (open or closed)	Logic operators
Electrical power	kW

Certain process variables are not available from sensor measurements. If insufficient measurements are available at a specific point to determine the energy at such point this is referred to as a location of uncertainty.

After the nodes and links have been chosen, it is necessary to evaluate if all measurements required to determine the attributes are available. If not, the required variables should be calculated from available variables; if it is not viable to calculate the unknown process variables, the nodes and links will need to be reevaluated.

The complexity of the system requires skilled staff to understand various processes and subprocesses to ensure that the system is operated safely and correctly. During the investigation of the THERESA system, the operators of the plant are valuable in understanding the working and operating principles of the system.

The THERESA system rarely maintains steady state operation for prolonged periods. Throughout a single operational cycle, the system dynamically transitions between various operational points. It is valuable to understand the system modes when transitions occur, and why a given transition occurs. From Chapter 3 the theoretical operation of the CPDS component is described, and the behaviour of the system throughout the mode changes requires further investigation.

Based on expert knowledge and literature on the operation of the system, the THERESA system's modes should be easily identifiable from data acquired from the valves controlling the flow paths connected to the CPDS. However, interpretation of the valve data shows that these valves do not function exactly as expected. This can be a result of the system control or updates to the system that have led to the information becoming outdated.

An additional source of information is the internal temperature of the CPDS, the system mode can be interpreted to a large extent by evaluating the internal temperature of the CPDS.. As stated in Chapter 3, the CPDS temperature is indicative of the system mode when compared between the system modes. In Figures 3.10 and 3.11 and preceding discussions the link between the temperature within the CPDS and the mode changes are described, thus utilisation of the system temperatures are important for mode determination.

4.3.2 System structural information

As discussed in Chapter 2, the measurement space includes structural information of the physical system. This structural information is the connection between components from an energy perspective. When a component transfers energy to another component through a pipeline or other form of structural connection such as mechanical or electrical coupling, it is said that these components are connected. This structural information is examined in the measurement space and transformed into the feature space by feature extraction and feature selection. System structural information and system process variables are utilised in the transformation from the measurement space to the feature space.

Figure 4.2 shows the THERESA system schematic. This schematic is a simplified representation of the system's structural information that relates components with flow paths. From the structural information of the system, nodes and links are selected (feature selection) to represent a complex system as a graph representation of the system.

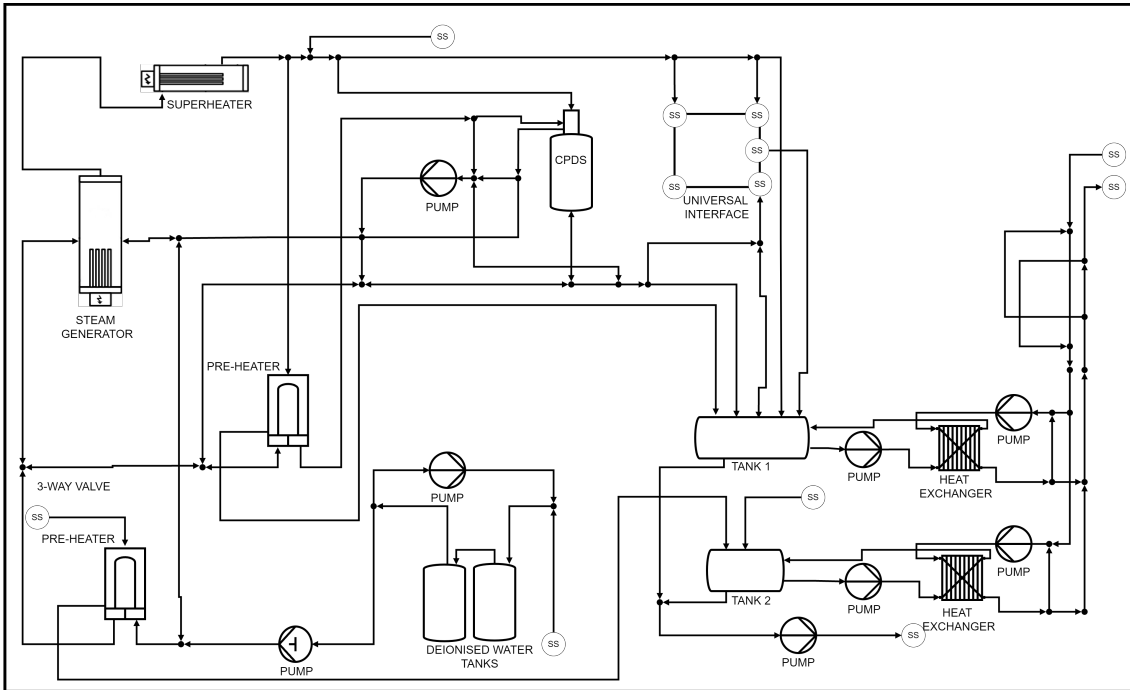


Figure 4.2: THERESA system schematic

This structural information is the basis on which the application of graph theory is applied.

The structural information is a fixed reference to which the energy and exergy attributes are linked.

4.3.3 Mass and energy balance equations

In the measurement space, energy balance equations are derived to be applied in the feature space, as a way to reduce dimensionality. The equations for mass and energy balance are formulated in Chapter 2 and verification employing these equations is demonstrated in Chapter 5.

The energy and mass balance equations serve as tools used to relate process variables between components of the system. These equations are used in the literature when setting up system models. In this case, the process is not modelled, however, it is required to determine values for process variables that are not measured by sensors.

The components of the THERESA system operate in the physical domain, allowing the use of thermodynamic relations to represent properties different from those in the unifying domain of energy. Physical domains are represented by measurement data and these measurements are related by energy and exergy balance equations to energy as a unifying parameter.

Within the MATLAB[®] implementation, the process data are related to energy, first by relating process variables to the energy domain. This is done with the equations as discussed in Chapter 2.

The MATLAB[®] function XSteam is used to determine the thermodynamic properties of water and steam. XSteam for MATLAB[®] is an implementation of the IAPWS IF97 standard formulation which provides accurate data for water, steam, and mixtures of water and steam properties from 0 - 1000 bar and 0 - 2000 °C [56]. This function is used for the calculation of density (ρ), enthalpy (h), and entropy (s). The function receives input values of temperature (T) and pressure (P) at various points and the output is determined based on the variable being calculated.

The following properties are calculated using this function to evaluate if the nodes and links are well-defined in terms of available data.

Fluid density: $\rho = \text{XSteam}(\text{'rho_pt'}, T, P)$

Enthalpy: $h = \text{XSteam}(\text{'h_pt'}, T, P)$

Entropy: $s = \text{XSteam}(\text{'s_pt'}, T, P)$

Although the attributes for the energy-attributed graph are determined in a later section, it is necessary to evaluate if the current iteration of the system graph is feasible for further application of the method; thus, it is necessary to understand the equations that will later be used to know if sufficient process variables are available at the required nodes and links to have a fully defined system graph.

If the required inputs are not available, and the undefined process variables cannot be calculated from known process variables through thermodynamic relations or sound assumptions, the system graph should be reevaluated to address this uncertainty. This is done through various graph reduction techniques as described in subsequent sections.

4.3.4 Node selection and link determination

Utilising system *a priori* knowledge, decisions regarding the selection of nodes and links from system structural information are informed. The criteria for the system graph setup, outlined by Smith [3], is a guide to the selection of nodes involving five main questions.

Firstly, the identification of components crucial for FD within the system is determined in consultation with plant operators. Ideally, plant operators would be able to detect faults within the entire system. However, the following components are considered priority in FD: Pump 1, preheater 1 and preheater 2, steam generator, superheater, MP, CPDS, Universal interface, tank 1 and tank 2, cooling circuit pumps, and heat exchangers.

Secondly, components with multiple inflows or outflows of energy are identified through a process involving system schematics and sensor data evaluation. The information available is sensor data information, time series operational system variables, and system schematics showing components and the sensor locations. This process involves using the system schematics to identify sensors around the components on the flow paths linked to that component. In addition, with the utilisation of the time series data from the selected sensors, the energy is qualitatively evaluated. The next step is to choose a new node and follow the qualitative energy evaluation, repeating this iterative process for all nodes considered. When the above process has been completed, the system has been qualitatively evaluated for energy values, thus the question of which components have many inflows and outflows of energy is answered. Simultaneously the question of at which components do large amounts of energy flow in/out can be answered as within the procedure outlined above the relative energy values can be determined and evaluated.

Thirdly, this evaluation extends to determining components where small energy changes can significantly affect the system. This question focuses primarily on specific components that may exist in large-scale systems heavily reliant on prompt FD, control-related components, and sensitive components related to energy changes. An example of this is a turbine in a power plant that operates at very high rotational speeds. If a fault occurs that causes a sudden change in pressure, this component may be directly affected, thus also considered as a fault propagating to this component. In terms of system control, this can be viewed as a component where system control can cause rapid changes in process variables that influence energy. Even a small change may result in changes in the overall energy across or within the component. The components identified within the THERESA system are the CPDS, boilers, and tanks where system control may lead to sudden pressure losses or increases that may affect the system components.

Finally, the complexity of the graph representing the system is considered, whereby the computational burden increases with a more complex graph. This entails careful selection of components for inclusion; within a complex thermal power plant with many components, the computational load can be substantially reduced by carefully selecting which components to include and which components to neglect to compose a graph that is representative of the system but does not take into account redundant or unnecessary information. Valves are often neglected under the assumption that faults within them may propagate to nearby nodes and be detected within these components. In general, this approach ensures a systematic and informed method for node selection and graph composition within the system.

The components represented by the nodes are identified based on process knowledge and criteria, as depicted in Figure 4.3. The system components outlined in red are identified as possible nodes: these include critical components such as pump 1, preheaters, steam generator, superheater, and other components selected based on established criteria.

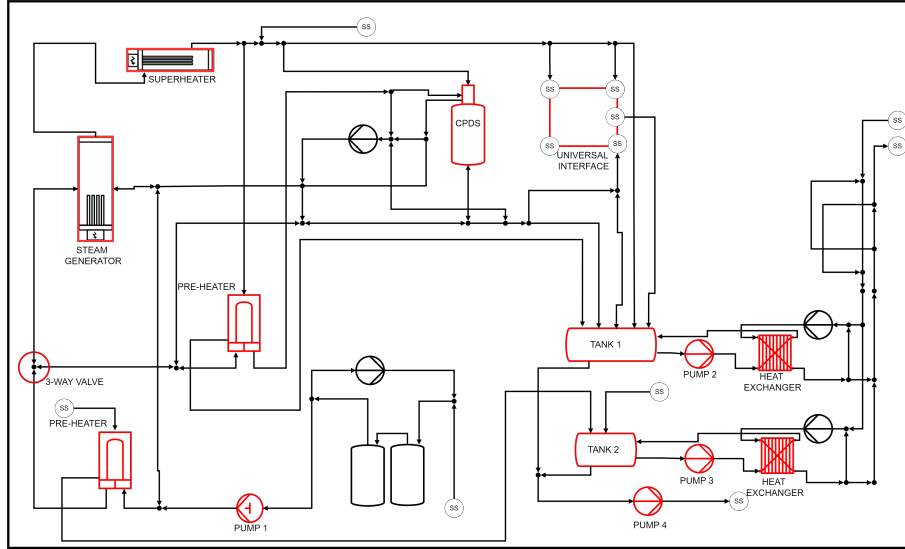


Figure 4.3: Possible system nodes in THERESA system highlighted in red

Links are defined by implementing process information, which is information about the connection of flow paths between components. These links are determined after nodes are selected, the links can be evaluated based on the available data, and the appropriate equations are set up for further use in the energy and exergy attribute determination. In cases where energy is transferred across the system boundary, the environmental node is required. The environmental node represents the system boundary; thus, if energy is transferred across the system boundary, such a node has a link to the environmental node.

Although this node is regarded as a single node, for simplified representation, the environmental node is represented as separate nodes with the same number. The power provided externally to the components $\dot{W}_{in}$ is considered to be a link from the environment to the component to which the power is provided.

In addition, fluid inlets and outlets to the environment, thus fluid that is externally sup-

plied or expelled to or from the system, are considered a link to the environment, since this inlet and outlet is determined as the system boundary.

After the schematics and flow information have been reviewed and discussed with the plant operators, the system can now be represented as a graph after certain nodes have been determined and their flow paths chosen.

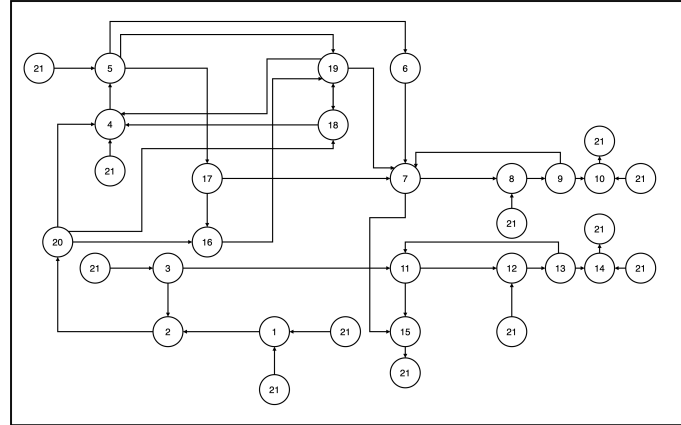


Figure 4.4: Full system graph with all possible THERESA system nodes included

Figure 4.4 shows the largest possible graph that can be determined for the THERESA system from the available information. Some components are not included in the final system graph from the initial selection of possible nodes. A lack of information on those components causes nodes to be combined into a single node, or process information that has led to certain nodes and links being neglected based on the availability of data or the fact that they are declared inactive for most of the process.

The nodes seen in Figure 4.4 are described below as to which nodes are which components.

Table 4.2: Description of final selected system graph nodes

Node	Description
1	Pump (1P02)
2	The secondary side of preheater 1 (1W01b)
3	The primary side of preheater 1 (1W01a)
4	Boiler (1B01)
5	Super-heater (1W03)
6	Universal-Interface (UVS-HD)
7	Tank 1 (3B01)
8	Pump (3P01)
9	The primary side of heat exchanger (3W01a)
10	The secondary side of heat exchanger (3W01b)
11	Tank 2 (3B02)
12	Pump (3P03)
13	The primary side of heat exchanger (3W02a)
14	Secondary side of heat exchanger (3W02b)
15	Pump (3P05)
16	The secondary side of pre-heater 2 (2W02b)
17	Primary side of pre-heater 2 (2W02a)
18	Constant pressure displacement storage (2B01)
19	Mixing pre-heater (2W01)
20	Temporary node (3-way junction connecting 1A21, 1A17 and 2A01)
21	Environmental node

In Figure 4.3, the system graph can be qualitatively configured based on the selected components and the links between the nodes.

In this case, the nodes highlighted in red are areas identified to be reduced to a single node as these nodes do not have sufficient information to be well-defined.

Nodes 18 and 19, the MP and the CPDS, have a certain flow rate from 19 to 18 or 18 to 19. This is very mode-dependent. The flow rate measurement between nodes 19 and 18 is non-directional, thus the sensor does not measure the direction of flow, only the value of that flow. This poses a problem for the energy balance of these nodes and the node signature matrix. For energy balance, the flow direction should be indicated by positive values in one direction and negative values in the other. This is not possible to infer, as the system mode cannot be fully defined within such a measurement set.

The MP and CPDS function as a single component, the nodes can be merged, and the energy interaction can be regarded as fully within the boundary of the node, thus nodes 18 and 19 will be referred to as node 19.

In the case of node 15, the information for the interaction between nodes 15 and 21 is insufficient to fully define the energy interaction, thus node 15 will become part of the

environment (node 21); there is a direct interaction between nodes 7 and 21, and nodes 11 and 21.

Nodes 10 and 14 are the secondary sides of the heat exchangers, which are part of the cooling circuit. These interactions on the secondary side have insufficient information to define the energy interaction between nodes 10 and 21 and nodes 14 and 21; thus in both cases, the primary side of the heat exchangers, which is nodes 9 and 13, will interact directly with the environment, which is node 21. Thus, nodes 10 and 14 will become part of environmental node 21.

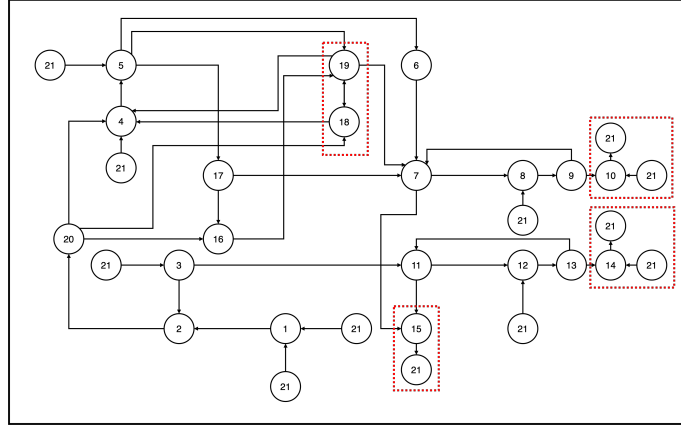


Figure 4.5: Full system graph before graph reduction, with reduction areas highlighted with red boundary

After the graph reduction methods are applied to the red outlined parts of Figure 4.5, the final system graph representation can be illustrated as in Figure 4.6. Where all nodes are well-defined and all links are well-defined in a qualitative manner.

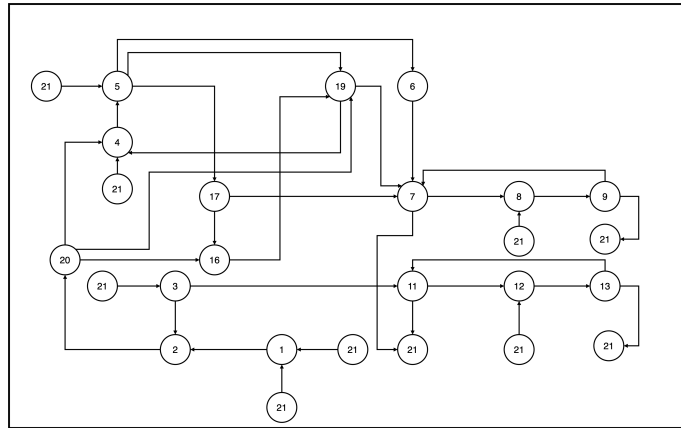


Figure 4.6: Full system after reduction

4.4 System-attributed graph setup

Process variables (pressure, temperature, mass flow, and others) are extracted from the measurement space, and energy and exergy values are calculated from the process vari-

ables.

Energy and exergy are attributed to the nodes and links of the system graph, and the system-attributed graph is formed. In addition, the system-attributed graph can be represented in an equivalent matrix format, which is referred to as the node signature matrix.

Dimensionality is reduced through the universal domain of energy and through the simplification of a complex system into a graph representation.

The reductive capabilities of the method are evident as the final system-attributed graph will consist of fewer nodes than the initial proposed system graph; however, for consistency, node numbers are not revised on each iteration. Each excluded node is justified through sound reasoning, and nodes are omitted either due to their inactivity within the operational datasets, insufficient available information (thus categorised as part of the environment), or consolidation into a single node (as exemplified by the merging of nodes 18 and 19). As stated in the Introduction, the objective of this chapter is to produce a representative graph, using the available information adequately and adhering to current best practices in the application of the method.

For an illustration of the different steps that are used in this study to determine the system-attributed graph, Figure 4.1 shows a process diagram. The relevant processes are discussed in the sections that follow.

4.4.1 Defining node energy interactions

Links are edges between nodes. To be able to keep track of possible interactions between nodes, it is noted that nodes may have more than one interface by which energy can be exchanged with other nodes. These various inputs or outputs are defined for each node and are labelled an IO (input/output).

In the case of a generic node n_i as shown in Figure 4.7 this node can have multiple IOs: there may be k possible inputs and outputs to node n_{i,IO_k} which refers to the i -th node's k -th IO. This notation proves useful for large systems with many components and thoroughly describes the energy interactions of each node within the system. To explain

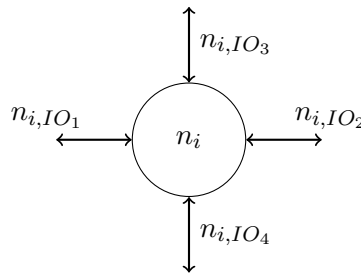


Figure 4.7: Generic node with four IOs

how this notation is useful for link determination, a case with two generic nodes, n_i and n_j , is considered. These links are said to be connected with a link $l_{i,j}$. This link is suitable for the construction of the system graph. However, in a complex system to assign weights or attributes to the link, the link needs to be well-defined in terms of the IOs utilised within such energy interaction. This is illustrated in Figure 4.8 where the nodes are connected by a single link. The link is defined as follows: From node 1 to node 2 through IO 2 of node 1 and IO 1 of node 2. Thus, $l_{i,j} = [n_{i,IO_2} \ n_{j,IO_1}]$

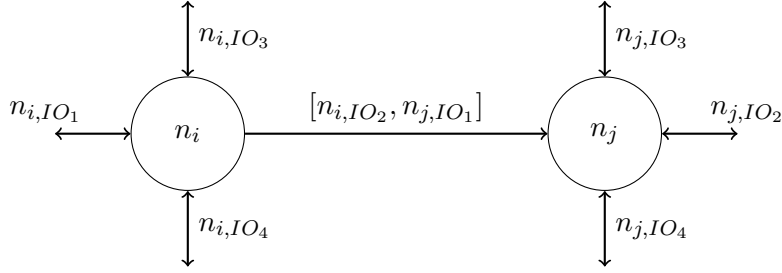


Figure 4.8: Generic representation of link definitions

The energy interactions table as in Table 4.3 within this table, an interaction is described in terms of the from node, the IO of the from node, the to node, and the IO of the to node.

The environmental node is not considered to have IOs, as seen in Table 4.3. In cases where energy interactions with the environment occur, the IO will be defined as electrical power or as fluid transfer from the environment. In cases where the heat exchange takes place between components, this interaction is not a measured value, but rather a calculated attribute, thus this notation makes provision for such calculations.

Table 4.3: System graph energy interactions

Interaction number	From Node	IO of From Node	To Node	IO of To Node
1	1	2	2	1
2	2	2	20	1
3	3	Heat_Exchange	2	Heat_Exchange
4	3	2	11	1
5	4	2	5	1
6	5	2	17	1
7	5	2	19	1
8	7	7	8	1
9	7	2	21	Fluid from environment
10	8	2	9	1
11	9	Heat_Exchange	21	Heat_Exchange
12	9	2	7	2
13	11	3	12	1
14	11	2	21	Fluid from environment
15	12	2	13	1
16	13	Heat_Exchange	21	Heat_Exchange
17	13	2	11	1
18	16	2	19	4
19	17	2	7	6
20	17	Heat_Exchange	16	Heat_Exchange
21	19	2	4	2
22	19	3	7	5
23	19	3	4	3
24	20	2	16	1
25	20	3	4	1
26	20	2	19	2
27	21	Fluid from environment	1	1
28	21	Electrical Power	1	Electrical Power
29	21	Fluid from environment	3	1
30	21	Electrical Power	4	Electrical Power
31	21	Electrical Power	8	Electrical Power
32	21	Electrical Power	12	Electrical Power

From this energy interactions table, a system-attributed graph can be calculated by ensuring that all energy interactions can be determined from the available flow as enthalpy, entropy, and power data. Thus for each interaction, the energy and exergy are calculated, and an algorithm implemented in MATLAB[®] sorts the energy values into the relevant positions within the system-attributed graph. The exergy change across each node is also determined in the relevant position.

4.4.2 Energy and exergy attribute determination

As the mass and energy balance equations are formulated in the measurement space, the energy equations are applied in the feature space to the measurement space to map onto the decision space.

Within the decision space, the energy and exergy values are attributed to the system graph, forming the system-attributed graph and node signature matrix. A single node signature matrix contains energy and exergy values that are attributed to nodes and links of the system graph within a particular instance. A node signature set describes the energy and exergy attributes of a system over a given operational period. From the node signature matrices, the next step in the diagnostic system for FDI will be to define references in the normal operational state and then compare normal references with operational data to form cost matrices.

As discussed in Section 2.2, defining the dead state in the exergy analysis is integral in the energy and exergy analysis, thus the dead state is defined as $P_0 = 101.325$ kPa or 1.0325 bar and $T_0 = 20^\circ\text{C}$, which is the ambient pressure and temperature that is experienced in Zittau, Germany, where the THERESA system is located. The ambient temperature is assumed to be constant as the system is housed within a building separated from the direct environmental temperature changes.

4.4.3 Assumptions underlying the energy characterisation

Within the THERESA system nodes are determined and these nodes are evaluated as suitable if the required energy interactions with other nodes can be well-defined. Upon gathering information on sensor types and sensor placement, a list of possible uncertain parameters can be created, which involves determining the component, the uncertainty, and the resolution to the uncertainty. This is also described as underlying assumptions to the analysis, which is essential in a data-driven approach and is summarised in Table 4.4.

The applied solutions to the underlying concerns at various nodes, is based largely on assumptions that provide a scope for what would be acceptable outputs for the system. These outputs cannot be measured to practical data quantitatively as this measurement is unknown. The lack of a process model also deems such a comparison impossible. In

Table 4.4: Table of underlying assumptions necessary for energy characterisation

Node	Component	Component ID	Underlying concern to address	Applied solution
1	Pump	1P02	Power input to pump unknown	Pump power calculation
3	Preheater primary side	1W01a	Flow rate out of primary side unknown	Tank calculations
4	Boiler	1B01	Flow rate into boiler unknown	Flow rate to boiler is zero for all experiments
6	Universal interface	UVS-HD	Not active in experiment	Flow rate assumed to be zero for all experiments
7	Tank 1	3B01	Output flow rate unknown	Tank calculations
8	Pump	3P01	Input pressure unknown	Input pressure is calculated from known pressures
			Input power unknown	Pump power calculation
9	Primary side of heat exchanger	3W01a	Input pressure unknown	Calculated from known pressures
10	Secondary side of heat exchanger	3W01b	More than two process variables unknown	Combine as part of the environmental node
11	Tank 2	3B02	Output flow rate unknown	Tank calculations
			Output pressure unknown	Calculated from known pressures and fluid pressure in tank
12	Pump	3P03	Input pressure unknown	Input pressure is calculated from known pressures
			Input power unknown	Pump power calculation
13	Primary side of heat exchanger	3W01a	Input pressure unknown	Calculated from known pressures
14	Secondary side of heat exchanger	3W02b	More than two process variables unknown	Combine as part of the environmental node
15	Pump	3P05	More than two process variables unknown	Combine as part of the environmental node
16	Secondary side of heat exchanger	2W02b	Output flow rate unknown	Tank calculations
17	Primary side of heat exchanger	2W02a	Output flow rate unknown	Output flow rate assumed to be equal to input flow rate
18	CPDS	2B01	Identification of direction of input flow rate is not possible	CPDS and MP assumed to be a single component
19	MP	2W01	Input flow rate and output flow rate unknown	CPDS and MP assumed to be a single component
			Flow paths of flow expected is reliant on experiment	Fixed expected flow paths for all data sets
20	3 way valve	Temp node	Flow rate to boiler unknown	For all experiments flow rate to boiler is zero

Chapter 5, these assumptions are discussed and the solutions evaluated to ensure that the solutions applied hold merit.

4.4.4 Calculation of unknown process variables

Unknown or unmeasured process variables are calculated in the sections below, including underlying assumptions and methodologies applied.

For pumps utilised to circulate working fluid (mostly water) the input power values are not available from the measurement data. Based on the available data the fluid flow rates, pressure, and temperature can be utilised to determine the energy of the flow streams into and out of the pumps and the equation for power delivered to a fluid stream can be utilised to determine the power delivered to the pump. This is given in (4.1), assuming no energy loss through inefficiencies.

$$\dot{W}_{in} = \dot{m}(h_{in} - h_{out}) \quad (4.1)$$

The power transferred to the fluid stream can be compared to the actual pump size to determine if the results are consistent with the expected values.

Within the THERESA system, there are four heat exchangers, two shell-in-tube U-tube heat exchangers operating as the main heaters for water transferred to the CPDS component, and two that form part of the heat sink and cooling circuit. The mass balance equation derived in Chapter 2 is utilised to determine heat transferred between the primary and secondary sides of each heat exchange component.

Within the system, many vessels can be described as tanks holding and distributing fluid to various parts of the system through the outlets. Some of these tanks do not have flow rate sensors at their outlets, thus determining suitable equations for calculating these values is necessary. These tanks include two preheaters, also referred to as shell-in-tube

heat exchangers with a constant effective area and vertically orientated. The calculations for these tanks are quite straightforward:

Two tanks that are horizontally orientated with end caps are more involved in calculating. A representation of these tanks can be seen in Figure 4.9. The equations for the volumes of these tanks are derived from torispherical tanks as this is deemed a close approximation to the actual components. These tanks can be modelled as a horizontal tank with a torispherical head [50, 57]. The two tanks have similar geometries, with different length values.

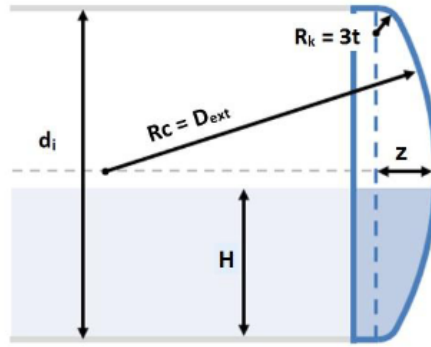


Figure 4.9: Horizontally orientated tank with torispherical head from [8]

Height of the water level measured in the tank is related to the volume of the tank through (4.2) .

$$V_p = d_i^3 C \frac{\pi}{12} \left[3 \left(\frac{H}{d_i} \right)^2 - 2 \left(\frac{H}{d_i} \right)^3 \right] \quad (4.2)$$

where V_p is the partial volume of liquid in the tank, d_i is the inner diameter of the tank, C is a constant obtained from the ASME code standard for torispherical heads and H is the liquid level above the bottom of the tank. To determine the mass flow from the tank output, given the various inputs (4.2) is used in conjunction with (2.1).

4.4.5 THERESA system-attributed graph

The attributed graph is constructed by associating energy flow ($\dot{E}$) and exergy change ($\Delta \dot{B}$) to links and nodes respectively. In the case of the THERESA system, energy and exergy of flow are determined and assigned to links, and exergy change is calculated from the links using exergy balance to determine the exergy change attributes to the nodes. A 21×22 , N_s matrix is constructed from the THERESA system by populating the matrix with the attributes. Figure 4.10 illustrates the final attributed graph used for subsequent sections of the study.

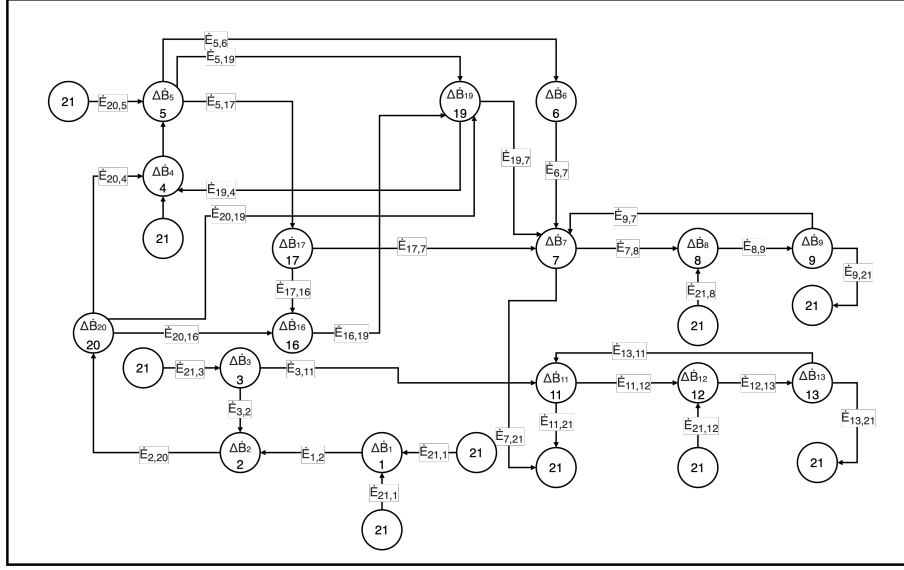


Figure 4.10: THERESA system-attributed graph

4.4.6 THERESA system node signature matrix

The node signature matrix is a matrix representation of the attributed graph in Figure 4.10. With various possible ways of representing the attributed graph in a matrix format, the decision space is where the decision is made on how to order attributes within a matrix. Notation is adopted from Wolmarans [41], with the rationale for the structure of the matrix discussed in Subsection 2.2.5.

The THERESA system node signature matrix is structured to have 22 columns and 21 rows, the energy attributes from the attributed graph are assigned to the relevant positions within the node signature matrix with the first column containing the exergy change across nodes and the subsequent columns containing the energy flow between components as determined by the attributed graph.

$$\mathbf{N}_s = \begin{bmatrix}
 \Delta B_1 & 0 & \dot{E}_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1,21} \\
 \Delta B_2 & -\dot{E}_{2,1} & 0 & -\dot{E}_{2,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{2,20} & 0 \\
 \Delta B_3 & 0 & \dot{E}_{3,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{3,11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{3,21} \\
 \Delta B_4 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{4,5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{4,19} & -\dot{E}_{4,20} \\
 \Delta B_5 & 0 & 0 & 0 & -\dot{E}_{5,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{5,17} & \dot{E}_{5,19} & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \Delta B_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{7,8} & -\dot{E}_{7,9} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{7,21} \\
 \Delta B_8 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{8,7} & 0 & \dot{E}_{8,9} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{8,21} \\
 \Delta B_9 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{9,7} & -\dot{E}_{9,8} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{9,21} \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \Delta B_{11} & 0 & 0 & 0 & -\dot{E}_{11,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{11,12} & -\dot{E}_{11,13} & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{11,21} \\
 \Delta B_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{12,11} & 0 & \dot{E}_{12,13} & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{12,21} \\
 \Delta B_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{13,11} & -\dot{E}_{13,12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{13,21} \\
 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 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \Delta B_{16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{16,17} & 0 & \dot{E}_{16,19} & -\dot{E}_{16,20} & 0 \\
 \Delta B_{17} & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{17,5} & 0 & \dot{E}_{17,7} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{17,16} & 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 \\
 \Delta B_{19} & 0 & 0 & 0 & 0 & \dot{E}_{19,4} & -\dot{E}_{19,5} & 0 & \dot{E}_{19,7} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{19,16} & 0 & 0 & 0 & -\dot{E}_{19,20} \\
 \Delta B_{20} & 0 & -\dot{E}_{20,2} & 0 & \dot{E}_{20,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{20,16} & 0 & 0 & \dot{E}_{20,19} & 0 \\
 0 & \dot{E}_{21,1} & 0 & \dot{E}_{21,3} & \dot{E}_{21,4} & 0 & 0 & -\dot{E}_{21,7} & \dot{E}_{21,8} & -\dot{E}_{21,9} & 0 & -\dot{E}_{21,11} & \dot{E}_{21,12} & -\dot{E}_{21,13} & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{bmatrix} \quad (4.3)$$

4.4.7 THERESA system node signature matrices

The node signature matrix data sets are constructed by populating a node signature matrix for each instance in a process with the relevant energy and exergy attributes, with a node

signature matrix representing a snapshot of energy and exergy attributes at a given point in time.

In this study the node signature data sets produced are data sets A, B and C. Using a MATLAB[®] script where all the above mentioned information is taken into account the program takes as input process variable data and as output generates time series matrix data to be used in subsequent chapters.

4.5 Conclusion

This chapter outlines the method followed to obtain the THERESA system-attributed graph and the node signature matrix of the THERESA system.

The class space forms part of the data analysis and is incorporated in this chapter for continuity. The class space deals with the analysis of the data obtained from this chapter. This data analysis will be explained in Chapter [6](#), where the EGBV analysis will be done, applying various methods that are part of the class space.

Chapter 5

Verification of energy characterisation

5.1 Introduction

This chapter addresses the verification of the energy characterisation in Chapter 4. Chapter 2 provides a review of the literature supporting EGBV as an energy characterisation tool, describing the methodology and research within the scope of this study. Additionally, Chapter 3 furnishes essential details on the THERESA system for implementing EGBV. Thus, the application of the energy characterisation in Chapter 4 should be verified to establish confidence in this application.

Verification ensures that values obtained within the node signature data are within expectations as evaluated according to defined criteria. Through this process, this study evaluates the node signature results for the EGBV methodology applied within the THERESA system context.

For this study, a process model is not available. Therefore, the verification of the applied method is applied concerning the available data, literature, and process knowledge.

The verification criteria for the application of EGBV on the THERESA system are based on agreed specifications and underlying assumptions necessary to achieve the deliverables specified in this study.

5.2 EGBV method verification criteria

Criterion 1: Does the application of the techniques and processes produce the results expected by the EGBV method applied on the THERESA system, based on previous research on the application of the method in the literature?

Based on the outcomes achieved and the methods applied in the EGBV literature, does this study produce the desired results?

Criterion 2: Where calculations and assumptions are used to determine unknown process variables, are the resulting values a sufficient approximation to the expected values?

Through the process of verification and validation, discrepancies are detected that require further debugging or inaccurate assumptions that require significant model modifications and further model re-verification. Thus this section is carried out constantly and iteratively even though these iterations are not depicted within the final work.

This criterion refers to the calculation of initially unknown process variables necessary to the application of the EGBV method on the THERESA system.

Several stages of the application process allow for the evaluation of a sound application of EGBV in a qualitative or quantitative evaluation. These stages include: Handling of data, development of assumptions for analysis, calculations for unknown process variables, calculation of energy and exergy and determination of node signature matrices.

Verification starts with acquired system data scrutinisation to ensure a thorough understanding. For raw system process data, the scrutinisation process includes determining the type of measurement taken, the location of the measurement, the unit the measurement is taken in and other relevant information about the sensor.

After scrutinising the data, system information and expert knowledge are utilised to interpret the data. This is referred to as the development of assumptions for analysis and also includes heuristic principles that are inherent to the EGBV methodology and applied to areas of uncertainty.

Using the necessary assumptions, the EGBV methodology is applied to the point where necessary unknown process variables are calculated. Here, the verification of the output parameters may be influenced by the necessary assumptions made, thus a quantitative verification accompanied by a qualitative discussion is utilised.

Further, the data and the calculated process variables are utilised to determine the necessary energy and exergy values. At this point verification methods similar to the methods applied to the calculation of unknown variables are utilised.

Finally, the output node signature matrices are verified using energy balance, and thus the verification process is completed and finalised for the application of EGBV on the THERESA system. Qualitative analysis based on judgment is utilised based on the expected values and behaviours of the system. This study aims to approximate these values to a point where usable data should be produced that approximates the behaviour of the actual system.

5.2.1 Process knowledge

The implications of the assumptions made are not easily quantified due to the lack of a model to verify the results by comparison with the outputs of a process model. In cases where operator process knowledge is used for assumptions made on the functioning of components and the system, this information is utilised as factual. However, in some cases, process knowledge may be limited to the full understanding of the actual functioning of the system; thus, at points where there is assumed to be zero flow rate, a finite flow rate may still be present causing a 'leak' of energy from certain components that are not easily determined.

Due to the vast complexity of the THERESA system, the documentation to be reviewed for such a system is tedious and may contain information that may be contradictory to operator knowledge or other documentation. An example may be that the system documents may indicate that a sensor measures in L/h. Where this may once have been factual, operator knowledge states that the sensor has been changed and the sensor now measures in kg/h.

5.2.2 Sensor placement

In the THERESA system measurements are taken continuously every second of an experiment and are taken mainly within components (such as the CPDS) or on flow paths such as on the inlets and outlets of components.

Detailed schematics are not included in this study, however the reduced schematic in Figure 4.2 gives context to the number of flow paths. And Table 4.4 indicates points where sensors or measurements are not available. All other positions within the system where pressure, temperature, fluid level or fluid flow rate is required is available. Sensor placement is exactly representative of an actual position within the system. Confidence in expert knowledge allows the assumption that all chosen sensors are within the correct placement as indicated by the system documentation. However, if a single sensor is not within the specified position, this may have negative results on the final output.

5.2.3 Ideal heat transfer and power transfer assumptions

In the application of the energy characterisation for mass balance and energy balance, it is evident that a system without losses is ideal and a real-world system such as the THERESA system, the system is not ideal. However, within thermodynamic modelling, it is possible to assume ideal conditions of energy and mass transfer to simplify the calculation of necessary parameters with limited information.

5.3 Assessment of underlying assumptions

In Table 4.4 the process variables that are calculated to be evaluated based on criterion 2 are outlined and the resolutions are determined. The process following this implementation of resolutions involves determining the necessary equations to ensure that this resolution is applied, and the results from these equations are to be evaluated.

The resolutions to the uncertainties are evaluated while considering the underlying assumptions. These assumptions applied to a single point affect the overall functioning of the system, thus in the holistic view of the system, certain parameters may be higher or lower than the expected values.

5.3.1 Verification by mass and energy balance

Mass and energy balances are qualitatively and quantitatively used as a verification method. The quantitative analysis refers to numerical results that can be interpreted on the basis of the first law of thermodynamics, referred to as the law of conservation of energy.

The general energy balance equation for a generic control volume can be expressed as

$$E - W = \Delta U_{\text{system}} \quad (5.1)$$

With E the heat added to the system, W is the work done by the system on its surroundings and ΔU_{system} the change in the total energy of the system, including the internal, kinetic, and potential energies.

Rewriting this equation into an equation relating the energy flow rate into a component, the energy flow rate out of a component, and the rate of energy change within a system, the following equation results.

$$\dot{E}_{in} - \dot{E}_{out} = \Delta \dot{E}_{\text{system}} \quad (5.2)$$

This energy balance equation aids in the evaluation of energy that is transformed and stored within a system, reflecting the impacts of heat transfer, work interactions, and the accumulation of energy within components.

Therefore, from (5.2), if $\Delta \dot{E}_{\text{system}} \approx 0$ the energy in and the energy out of a component are equal, or approximately equal. If $\Delta \dot{E}_{\text{system}} \neq 0$ there is a difference between the energy in and out of the component, and the difference is attributed to changes within the system, such as heat transfer and energy accumulation. The components within the THERESA system can be evaluated using (5.2).

In Table 5.1, each component is evaluated for the average energy flow rates in and out of the component. With (5.2) the energy balance column is computed. Further, $\Delta \dot{E}_{system}$ is represented as a percentage of the maximum value between $|\dot{E}_{in}|$ and $|\dot{E}_{out}|$.

The component type column indicates the possible rationale behind a $\Delta \dot{E}_{system}$ %. Accumulating components include the shell side of a heat exchanger, the steam generator, the superheater, tanks, and the CPDS.

Non-accumulating components include pumps, tube-side of heat exchangers, and 3-way valves.

The quantitative analysis of the energy within the system is represented in Table 5.1 showing the absolute value of the average energy values at each point in the THERESA system for this study. Based on this information, a qualitative analysis can be used to discuss the results of the energy characterisation.

From the analysis, an approximate energy balance can be formulated between the nodes to evaluate these values using energy conservation. This is approximate as the average of these values does not represent the changes within the time-series data. However, the results are sufficient for visualisation of the following qualitative analysis.

In the context of this study, assumptions underlying the analysis and the application of process knowledge as the basis for key decisions may influence the results from such an energy or mass balance, thus qualitative reasoning is utilised to ensure a thorough understanding of possible discrepancies.

As seen in Table 5.1, particular nodes have energy values that are equal to zero. These nodes are considered part of the system graph, however, they are not currently active within the system.

An energy balance across nodes 1, 2, 3, 5, 8, 9, 11, 12, 13, 16 and 17 yields an energy balance of approximately zero, indicating that the calculated process variables within these nodes are as expected. These components are at a position within the system and in connection to components where flow paths are unambiguous and relatively straightforward to determine.

In cases where nodes represent pumps these nodes are evaluated on the power delivered to the pump being approximate to the specifications of the pumps: this is the case for node 1, where the average power delivered to the pump is 1.27 kWp per the pump size as expected. For pumps represented by nodes 8 and 12, the average power delivered to them is approximately 72 Wp and 53 Wp; this is also within the expected range for these two pumps.

For nodes located within parts of the system where flow paths require tedious review and expert knowledge, the energy balance does not yield the expected values of around zero. For nodes 4, 7, 19 and 20 the respective energy balance equations yield values of around

Table 5.1: Energy balance for data set A

Node number	Component type	$\dot{E}_{in}$ [kW]	$\dot{E}_{out}$ [kW]	$\Delta\dot{E}_{system}$ [kW]	$\Delta\dot{E}_{system}\%$
1	Pump	19.131	19.230	0.100	0.52%
2	Tube-side of heat exchanger	273.871	286.281	12.411	4.34%
3	Shell side of heat exchanger	298.181	428.988	130.807	30.49%
4	Steam generator	261.747	215.401	-46.345	17.71%
5	Superheater	202.324	208.555	6.231	2.99%
6	Component Inactive				
7	Tank	543.520	406.486	-137.035	25.21%
8	Pump	178.850	178.850	0.000	0.00%
9	Tube side of heat exchanger	178.850	169.089	-9.761	5.46%
10	Component represented as part of the environmental node 21				
11	Tank	86.262	123.911	37.649	30.38%
12	Pump	86.916	86.916	0.000	0.00%
13	Tube side of heat exchanger	86.916	46.475	-40.440	46.53%
14	Component represented as part of the environmental node 21				
15	Component represented as part of the environmental node 21				
16	Tube side of heat exchanger	431.275	426.047	-5.228	1.21%
17	Shell side of heat exchanger	208.250	281.388	73.138	25.99%
18	Component represented as part of node 19				
19	CPDS	489.198	588.646	99.448	16.89%
20	3-way valve	286.281	348.839	62.558	17.93%

–81, 151, 273, and 68 kW respectively. This can be explained through a qualitative analysis that takes into account the assumptions made.

For node 4, the boiler, the key assumption that is made is that the boiler is currently not supplied with electricity, thus the energy output of the boiler should be less. If this is the case, then the second assumption that certain flow paths within the system are not being utilised might not be accurate, and some energy is flowing from another component to node 4, for which the flow is not accounted for. For tanks 7 and 11, two assumptions may have an impact on the energy balance. These tanks accumulate high-temperature water; however, the energy balance across an accumulator does not equal zero unless the component returns to the initial state. The same rationale is used to describe the discrepancy in the energy balance of node 3. Also, losses are not taken into account which may be significant within tanks of such large size.

For node 19 which is the CPDS and MP, the assumptions made throughout all components and flows to and from this component may cause the energy balance to be obscured from the expected balance of zero. An assumption that is taken into account within this component is that zero losses occur through heat transfer, which may be accepted as the component is well insulated. However, the second assumption is that all flow paths considered sum up all the energy flows throughout the component across all operational conditions. Based on the knowledge used within this study and for this study, this assumption is deemed valid; however, when considering the quantitative analysis of the model and the energy and exergy balances, this assumption is the first to consider should further research be conducted utilising the exact methodology followed within this research.

5.3.2 Data availability and process knowledge limitations

The THERESA system has sensors that are selectively placed within the system. Utilising energy as a unifying domain requires a minimum amount of process variables to be known to calculate energy and exergy. Within this system, there are areas where all required process variables are measured by dedicated sensors. This provides the necessary process variables for straightforward energy and exergy calculations. Conversely, there are areas where all required process variables are not measured by dedicated sensors and it is necessary to calculate those variables either from known variables at other points in the system or utilising assumptions from process knowledge that leads to a simplification of calculations.

It is necessary to ensure a thorough understanding of the system measurements and operation to make informed decisions on such assumptions and calculations, however, without a verified process model for verification of such calculations there are fundamental principles that need to be followed to ensure such calculations are thorough, accurate, and reliable.

5.3.3 Data handling

The first step is to ensure that good practice is followed in the data handling procedure. In this case, many possible complications may arise if the process raw data are not converted to the appropriate units when utilised in calculations. It is necessary to ensure that verification is done continuously at various points throughout the modelling and application processes.

The process measurement data acquired for use in this study includes measurements of fluid flow. This includes sensors measuring fluid mass flow rate ($\dot{m}$) and fluid volume flow rate (Q). The formulation of the energy and exergy equations requires the fluid flow to be used in kg/s. Where applicable all flow data are evaluated to be converted to kg/s, this ensures that all operations to follow are consistent in the calculation of energy and exergy values. Further, temperature data are measured in °C and where applicable converted to K by the addition of 273.15.

In terms of fluid level measurements in tanks, there are two main methods of measurement. The first unit is the percentage (%) of fluid about the maximum limit of the tank. Within these calculations, it is necessary to ensure the geometry of the different vessels is obtained. The other method is the vertical height of the fluid column within a tank in cm. These measurements are utilised in the process variable modelling of the flows from various vessels and are converted to the relevant units through the use of the tank geometries, or used in the necessary unit within each tank.

Electrical power measurements are calculated from $\dot{m}$ and enthalpy for each component as described in Chapter 2, thus the units are maintained consistently as kW. Pressure measurements are measured in bar or mbar and for consistency converted to bar.

5.4 Conclusion

In conclusion, the values presented in Table 5.1 in conjunction with the motivation for these values based on the process knowledge provide sufficient confidence that the method applied offers a reasonable approximation of the expected results. The verification and validation of the energy characterisation were conducted using a methodology comprising two key components: the establishment of verification criteria and the evaluation of underlying assumptions. The criteria for verifying the EGBV method applied on the THERESA system were assessed based on two aspects: first, whether the outputs align with the expected behaviour derived from known system parameters; and second, whether outputs dependent on calculated values sufficiently approximate the expected outcomes. These criteria were evaluated and found to be sufficiently met through a detailed examination of process information and system operation, as well as an energy balance across components.

Chapter 6

EGBV analysis

6.1 Introduction

Following the application of the EGBV method on the THERESA system, this chapter explores analytical methods for evaluating the node signature matrices composed in Chapter 4. The methods investigated include the fixed reference DP approach and the SBS SVD approach. Two data sets are used for the analysis and discussion of the node signature matrices. Subsequently, a single data set is employed to conduct a sensitivity analysis, assessing the performance of each method in terms of a fixed threshold AR.

6.2 EGBV analysis methodology

Based on the literature discussed in Chapter 2, it is clear that the EGBV method is inherently sensitive to changes in the node signature matrix. The node signature matrix at a specific time contains real-time step information about the state of the system at that moment. Therefore, when analysing the full time-series data of a system during operation, the node signature matrix can be used to identify anomalies through comparison with another matrix. These anomalies may indicate faults within the system or changes in system operational state.

The analysis applies two distinct methods of analysis on the same data to facilitate comparisons and ultimately draw conclusions from the results. This is the DP method and SBS SVD method. To compare the two methods of analysis, a metric of comparison and a methodology by which to conduct such comparison is required.

In the context of EGBV and FDI a viable characteristic to compare across the two methods is sensitivity. Sensitivity in this context is defined as the responsiveness or degree of change in the output of a method in response to variations or changes in its inputs. That is, observing the possible sensitivity of the outputs to mode changes within the system.

The sensitivity can be investigated utilising a methodology of a perturbation analysis. Perturbations of the node signature matrix are used to facilitate the sensitivity analysis given the absence of a model to generate new data sets. The necessity of composing a comparative data set against a reference set, coupled with the constraints of using only available data sets, underscores the rationale for this approach. Thus, it is proposed that perturbations of the node signature matrix can adequately generate a data set suitable for comparative analysis, allowing the examination of system responsiveness under varied operational conditions without the need for additional experimental data.

Due to the nature of the two approaches, there are differences in the outputs of each method in terms of the scale of the output data as well as the dimensions of the output data. The DP method has an output in the form of a single vector with rows equal to the number of time intervals of the data. The SBS SVD approach has an output of multiple singular values in the form of vectors of rows equal to the number of time intervals of the data.

The scale of the data are also different, with the DP method output values between 0 and 1, and SBS SVD method output scales differing between singular values, in this case typically between a value of 0 and 25.

Due to the data representing a system's operation, a metric can be introduced by which the methods are compared without converting to a single scale. This metric is known in FDI as alarm rate and can be divided further into false alarm rate (FAR) and true alarm rate (TAR).

The methodology outlines the procedure by which the results of this chapter will be obtained, the data utilised will be discussed in the next section and subsequent sections will detail the methodologies followed for the application of each approach.

6.3 Data utilised

For all subsequent analyses in this section two data sets are used, namely data set A and data set B. The raw process data have been processed into time series node signature data through the method discussed in Chapter 4. These data sets are considered to have no faults present at any time and should represent the data expected from the system during normal operation.

Table 6.1 summarises the operational parameters used during the experiments to acquire data set A and data set B. During operation, these values are the governing parameters, and in the case of data set A and data set B, the only difference in parameters is the CPDS loading flow rate of $1.5 \text{ m}^3/\text{h}$ for data set A compared to $1 \text{ m}^3/\text{h}$ for data set B.

As each operational data set is divided into intervals where the operational point is expected to change due to changes in system mode. The mode of the system at a given time

Table 6.1: Operational parameters for data sets A and B during experimental data acquisition

	Data set A	Data set B
Main component	Storage component	
Experiment type	Stationary loading	
Stationary starting temperature (°C)	140	140
Stationary system pressure (bar)	38	38
Stationary loading temperature (°C)	242	242
loading volume flow (m ³ /h)	1.5	1
Fluid flow description	Turbulent	Turbulent

should should play a role in contextualising the data as the mode of the system may have significant implications for energy flow within the system.

6.3.1 Intervals for modes in data set A and data set B

The modes for each data set is determined and can be represented as intervals for each data set as seen in Table 6.2 which shows the starting points for each interval as identified in Subsection 3.4.3, which are utilised to analyse intervals within subsequent operations for data analysis.

Table 6.2: Data-set A and B approximate mode starting positions

Mode description	Dataset A (start point in seconds)	Dataset B (start point in seconds)
Startup	1	1
Warm-up 1	2191	1648
Warm-up 2	12101	4225
Charge	14916	5758
Hold	16476	8048
Discharge	25569	19673
Shut down	27555	22585

These parameters and intervals hold value within mode and operational point variations within the system associated with normal operational conditions that will be used as reference intervals for identifying changes across modes.

A third data set C is also introduced in this chapter, where the data set undergoes some of the same operations as data sets A and B as preparation for Chapter 8. Data set C is a set containing faults and will be used to determine fault detection and isolation performance of the two methods discussed in this chapter. To avoid redundancy and maintain relevance this set is discussed in detail in Chapter 8 Where the methods are applied for FDI.

6.4 EGBV analysis using fixed reference DP method

The fixed reference DP is a well researched approach used in multiple studies in the EGBV for FDI field. These studies have reported varying results, but has provided key insights in terms of changes in operational point [15]. The implementation of the approach does not have high computational demands, thus is a good starting point for analysis of a data set where it is unknown what to expect.

The methodology for applying the fixed reference DP method can be split into steps:

1. Time series node signature data are processed into time series cost matrices using a single fixed reference produced from steady state operation in the holding mode of each data set.
2. DP calculated from each cost matrix through comparison with the determined node signature matrix of each set.
 - (a) A reference for data set A was produced, all node signature matrices in data set A are compared to the reference to yield time series DP values for this data set.
 - (b) A reference for data set B was produced, all node signature matrices in data set B are compared to the reference to yield time series DP values for this data set.
3. Moving average filter applied to time series DP data.
 - (a) Determine run length L
4. Visualisation of the DP over all intervals for each data set A and B respectively.

The DP is a simple way of evaluating the distance between two matrices by utilising the average value of the diagonal entries of the cost matrix formed when two matrices are compared. This is a result of the cost matrix diagonal holding important information, which is reduced to a single value as a measure of distance. The equation utilised is given by

$$DP = \frac{\sum_{k=1}^n C_{kk}}{n}, \quad (6.1)$$

where kk refers to the relevant diagonal elements in the cost matrix and n represents the number of diagonal elements in the square cost matrix thus, the number of nodes of the system graph. This method has proven useful in literature as a predecessor to eigen-decomposition and SVD methods.

6.4.1 Reference node signature matrix determination

With the DP used in previous studies, a single reference node signature matrix was sufficient within a single data set as no mode changes were present. However, in this study, a decision is made to evaluate reference node signature matrices to gain an understanding of the relationships between each node signature matrix as a reference.

From the literature covered in Chapter 2 it is established, that computing the distance metric among potential node signature matrices across various modes may yield significant insights.

Using the intervals for data set A and data set B as determined from Chapter 2 Section 3.8, the average of the node signature matrices across an interval where the system mode is considered constant can be computed and compared to that of other modes of the same data set using the DP. Thus, for each data set a reference node signature matrix is determined as an average of all the node signatures in a specific mode, then compared to all other node signature matrices with the DP as an indication of the variability between the time series data and the single reference. This is done for each data set.

This comparison does not consider start-up and shut-down modes and the objective of such comparison is to determine the degree of difference between modes.

These results are given in Tables 6.3 and 6.4 with a colour scale of green to red showing low and high values of distance respectively. The average node signature matrix over each interval can be quantitatively compared to the other reference node signature matrices. This measure of distance shows that each reference compared with itself, yields a measure of 0 distance, thus the comparison matrix has zeros on its diagonal.

Further discussing Table 6.3 over all modes it is seen that the discharge mode and warm-up mode 1 yield overall higher values compared with all other modes, where Warm-up 2 mode, Charge mode and Holding mode all have relatively low values when compared to each other.

Table 6.3 shows the values obtained from this comparison for data set A.

Table 6.3: Reference comparison by distance parameter method data set A

	Warm-up 1	Warm-up 2	Charge	Hold	Discharge
Warm-up 1	0.000	0.256	0.271	0.260	0.286
Warm-up 2	0.256	0.000	0.121	0.022	0.269
Charge	0.271	0.121	0.000	0.122	0.328
Hold	0.260	0.022	0.122	0.000	0.277
Discharge	0.286	0.269	0.328	0.277	0.000

Table 6.4 shows the values obtained from this comparison for data set B. The same is true as for data set A in terms of similar modes. However, the degree of difference changes and the overall distribution remains relatively the same.

From these tables, the conclusion is made that those modes with the most similar node

Table 6.4: Reference comparison by distance parameter method data set B

	Warm-up 1	Warm-up 2	Charge	Hold	Discharge
Warm-up 1	0.000	0.176	0.209	0.202	0.253
Warm-up 2	0.176	0.000	0.091	0.060	0.188
Charge	0.209	0.091	0.000	0.114	0.232
Hold	0.202	0.060	0.114	0.000	0.164
Discharge	0.253	0.188	0.232	0.164	0.000

signature matrices when evaluated with the DP are Warm-up 2, Charge, and Hold. Warm-up 1 and Discharge exhibit DP values that indicate significantly different node-signature matrices to the other modes.

Ultimately the comparison matrix indicates that utilising a single average node signature matrix may lead to significantly different results depending on the reference that is chosen. To determine a framework for selecting such a reference for FDI is not the objective of the study and could be a topic of further investigation for future studies.

In this study, the reference node signature matrix is selected as the average node signature over the holding mode for the following reasons:

A large portion of each data set is in the holding mode and this mode is a steady state, thus the average node signature matrix will be more representative of the mode over the full interval. The sensitivity analysis to follow later in the study will best be done in a steady state mode.

6.4.2 Cost matrix calculation - DP method

As part of meeting the defined objectives of this study the use of the DP as a method of acquiring useful information from the node signature data produced from the THERESA system should be demonstrated. However, the investigation into the determination of the optimal reference node signature matrix remains a topic of further study. For this study, a steady-state interval is chosen as the interval over which the reference node signature matrix is chosen, namely, the holding mode, as this mode is considered the best-defined steady-state mode within the system.

For data set A, the reference node signature matrix is determined and all node signature matrices are compared to this reference node signature matrix. The DP is calculated and the data are filtered using the moving average function with an L value of 15.

The cost matrix at a given sample k in a time series set is the comparison of two node signature matrices, using the HEOM function.

$$\mathbf{C}_{s,k} = \mathbf{N}_{ref} \Psi \mathbf{N}_{s,k} \quad (6.2)$$

This is the operation $\mathbf{N}_{ref} \Psi \mathbf{N}_{s,k}$, and is equivalent to $HEOM(\mathbf{N}_{ref}, \mathbf{N}_{s,k})$, where $\mathbf{N}_{ref}$ is used as a reference for the full operational set of $\mathbf{N}_{s,o}$. This produces the cost matrix set

$\{\mathbf{C}_{\mathbf{N}_{ref}\Psi\mathbf{N}_{s,k}}\}_{k=1}^K$ with $k = 1, 2, \dots, K$ samples.

6.4.3 Fixed reference DP method visualisation

The visualisation of the resulting DP values is shown in Figure 6.1. The coloured background of the plot indicates specific intervals observable throughout the process. Green indicates the startup interval, yellow indicates the first warm-up interval, pink indicates the second warm-up interval, magenta indicates the charging phase, cyan indicates the holding interval, grey indicates the discharge interval and red indicates the shut-down interval.

As expected in Figure 6.1, the value of the DP is lowest within the cyan interval or holding mode because this is the interval within which the reference node signature matrix is chosen and the most likely interval for the DP to be zero.

Upon first inspection it is evident that the DP value fluctuates across modes. The DP value reaches areas where the average value is relatively constant with the exception of transient movement at the beginning or end of a mode.

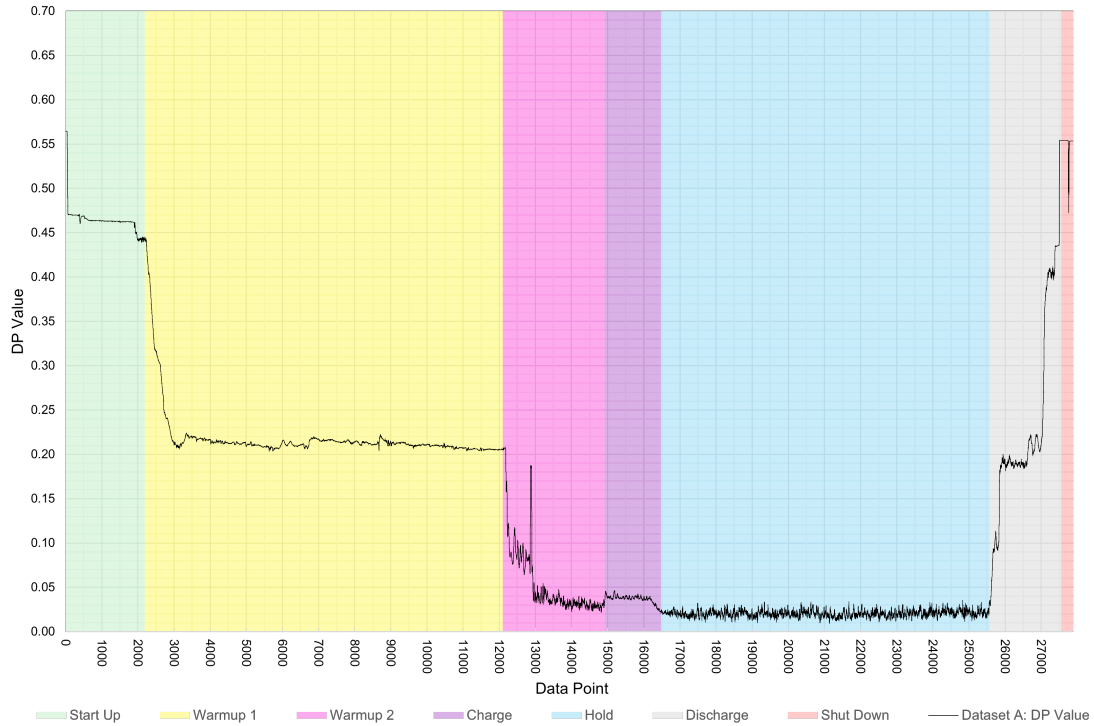


Figure 6.1: DP Approach - distance parameter of data set A visualised over all intervals with fixed reference in holding mode

The same method is followed for Figure 6.2 where it is also evident that within the holding mode the DP is at the lowest.

Comparing the two data sets is difficult as the modes have different interval lengths and

the intervals do not match up.

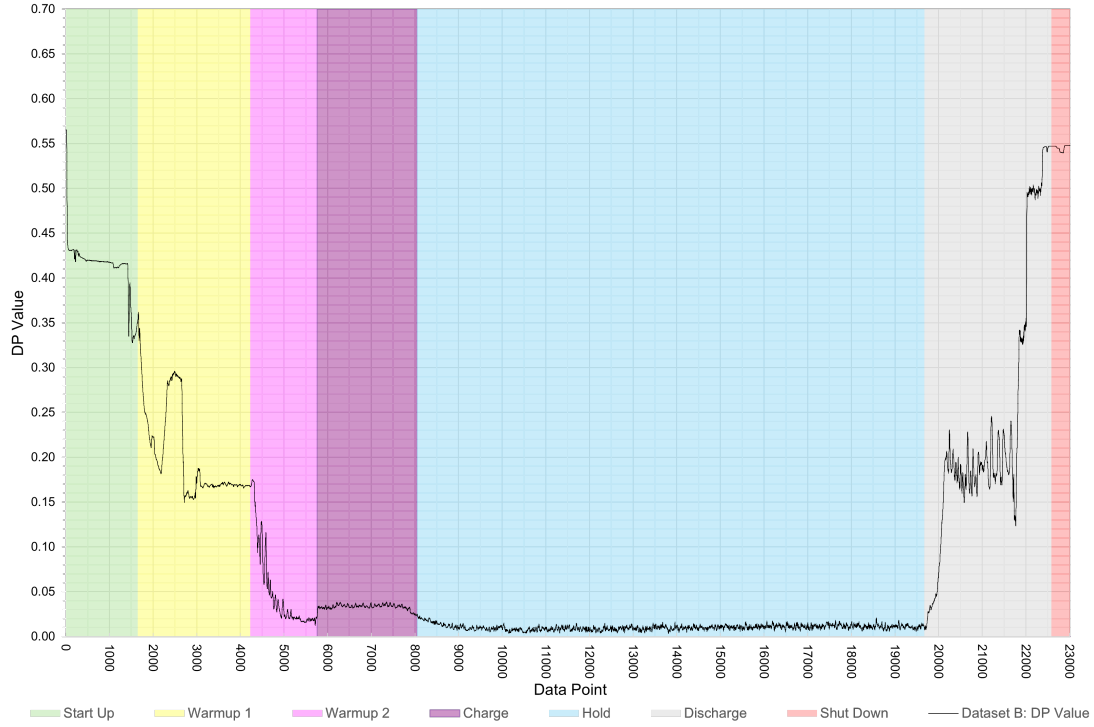


Figure 6.2: DP Approach - distance parameter of data set B visualised over all intervals with fixed reference in holding mode

As an initial observation, the data seems to follow the same pattern or shape within each mode regardless of the time interval. The next part of the investigation is to determine some of the characteristics of the data to determine if these two data sets share common characteristics that can be used in further analysis to extract value.

6.4.4 Characteristics of the DP method

The DP method characteristics observed across all modes are similar to the results observed in the reference node signature comparison tables: Table 6.3 and Table 6.4. Figure 6.3 shows that the DP values for both sets follow the same trend and that the DP value in the holding mode is the lowest compared to all other modes. This is indicative of the reference being chosen in the steady state holding mode.

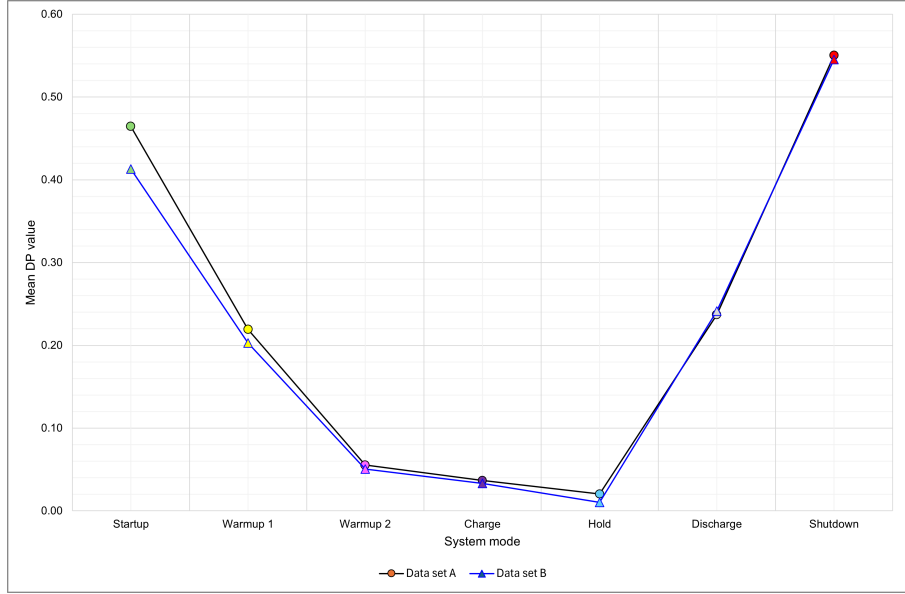


Figure 6.3: Mean DP value for data sets A and B over all modes

Data points where the operational node signature matrix is similar to the reference node signature matrix of the holding mode will result in lower DP values. This is observed in Table 6.5 and graphically represented in Figure 6.3

Table 6.5: Data set A - Mean DP value

	Mean DP Value
Startup	0.464
Warmup 1	0.219
Warmup 2	0.055
Charge	0.037
Hold	0.021
Discharge	0.237
Shutdown	0.550

Data set A and data set B should yield similar results. However, it is expected that the results may be different in magnitude across the set given that the sets were not produced with the same parameters.

Table 6.6: Data set B - Mean DP value

	Mean DP Value
Startup	0.413
Warmup 1	0.203
Warmup 2	0.051
Charge	0.033
Hold	0.010
Discharge	0.241
Shutdown	0.545

6.5 EGBV analysis using the SBS SVD method

The SBS SVD methodology is derived from Wolmarans et al. [58] referred to as sample self comparison or SSC in the context of spectral conditioning in EGBV analysis.

The selection of a fixed reference as in the case of the DP approach used in the previous section can be adapted to use a reference node signature matrix that changes with each sample, in this context the method relies on each point to be used as a reference for itself.

Utilising this method for creating cost matrices allows the decomposition of the cost matrices into singular values whereby the singular values correspond to changes in the system from one point to the next. This will be referred to as the SBS methodology.

For this approach the methodology can be divided into steps:

1. Time series node signature data are processed into time series cost matrices using the SBS method of using a node signature instance as a reference for itself.
2. SVD is applied to the time-series cost matrices to determine singular values for each cost matrix, leading to time-series singular values.
3. Moving average filter applied to time series SV data.
 - (a) Determine run length L for moving average filter
4. The dimensional reduction by singular value (SV) selection leads to fewer singular values being considered
 - (a) Determine cumulative variance percentage (CPV) for singular values
5. Visualisation of chosen singular values over all intervals.
6. Visualisation of chosen singular values over selected intervals.

6.5.1 Cost matrix calculation - SBS SVD method

The cost matrix at a given sample k is the comparison of two node signature matrices, using the HEOM function is given by

$$\mathbf{C}_{s,k} = \mathbf{N}_{s,k} \Psi \mathbf{N}_{s,k}. \quad (6.3)$$

This is the operation $\mathbf{N}_{s,k} \Psi \mathbf{N}_{s,k}$, and is equivalent to $HEOM(\mathbf{N}_{s,k}, \mathbf{N}_{s,k})$, where $\mathbf{N}_{s,k}$ is used as a reference for itself. This produces the cost matrix set $\{\mathbf{C}_{\mathbf{N}_{s,k} \Psi \mathbf{N}_{s,k}}\}_{k=1}^K$ with $k = 1, 2, \dots, K$ samples.

6.5.2 Singular value decomposition - SBS SVD method

SVD analysis is applied to the time series cost matrix calculations which result in time series SVD elements. Within each time series and the given dimensions of the cost matrices, the SVD analysis gives twenty-one singular values for each cost matrix.

Dimensional reduction by singular values

SVD elements are arranged in descending order in terms of the variance represented by each SV. This is a useful property and justifies the reduction of SVD elements to a finite set less than the total number of singular values, as done by Wolmarans [2]. However, a trade-off exists between robustness and sensitivity to the number of singular values retained [41]. Thus, to retain the sensitivity of the system, mindful selection of the extent of reduction is important.

The cumulative percentage variance of singular values offers a threshold-based solution to dimensional reduction. A threshold can be chosen, and based on this choice a sufficient amount of singular values can be used. Figure 6.4 shows a plot of CVP of all data sets available for use in this study.

Dimansionality reduction is facilitated through choosing the least amount of singular values that cumulatively represent the majority of the total variance. Choosing a CVP of 95% provides for three singular values that are sufficient in expressing at least 95% of the total variance for all data sets used within this study. Using this method allows the reduction of the singular values while ensuring sufficient data are preserved within the selected singular values. Thus, the total amount of singular values that will be considered for further analysis is reduced from twenty-one to three, which is a reduction of approximately 85% while still maintaining approximately 95% of all the variance with each data set.

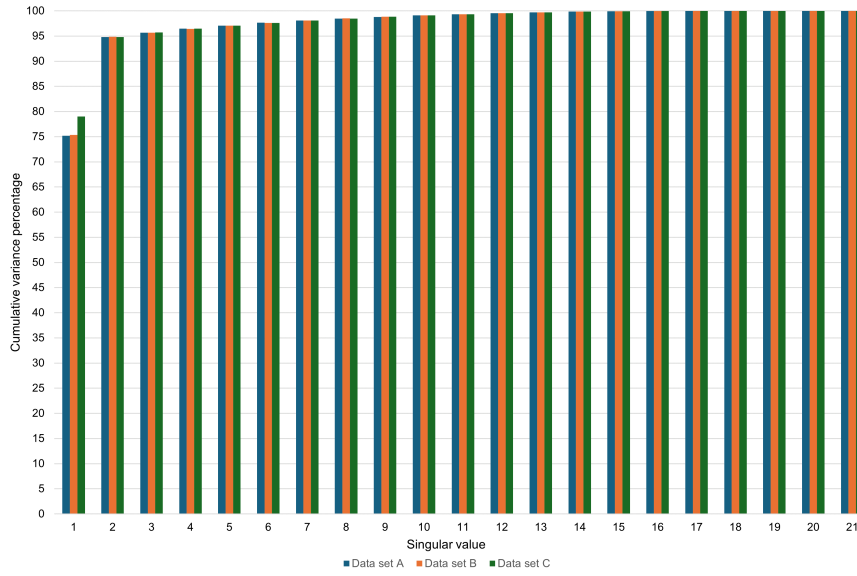


Figure 6.4: Cumulative variance percentage vs number of singular values for data sets A, B and C

Noise reduction for singular values

Inherently, the composition of node signature matrices derived from process data encompasses process noise. This noise can stem from various sources, including sensor disturbances, process variability, and environmental factors.

To address the noise observed in the singular values caused by these disturbances, a filter is applied to the SV data. This study explores two filtering techniques: the exponentially weighted moving average (EWMA) filter and the moving average (`movmean`) filter.

In this study, the moving average filter is employed, requiring precise selection of the window length L . Unlike the EWMA used by Wolmarans [2], which combines the most recent observation with the previous average, the `movmean` function calculates the average across all points within a sliding window. By using the MATLAB[®] `movmean` function, the moving average is computed to smooth out short-term fluctuations and highlight longer-term trends. The equation for calculating the simple moving average is given by

$$y_k = \frac{1}{L} \sum_{i=k-\lfloor \frac{L-1}{2} \rfloor}^{k+\lfloor \frac{L-1}{2} \rfloor} x_i, \quad (6.4)$$

where y_k represents the moving average at position k , x_i denotes the data points within the averaging window and L is the run length which indicates the total number of data points used for each average calculation. The indices $i = k - \lfloor \frac{L-1}{2} \rfloor$ to $i = k + \lfloor \frac{L-1}{2} \rfloor$ define a window centred on the k -th element.

In (6.4) it is assumed that the run length is odd for simplicity. If L is even, adjustments are necessary to maintain symmetry in the averaging window. When k is a positive integer

scalar, the centred average encompasses the current element and its adjacent neighbours. The determination of L is crucial for balancing data smoothing and preserving information in the dataset.

To determine the value of L that adequately filters the data, but does not cause significant data loss, a range of potential run lengths are explored. The evaluation of the SAD value and the value L for the first SV of each data set is shown in Figure 6.5

In this case $L = 2$ to $L = 100$ data points. Each run length within this range is iteratively tested using the moving average filter. The effectiveness of each run length is assessed by calculating the sum of absolute differences (SAD) between the original dataset and its smoothed counterpart. A plot of L vs SAD shows that increasing the run length from $L = 2$ to about $L = 10$ results in rapidly increasing SAD values. However, from $L = 20$ to $L = 100$ the gradient changes and decreases, causing an elbow that defines an interval where smoothing is sufficient and data loss is minimised. Choosing a L value within this window will result in a smoother output and will preserve information in the data.

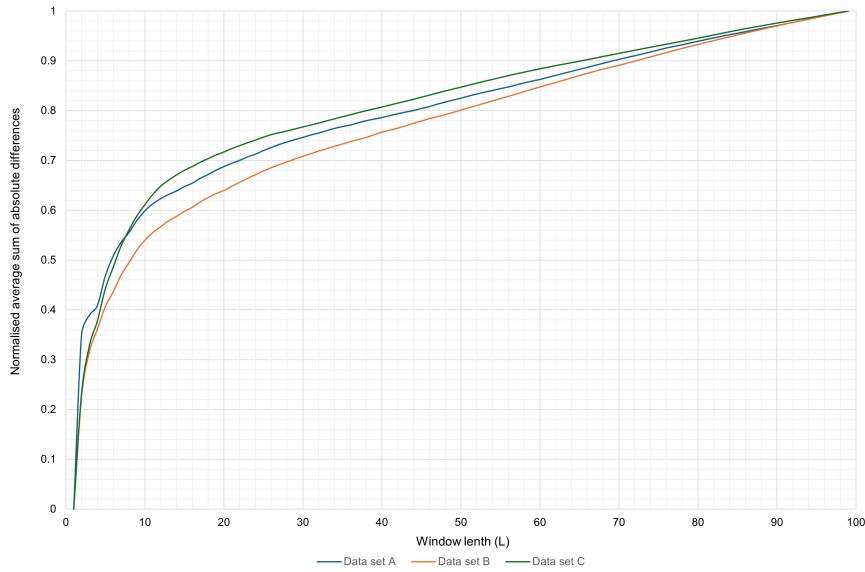


Figure 6.5: Normalised average sum of absolute differences vs window length (L) for data sets A, B and C

The graph of window length to normalised sum of absolute differences show a steep gradient between $L = 1$ and $L = 10$ with a change in gradient occurring between $L = 10$ and $L = 20$ from $L = 20$ to $L = 100$ the gradient is less than that of the first interval. A trade-off exists between smoothing the data and losing information from the filtered data set. In this case, the window length considered is $L = 15$ which is in the gradient transition which indicates that here the sufficient smoothing can be achieved while also minimising data loss compared to choosing a longer window length.

6.5.3 SBS SVD method visualisation

Figure 6.6 shows a combined plot of the first three singular values obtained from the time series SBS evaluation of data set A.

The coloured background of the plot indicates specific intervals observable throughout the process. Green indicates the startup interval, yellow indicates the first warm-up interval, and pink indicates the second warm-up interval. Magenta indicates the charging phase, cyan indicates the holding interval, grey indicates the discharge interval, and red indicates the shut-down interval.

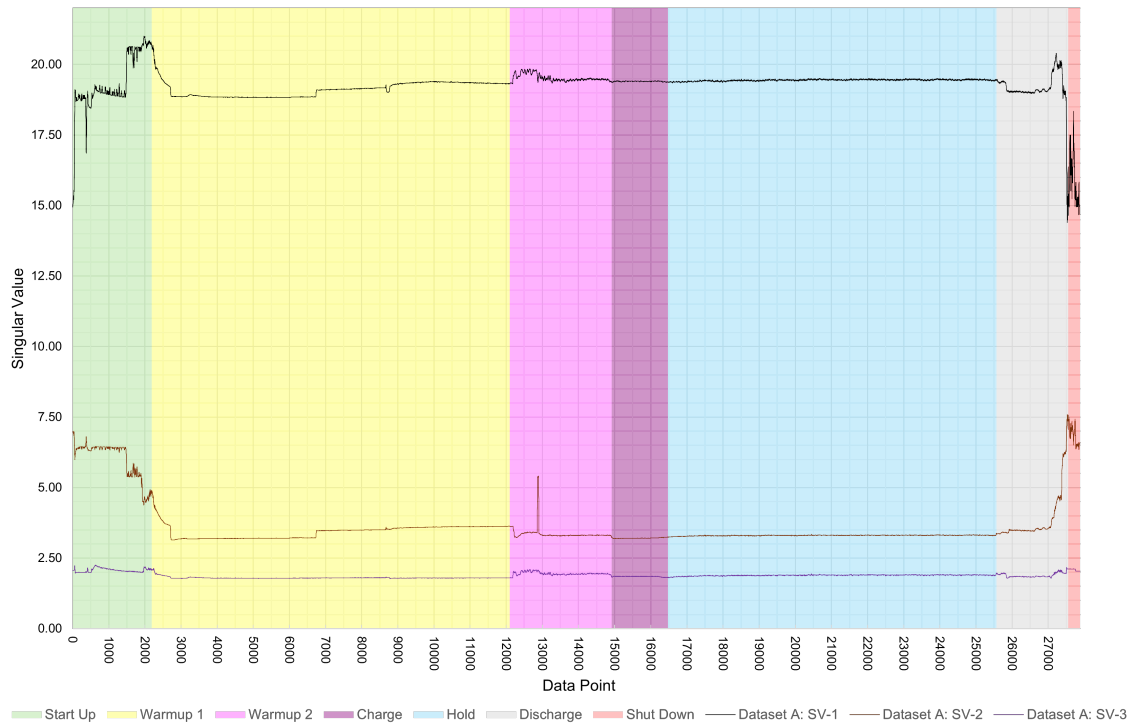


Figure 6.6: SBS Approach - first three singular values of data set A visualised over all intervals, with modes indicated

For data set B, the same visualisation methodology is followed and can be observed in Figure 6.7, a combined plot of the first three singular values obtained from the time series SBS evaluation of data set B. Compared to data set A it is observed that mode interval lengths vary. Initial observations of Figures 6.6 and 6.7 are that the singular values, as expected, are arranged in decreasing magnitude and can be distinguished based on the average magnitude of the singular values. The first SV has the highest value, secondly, the second SV, followed by the third SV, which is consistently the lowest magnitude.

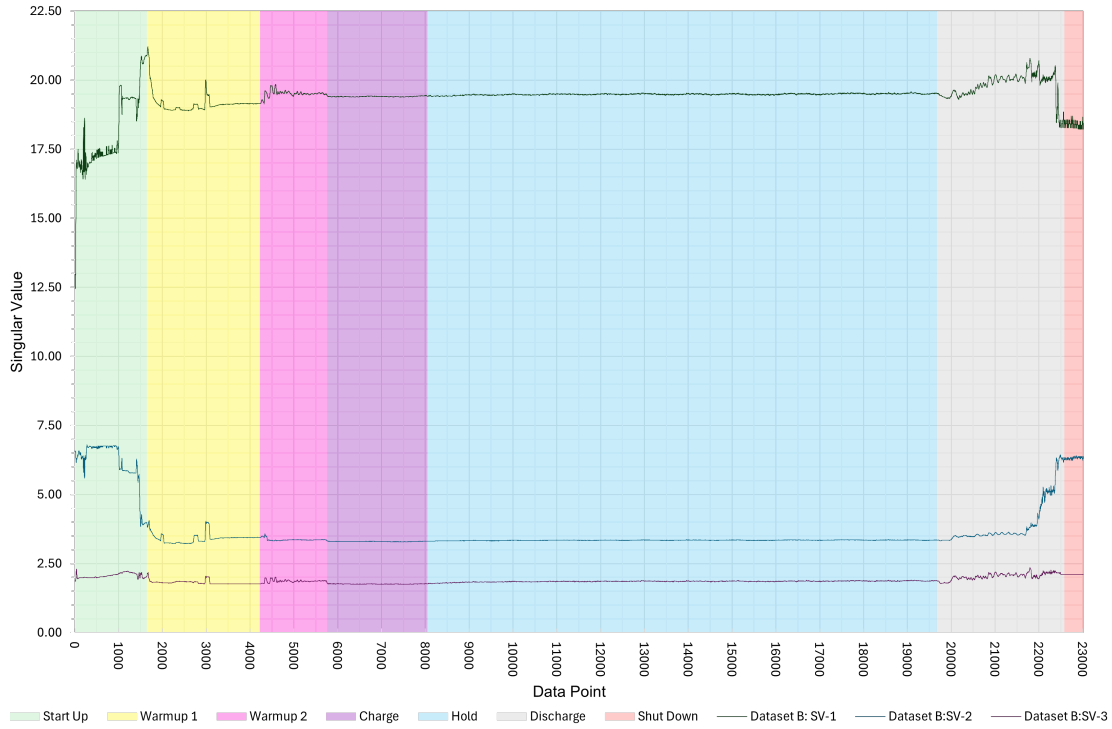


Figure 6.7: SBS Approach - first three singular values of data set B visualised over all intervals, with modes indicated

In large-scale systems, transient behaviour is expected during the startup and shutdown phases. This occurs as certain components need to warm up, reach specific speeds, or attain desired pressures or temperatures. For example, the THERESA system requires components to warm up and pressure to build to a set point during startup. Conversely, during shutdown, the system experiences a significant loss of energy through pressure and temperature decreases. These transient areas can be observed in Figures 6.6 and 6.7 as sharp peaks and valleys in the initial and final intervals of the data.

These areas of transient behaviour is not the only areas where this is observed, as transient system movement can be observed at mode changes. In the case of the start up to warmup 1 mode it is clear that there is a sharp drop off from the end of the startup mode into the warmup 1 mode.

6.5.4 Characteristics of SBS SVD method

Figures 6.8 and 6.9 shows the average value of each SV over each mode for data sets A and B respectively. These graphs show similarities between data sets A and B and could be useful within further analysis or for results interpretation.

To be able to verify if the sets produce similar average values in a given mode Figures 6.10, 6.11 and 6.12 are plotted to show the comparison between data set A and B for each of the singular values.

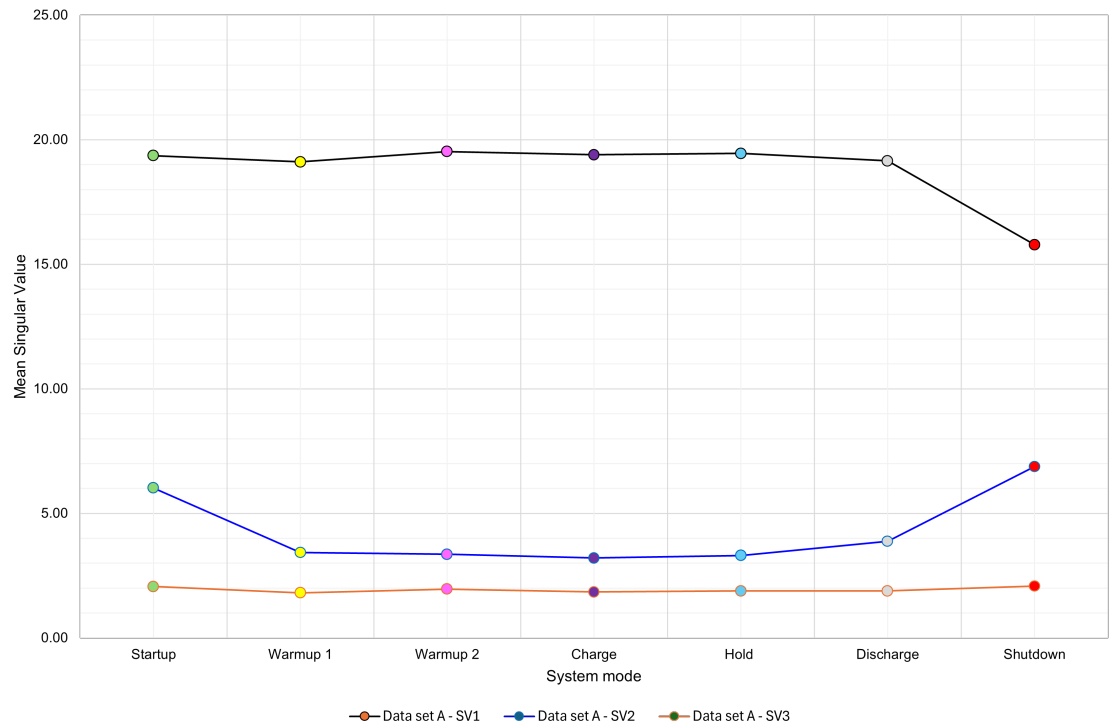


Figure 6.8: Mean of each singular value over all system modes for data set A

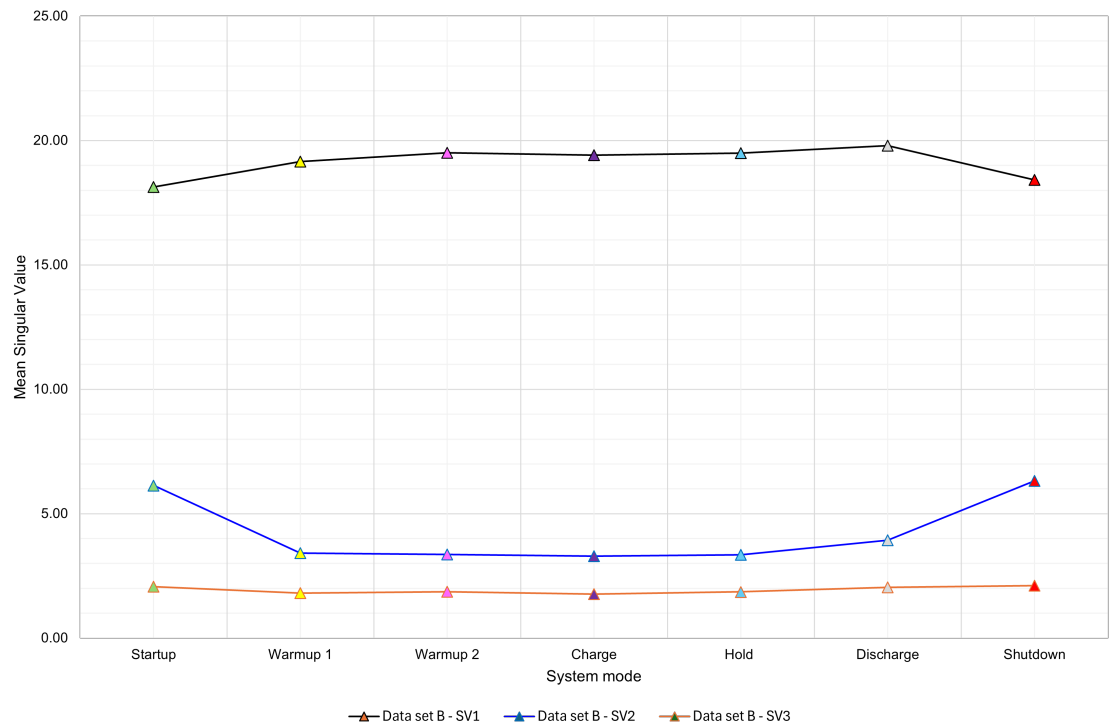


Figure 6.9: Mean of each singular value over all system modes for data set B

Due to scale differences between singular values, comparison across data sets will be done per SV as in Figures [6.10](#), [6.11](#) and [6.12](#).

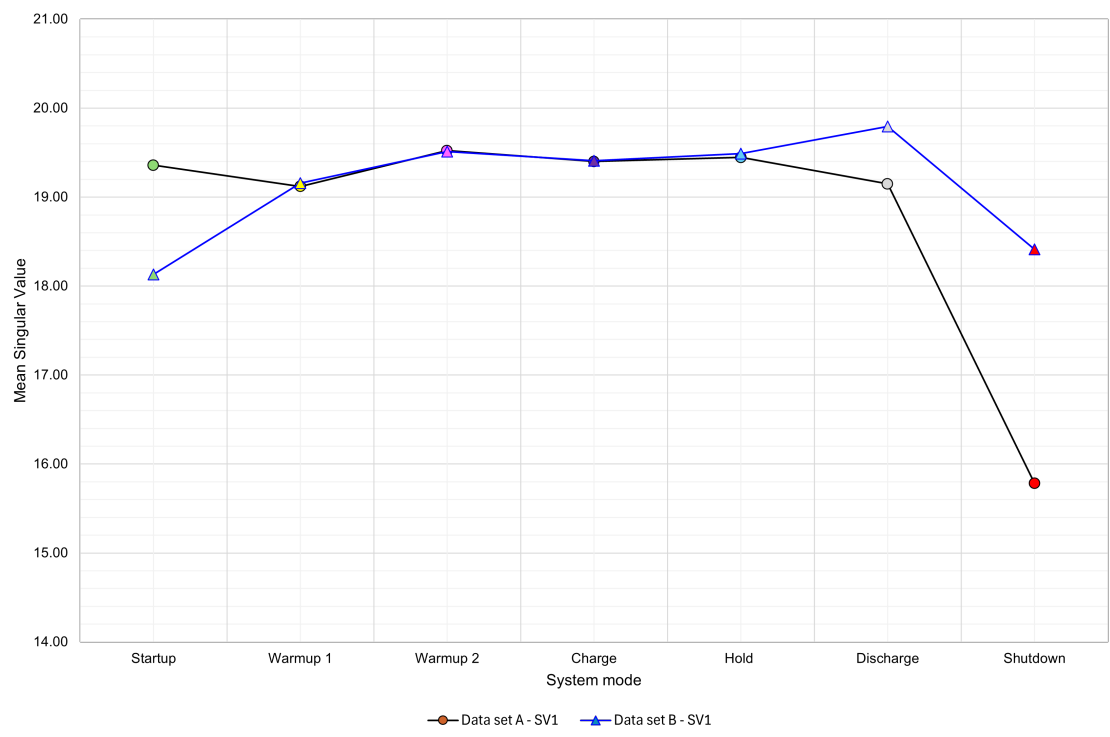


Figure 6.10: Mean value of SV-1 for data sets A and B over all modes

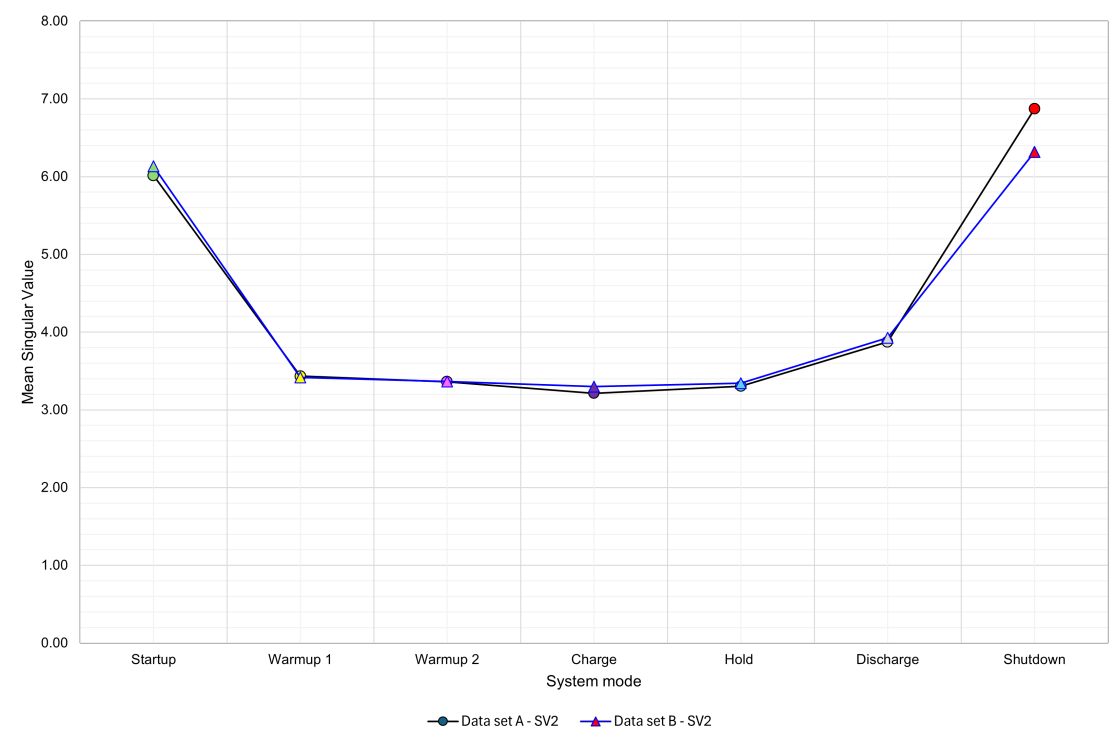


Figure 6.11: Mean value of SV-2 for data sets A and B over all modes

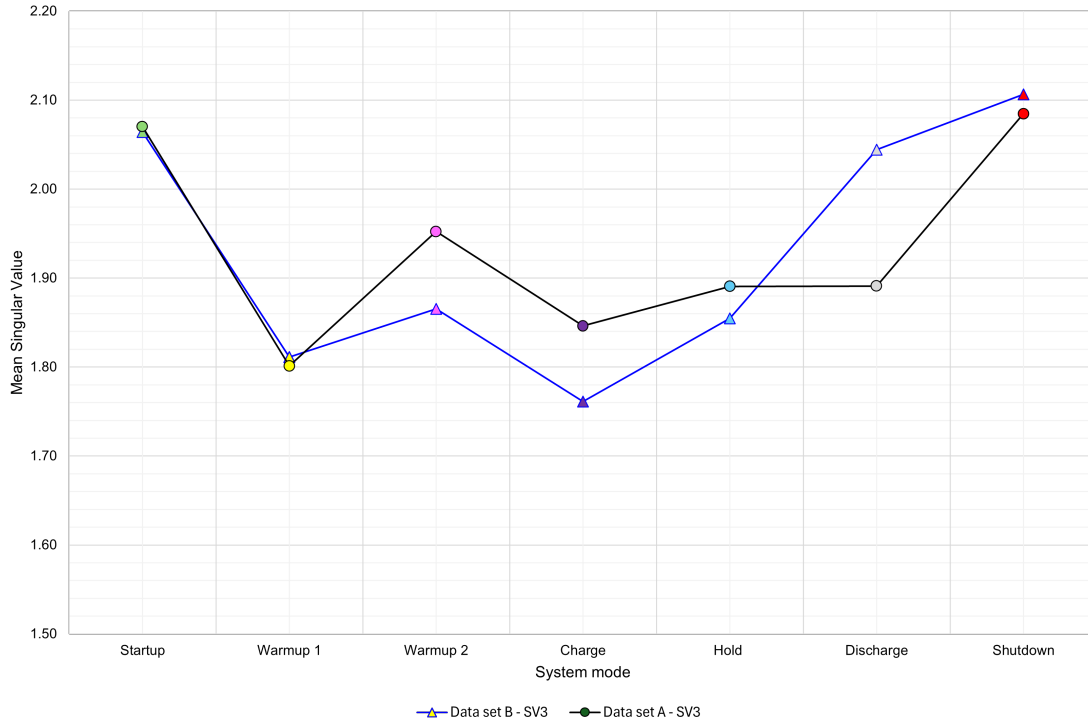


Figure 6.12: Mean value of SV-3 for data sets A and B over all modes

Overall, each SV shows some correlation with the average value across each data set. It is clear that not all points correlate 100%, this is attributed to the operational parameters of the data sets being different and thus, slight differences are expected. It is thus clear that the method is mode dependent and not mode independent meaning that the method should be able to differentiate between set points of a system.

Table 6.7: Average singular value of first 3 singular values over mode intervals for data set A

	SV1	SV2	SV3
Startup	19.36	6.02	2.07
Warmup 1	19.12	3.44	1.80
Warmup 2	19.53	3.36	1.95
Charge	19.40	3.21	1.85
Hold	19.45	3.30	1.89
Discharge	19.15	3.88	1.89
Shutdown	15.78	6.87	2.08

The mean value of the first SV is substantially higher than that of the second and third singular values. It is, however, inferred that this is a result of a system variance represented by the first SV being substantially more important in the overall system variability than that of subsequent singular values.

Table 6.8: Average singular value of first 3 singular values over mode intervals for data set B

	SV1	SV2	SV3
Startup	18.13	6.13	2.06
Warmup 1	19.16	3.42	1.81
Warmup 2	19.51	3.37	1.87
Charge	19.41	3.30	1.76
Hold	19.49	3.34	1.85
Discharge	19.80	3.93	2.04
Shutdown	18.41	6.32	2.11

6.6 Sensitivity analysis by node-signature perturbation

Sensitivity is regarded as a key FDI performance parameter and is often discussed in conjunction with robustness [3, 41]. Sensitivity of an analysis approach or FDI method is in simple terms how well a method detects small changes in system parameters. The sensitivity of the system can be deduced from the evaluation of the FDI method output compared to known faults within the input data [3].

Thresholds for upper and lower limits of the FDI method output are set up using the normal operational data, fault data are then evaluated using the normal data thresholds, and the performance of the method is evaluated based on how well faults can be isolated in terms of changes in fault magnitudes.

Ideally, if it were feasible to conduct a range of experiments with the THERESA System under controlled experimental parameters, the sensitivity analysis could directly assess how variations in process variables influence the singular values of the process. Similarly, a system process model can produce data that approximates the system's behaviour based on defined experimental parameters. These methods are combined by Smith [3]. However, in the absence of a process model, this study's sensitivity analysis must rely on existing data.

The objective is to create multiple perturbed data sets from the original normal data set by manipulating certain entries within node signature matrices. This is similar to the approach by Smith [3] where fault sets are evaluated. However, instead of evaluating for specific intervals a performance parameter is chosen that will be evaluated. Matrix perturbation and analysis by SVD is employed by Turan et al. [59], and is a well-researched FDI method in EGBV, as singular values indicate changes within a matrix over time steps.

The rationale for focusing perturbations on the node signature matrix rather than directly on existing process variables is to simplify the analysis from changes in system variables to changes in energy levels. Consequently, perturbing a single element within the node signature matrix can simulate a specified percentage change in a component's energy value,

instead of altering multiple process variables (P , T , $\dot{m}$) for the same energy component. This method allows for an examination of the sensitivity of the SBS method to changes within the node signature matrix, providing insights into the system's responsiveness to perturbations.

6.6.1 Sensitivity analysis methodology

A sensitivity analysis is conducted to evaluate the responsiveness of the two methods to changes of varying magnitude in the node signature matrices. This analysis is implemented utilising a single data set to produce multiple newly perturbed data sets. The data set selected is a steady state interval within data set A and by following the method below, 21 new data sets are produced, each equivalent in length to the original set. Each set will represent a specific perturbation of a group of entries within each node signature matrix at a given time step.

The notation for indicating the dimensionality of a node signature matrix is, $\mathbf{N}_s \in \mathbb{R}^{n \times n+1}$ [41]. Now, let $\{\mathbf{N}_{s,k}\}_{k=1}^K$ be a set of $n \times (n+1)$ matrices at samples $k = 1, 2, \dots, K$. The element at position $\mathbf{N}_{s,k}(i, j)$ is defined for $1 \leq i \leq n$ and $1 \leq j \leq n+1$. Suppose a perturbation is applied to the element at position (i, j) across all time steps by increasing its value using a perturbation factor (PF). The perturbed element can then be expressed as

$$\tilde{\mathbf{N}}_{s,k}(i, j) = \mathbf{N}_{s,k}(i, j) \times \text{PF}. \quad (6.5)$$

This modification is applied at each sample k , resulting in a new set of matrices $\{\tilde{\mathbf{N}}_{s,k}\}_{k=1}^K$.

Utilising this methodology, and taking into account how each node signature matrix is composed, it is possible to infer perturbations due to the relation between different elements within the node signature matrix. The perturbation of multiple elements within each node signature matrix is inferred to ensure that perturbations take into account the relations between different elements in the node signature matrix. For example, if the energy changes from node a to node b the energy from node b to node a will also change. Similarly, the corresponding exergy values within the first column of the node signature matrix may also change. Exergy values can be considered to be dependent on the energy values, thus if at a given sample k a node signature matrix $\mathbf{N}_{s,k}$ is perturbed, then the target entry $\mathbf{N}_{s,k}(i, j)$ is perturbed, with the other entries that are dependent on that energy or exergy value, which means that four elements are perturbed simultaneously.

Firstly, the new perturbed matrix is set equal to the unperturbed matrix $\{\mathbf{N}_{s,k}\}_{k=1}^K = \{\tilde{\mathbf{N}}_{s,k}\}_{k=1}^K$, the perturbed matrix is consistently perturbed over each time interval at four positions: the target element (i, j) , the exergy value in the same row as the target element $(i, 1)$, the opposite energy value to the target element $(j-1, i+1)$, and the exergy value in the same row as the opposite energy value of the target element $(j-1, 1)$. Thus, the operations required to determine the i -th row entry and j -th column entry for the perturbed matrix is given as

$$\tilde{\mathbf{N}}_{s,k}(i, j) = \mathbf{N}_{s,k}(i, j) \times PF, \quad (6.6)$$

$$\tilde{\mathbf{N}}_{s,k}(i, 1) = \mathbf{N}_{s,k}(i, 1) \times PF, \quad (6.7)$$

$$\tilde{\mathbf{N}}_{s,k}(j - 1, i + 1) = \mathbf{N}_{s,k}(j - 1, i + 1) \times PF, \quad (6.8)$$

and,

$$\tilde{\mathbf{N}}_{s,k}(j - 1, 1) = \mathbf{N}_{s,k}(j - 1, 1) \times PF. \quad (6.9)$$

In these equations PF denotes a constant of perturbation referred to as the perturbation factor. The unperturbed data are equal to a $PF = 1$, for subsequent perturbed data $PF = 1.1$, $PF = 1.2$ and $PF = 1.3$ for three magnitudes of perturbation.

For each PF , when performed over all time instances k , a new set $\{\tilde{\mathbf{N}}_{s,k}\}_{k=1}^K$ can be constructed. When done over all entries, $21 \times 22 = 462$ new data sets are formed from the original unperturbed set. When done over three perturbation factors the total number of new data sets is 1386.

A perturbed set differs from the normal set as a result of the perturbation applied, thus FDI methods for sensitivity analysis can be utilised comparing the new 'fault' set to the original 'normal' set.

Further analysis is primarily based on the two methods, SBS SVD analysis and fixed-reference DP method.

The chosen performance metric is AR, which for the unperturbed data are set at a constant value using a fixed threshold approach using an upper threshold (UTH) and a lower threshold (LTH). The AR is a measure of FDI performance that can be implemented to showcase the difference in sensitivity between the two methods.

An UTH and LTH are set based on the metric of FAR. A FAR of 5% is chosen for the reference unperturbed data. For each of the three observed singular values, unique reference UTH and LTH are set.

Utilising the perturbed data, each set is evaluated with the corresponding reference UTH and LTH, which, considering perturbed data as fault data, gives a TAR. The average TAR over all the perturbation sets is then calculated and should be indicative of the change in AR from the reference data compared to the perturbed data evaluated against the reference UTH and LTH.

6.6.2 Sensitivity analysis using the DP method

Utilising the perturbation method as described in Subsection [6.6.1](#) a sensitivity analysis is done using the DP method with a single reference in the holding mode. The interval over

which the DP method is evaluated is also the same as in Subsection 6.6.3, the holding mode interval.

Figure 6.13 shows the reference DP data with the reference UTH and LTH indicated, which is used to evaluate the perturbed data.

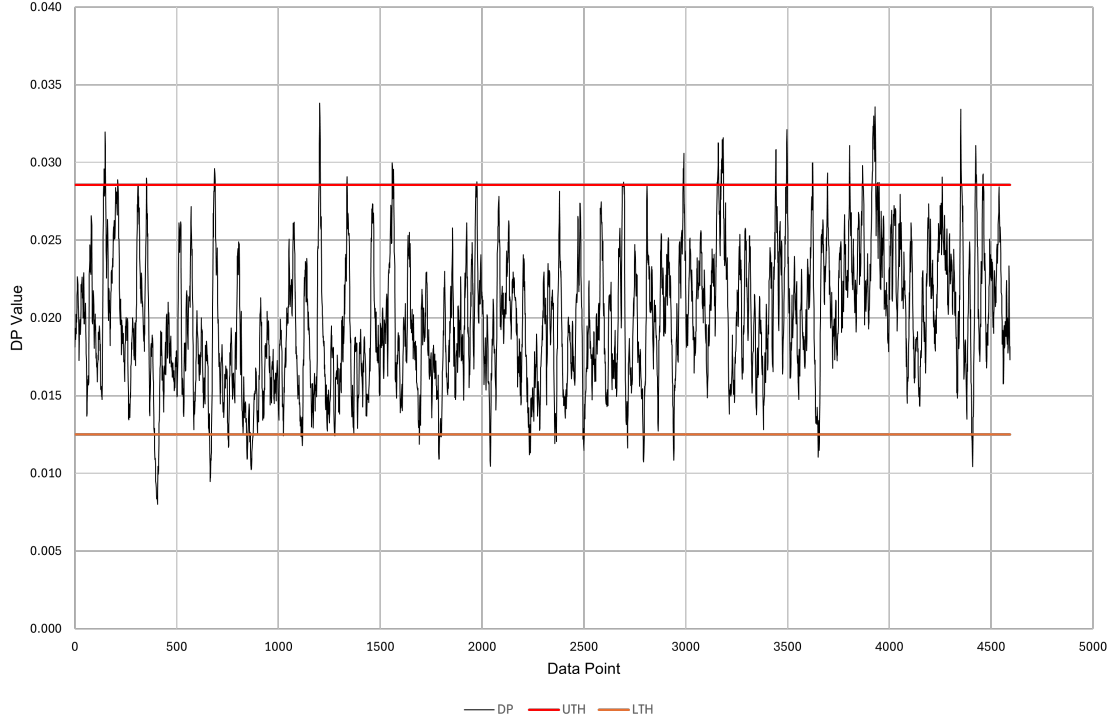


Figure 6.13: Reference thresholds for DP method sensitivity analysis of dataset A

6.6.3 Sensitivity analysis using the SBS SVD method

Given a perturbed node signature time series set $\{\tilde{\mathbf{N}}_{s,k}\}_{k=1}^K$ and the original unperturbed node signature time series data set $\{\mathbf{N}_{s,k}\}_{k=1}^K$ the SBS methodology is performed for each set, which gives a perturbed cost matrix set $\{\tilde{\mathbf{C}}_{s,k}\}_{k=1}^K$ and the reference cost matrix set $\{\mathbf{C}_{s,k}\}_{k=1}^K$. Following the SVD methodology for each of the cost matrix sets, the first three singular values can be determined over all time instances. The values are filtered with a moving average filter. Table 6.9 shows the reference thresholds for each singular value.

Table 6.9: Reference threshold values for singular value sensitivity analysis

	SV1	SV2	SV3
Upper threshold value	19.490	3.322	1.915
Lower threshold value	19.395	3.284	1.857

Figure 6.14 indicates the first SV of the SBS reference data and the reference UTH and LTH for SV 1.

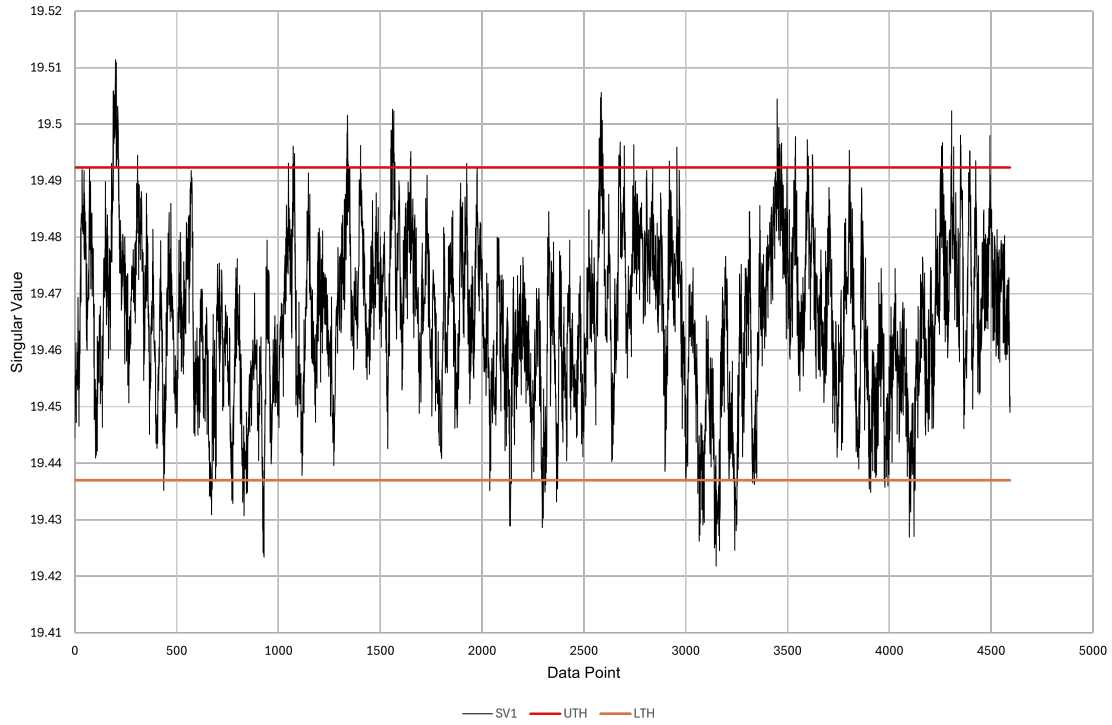


Figure 6.14: Reference thresholds for SBS SVD method sensitivity analysis of dataset A - First singular value

Figure 6.15 indicates the first SV of the SBS reference data and the reference UTH and LTH for SV 2. Figure 6.16 indicates the first SV of the SBS reference data and the reference UTH and LTH for SV 3. The perturbed data are evaluated using the same thresholds to determine the change in AR from the unperturbed set of 5% AR to the perturbed set.

6.7 Conclusion

This chapter meets the objective of applying EGBV analysis on the THERESA system by evaluating node signature data sets using the SBS SVD method and the fixed reference DP method. Additionally, the objective of conducting a sensitivity analysis comparing these two methods is addressed, aimed at comparing the robustness of these two methods.

The results are thoroughly examined in the subsequent chapter, Chapter 7, which provides a detailed discussion of the insights gained from the study and draws conclusions on the value of these methods for FDI in large-scale industrial systems.

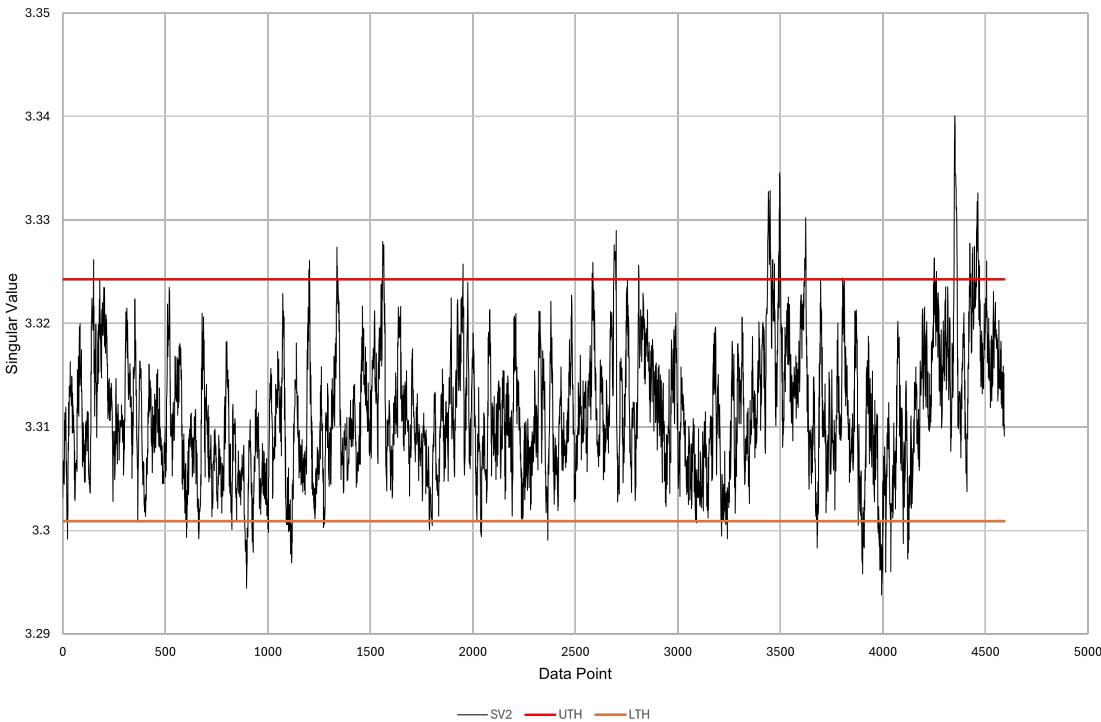


Figure 6.15: Reference thresholds for SBS SVD method sensitivity analysis of dataset A
- Second singular value

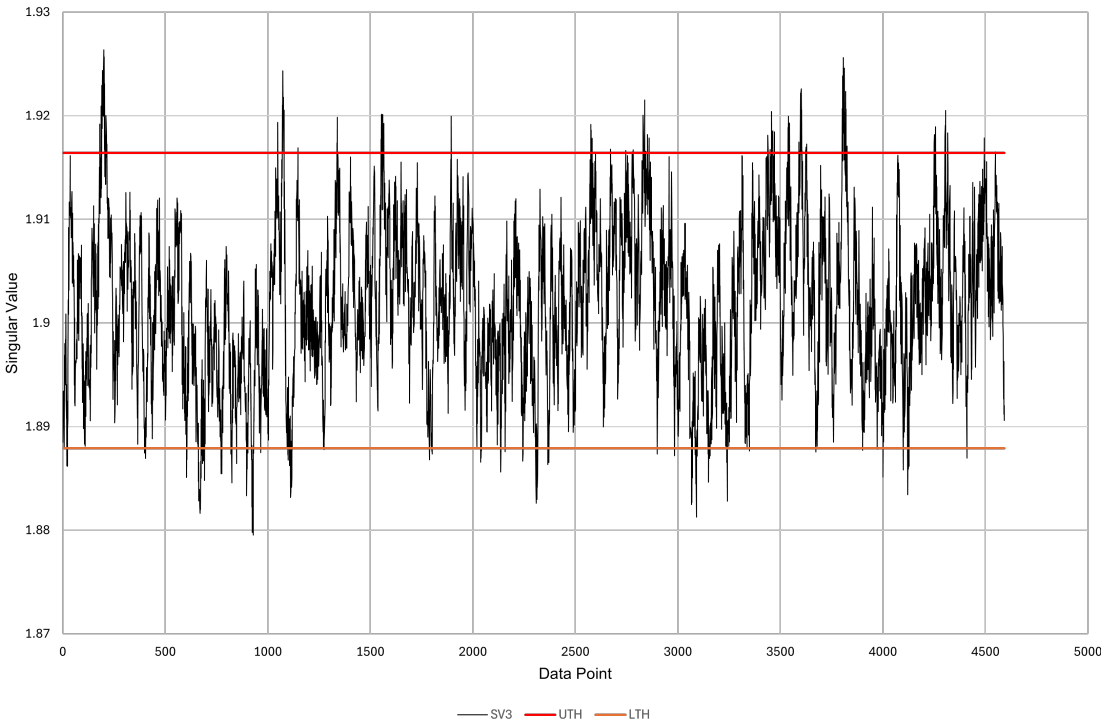


Figure 6.16: Reference thresholds for SBS SVD method sensitivity analysis of dataset A
- Third singular value

Chapter 7

Results

7.1 Introduction

This chapter presents the findings of the study on the application of EGBV on the THERESA thermal power plant system with integrated heat storage capabilities. The results are structured to address the research question and objectives outlined in Chapter 1, focusing on the evaluation of the fixed reference DP method and the SBS SVD approach. The results of the sensitivity analysis using the AR metric are also discussed, providing insight on the robustness and potential of these methods for future FDI applications. The data and analyses presented in this chapter form the foundation for the subsequent discussion and recommendations, illustrating the effectiveness of EGBV in characterising complex large-scale thermal systems.

7.2 EGBV application on the THERESA system

A significant component of this research is the application of the EGBV method on the THERESA system.

The main deliverables expected from the application of EGBV on the THERESA system are discussed in Chapter 4. A system-attributed graph is constructed and the process measurement data are used to set up data sets of time series node signature matrices for further analysis in Chapter 6.

Following a hybrid method where process data are used merits incorporating a process model which would result in reliable calculation of unknown process variables for the application. This could reduce the degree of graph reduction applied in this study while possibly increasing the resolution of the node signature matrices produced from a higher-order system-attributed graph. The data-driven nature of arranging known process variables into the EGBV framework could be made significantly less complex and data-intensive if

a process model is implemented to allow for the application of these unknown variables. This could reduce uncertainty experienced when the system does not have measurement data available.

The EGBV methodology possesses the potential for expansion to encompass larger, more complex systems. This expansion involves the development of a framework for graph reduction, establishing a standardised notation for denoting and describing multiple energy interactions within the system. Additionally, an inherent mechanism could be devised within the EGBV framework to explicitly represent the energy stored within individual components.

7.3 EGBV analysis results

7.3.1 Results of EGBV analysis using the DP method

Initial results from utilising the DP for visualisation of data sets A and B show consistency in the output when data sets A and B are evaluated. As seen in Figure 6.3, the average DP value over modes are consistent across data sets A and B, with minor discrepancies that can be attributed to the transient modes or differences in data set operational points.

It is thus determined that due to the fluctuations observed within the data across one mode to the next, the DP method should show differences in modes, provided an adequate classification method is used. This property is derived from the graph of DP over intervals not being a consistent average value, however, further analysis is required to determine the performance of the method as a mode determination tool. Also, such analysis can be thoroughly done with more data sets for normal operational data.

The DP output is a single parameter that varies over time. In a complex system such as the THERESA system, the degree of data reduction, should be evaluated against the value of the output. It is unclear what the consequences of such reduction would be in terms of FDI capabilities and, as proposed in the literature, the possibility of process control utilising this method.

This parameter is also less sensitive to changes within the node signature matrices compared to the SBS SV method.

The possibility that a less sensitive output could be beneficial in the analysis of long-term change such as mode detection within the THERESA system seems plausible utilising this method. This approach could possibly be used in combination with other methods for FDI or process control.

7.3.2 Results of EGBV analysis using SBS methodology

Table 6.7 and Table 6.8 show the mean value of each SV over each mode as visualised in Figures 6.6 and 6.7.

From this table and the visual representation provided in Figures 6.8 and 6.9 trends across singular values and data sets can be observed. From the fluctuations in the average value over modes it is determined that the output is not mode independent, the output fluctuates or differs with the input parameters. Provided that an adequate method is implemented, this property can be used to determine the system mode or operational set point.

Data sets A and B exhibit similar trends over the data set, however, during modes of high system variability such as startup, warm-up, discharge, and shutdown, the values of the sets deviate. Whereas over modes considered to have less process variability, such as warm-up 2, charge, and hold, the differences between data sets are less. With threshold functions it is possible to distinguish the modes of high process variability from the modes with low process variability. However, this is insufficient to provide an accurate method for determining the individual system modes.

The results from the SBS SVD analysis do not show the ability to differentiate between modes; however, this method does show promise for FDI. If the method is not highly sensitive to mode changes, which is evident from the visualisation, but proves to be sensitive to perturbations, the method has properties that benefit FDI analysis methods. The sensitivity analysis results that are discussed in a subsequent section should provide context for the sensitivity of the method and the possible value for FDI.

7.4 Sensitivity analysis results

Figure 7.1 shows a plot of the AR versus the percentage perturbation for data set A using two methods of analysis. Singular values 1, 2, and 3 form part of the SBS methodology.

Table 7.1: Table of perturbation factors and distance parameter percentage AR

Percent increase by perturbation	Distance parameter
Alarm rate percentage	
0% (Reference)	5.01%
2% (PF = 1.02)	5.21%
4% (PF = 1.04)	6.32%
6% (PF = 1.06)	7.81%

The graphs start at a 5% AR as the 0% perturbation percentage increase represents the reference data where the UTH and LTH are determined at a 5% AR. The initial perturbation of a 2% increase yields a new AR, now considered to be a TAR which is

Table 7.2: Table of perturbation factors and singular value percentage AR

Percent increase by perturbation	SV1	SV2	SV3
Alarm rate percentage			
0% (Reference)	5.01%	5.01%	5.01%
2% (PF = 1.02)	6.60%	7.03%	6.59%
4% (PF = 1.04)	11.50%	12.32%	10.95%
6% (PF = 1.06)	18.40%	19.34%	17.01%

higher in all cases than the reference AR of 5%. Further, it is seen that the initial changes within the singular values and the DP are such that all singular values respond more to the perturbation than the DP.

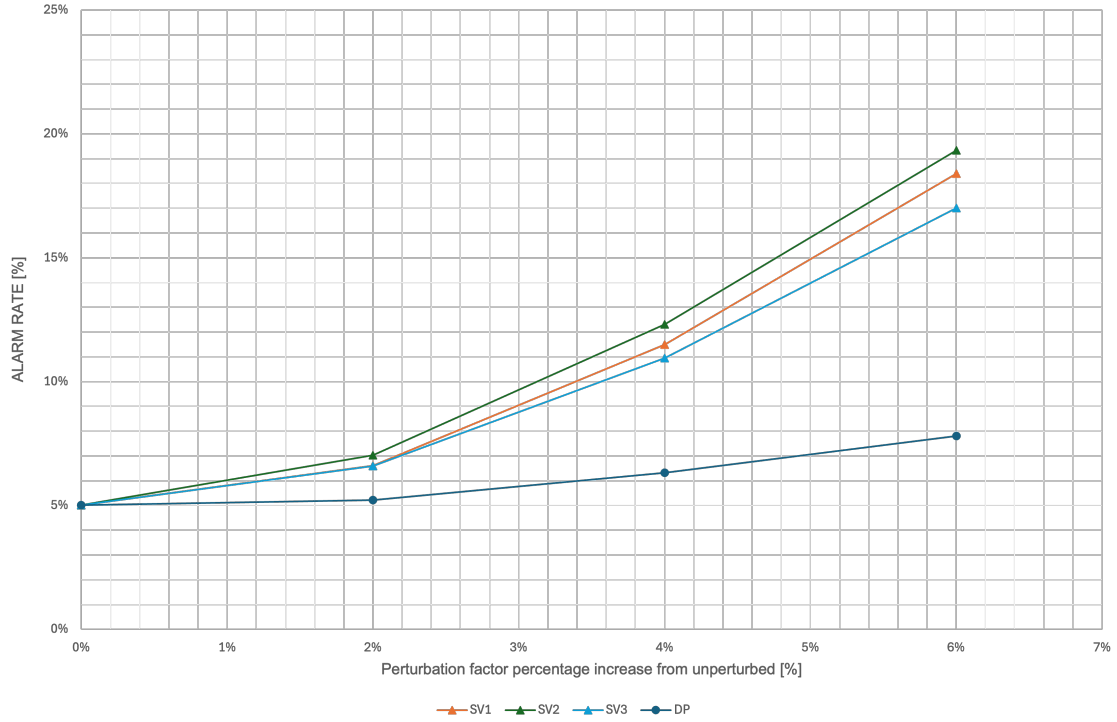


Figure 7.1: Distance parameter method and SBS method alarm rate with percentage perturbation

It is determined from this sensitivity analysis that the changes in the node signature matrix for various perturbations, that the SBS SVD methodology is more sensitive to changes than the DP in terms of the increase in AR caused by a 2%, 4%, and 6% perturbation based on the initial AR of 5% of the unperturbed data.

7.5 Conclusion

This chapter conveys and discusses the results obtained from the EGBV analysis of the node signature data sets produced by the energy characterisation applied in Chapter 4. The results stem from the application of the fixed reference DP method and SBS SVD method, where the data used have been visualised and evaluated. This chapter has also demonstrated the sensitivity of each method showing that the SBS SVD method is more sensitive than the fixed reference DP method. All three singular values of the SBS method are more sensitive than the DP method to perturbations in the node signature matrix when observed over a steady-state operational window for data set A.

Chapter 8

The application of EGBV for FDI purposes

8.1 Introduction

This chapter serves as an additional chapter to the study, following Chapter 6 that addresses this study's objectives in detail. Here, a threshold-based FDI methodology is implemented on the node signature matrix data sets produced in Chapter 4 for the THERESA system. This is to illustrate the usefulness of the EGBV method for FDI.

Threshold-based methods for FDI are widely employed and involve setting predetermined thresholds for specific variables. Any deviation from these limits signals the potential presence of a fault. Thresholds can either be static or dynamic, with the former employing fixed threshold values and the latter adjusting based on operational conditions [60, 61]. Operational data are evaluated based on these thresholds to detect deviations in system performance by comparing measured parameters against the defined limits. Violation of these limits by any point results in an alarm that is classified as a true alarm or a false alarm. False alarms occur when a sample is classified as a violation, but that point is known to be in normal operation, whereas true alarms occur when a sample is classified as a violation and the sample is known to be in a fault condition. False alarm rate and true alarm rate are used as performance metrics for such FDI methods in the form of FAR and TAR. These metrics are calculated as:

$$\text{FAR} = \frac{\text{Number of false alarms}}{\text{Total number of alarms}} \times 100\% \quad (8.1)$$

and,

$$\text{TAR} = \frac{\text{Number of true alarms}}{\text{Total number of actual faults}} \times 100\%. \quad (8.2)$$

In this study a fixed threshold method is utilised similar to the methods implemented by Wolmarans[2]. The approach utilises three distinct data sets: two representing normal operational conditions (NOC) and one encompassing faults. These sets are referred to as data sets A, B and C respectively. The data sets are produced through the fixed reference DP method and the SBS SVD method. Both methods have been described thoroughly in Chapter 6. These methodologies provide different perspectives on system behaviour, and their performance is compared to form an understanding of the possible characteristics of each method for FDI in future research.

To facilitate a comprehensive analysis, the NOC data sets are divided into seven operational modes: startup, warmup 1, warmup 2, charge, hold, discharge and shut-down. The segmentation of datasets into modes allow for a detailed understanding of the system's behaviour across different operational points, to determine mode-dependent thresholds. For each operational mode, upper and lower thresholds are developed using statistical techniques to produce thresholds that allow for a False Alarm Rate of 5%. This FAR is based on the methodology employed by [2] and this threshold development process involves determining the 97.5th percentile and the 2.5th percentile for the data within a mode and is done for each data set independently to produce 4 thresholds for each mode across 2 data sets. This process is repeated for each mode such that data set A has a total of 14 thresholds and data set B has a total of 14 thresholds.

Once these thresholds are established for the two NOC sets, they are combined by selecting the highest UTH and the lowest LTH for each mode across the two data sets. This strategic selection process ensures a representative evaluation of both NOC sets. In this process the total FAR produced when evaluating either NOC data set will be lower than the calibrated 5%, but ensures that the thresholds are not stricter than the 5% threshold, balancing the trade-off between maintaining a low FAR and achieving high detection rates.

The metrics of FAR and TAR will be implemented for a comparison of the two EGBV for FDI approaches and their ability to identify faults within the fault data set. It is important to note that this chapter, based on limited data, provides a preliminary exploration of what the FDI approach might look like in the context of energy graph based visualisation for FDI. Ultimately, this methodology seeks to establish a reliable framework for detecting anomalies in the THERESA system, addressing the pressing need for timely and accurate identification of faults that could impact system performance.

The data that could be used for implementing an FDI method are 3 data sets, 2 normal sets and 1 fault set where the operational parameters are sufficiently similar for comparison across sets. These sets are data sets A and B as used in previous chapters as normal operational data. This chapter introduces a new data set, data set C, which is the fault set. Table 8.1 summarises the mode intervals for data set C.

Table 8.1: Data-set C approximate mode starting positions

Mode description	Start point
Startup	1
Warmup 1	2156
Warmup 2	11376
Charge	14061
Hold	15624
Discharge	24364
Shutdown	28271

In this data set a fault refers to an anomaly that is not expected in normal operation. In this case Fault 1 is an over pressure within pump 1P02 or node 1. This is a point where the pressure exceeds the system warning pressure limit. Fault 2 is a stuck valve 1A19 on the outlet of node 3 to node 11.

Table 8.2: Data set C - Fault descriptions and intervals

Fault	Fault Description	Start Point	End Point
Fault 1	Pressure Fault	6020	6269
Fault 2	Stuck Valve	9554	11415

In the THERESA system these faults are only considered faults because of thresholds that are exceeded for certain parameters. It is uncertain what the cause is of the sharp pressure rise in Fault 1 at node 1, however because this component is mainly for circulation of water on the cold side of the preheater, no significant thermodynamic consequences are introduced, compared for instance with a fault such as a leak in the steam generator. In the case of fault 2 a stuck valve causes the pressure to rise sharply within the system, this sharp pressure increase oscillates through the system and is observable in other components, however in terms of operation the system is unaffected, possibly through the intervention of active control systems.

8.2 FDI on the THERESA system using the DP method

The fixed threshold method applied to the fixed reference DP where the data includes a single time series variable referred to as the DP. The thresholds based on a 5 % FAR for each mode are summarised in Table [8.3](#)

Table 8.3: Individual upper and lower NOC threshold values for each mode for data set A and B - DP approach

Mode	UTH-Data set A	LTH-Data set A	UTH-Data set B	LTH-Data set B
Startup	0.489	0.442	0.438	0.334
Warmup 1	0.330	0.205	0.320	0.155
Warmup 2	0.205	0.025	0.173	0.016
Charge	0.044	0.024	0.038	0.025
Hold	0.029	0.014	0.018	0.005
Discharge	0.513	0.046	0.547	0.019
Shutdown	0.554	0.493	0.548	0.540

Using the individual NOC thresholds in Table 8.3, a selection process is used to determine the combined NOC thresholds. The combined upper NOC threshold is chosen as the highest threshold across the 2 data sets in each mode. The combined lower NOC threshold is chosen as the lowest threshold across the 2 data sets in each mode. Which results in the final upper and lower thresholds shown in Table 8.4

Table 8.4: Combined upper and lower NOC threshold values for each mode - DP approach

Mode	UTH-NOC	LTH-NOC
Startup	0.489	0.334
Warmup 1	0.330	0.155
Warmup 2	0.205	0.016
Charge	0.044	0.024
Hold	0.029	0.005
Discharge	0.547	0.019
Shutdown	0.554	0.493

With the combined thresholds for NOC developed, a visualisation and analysis of the thresholds to data set A and B can be done to evaluate the FAR produced from the combined thresholds. Figure 8.1 and 8.2 illustrate a visual representation of the thresholds. The NOC thresholds produce a FAR of 2.67% in data set A and a FAR of 3.02% in data set B, indicating that the thresholds are less sensitive than the individual thresholds calibrated at 5%.

To evaluate the approach in terms of FD performance metrics, a third data set containing faults is introduced and evaluated with the NOC thresholds. Figure 8.3 plots data set C with the relevant NOC thresholds. Highlighting the known faults within the system, Figure 8.4 shows the known fault intervals with the DP parameter across all samples. The thresholds detect both faults, but produce false alarms in mode transition intervals. Significant known fault points violate the thresholds for fault 1 and fault 2.

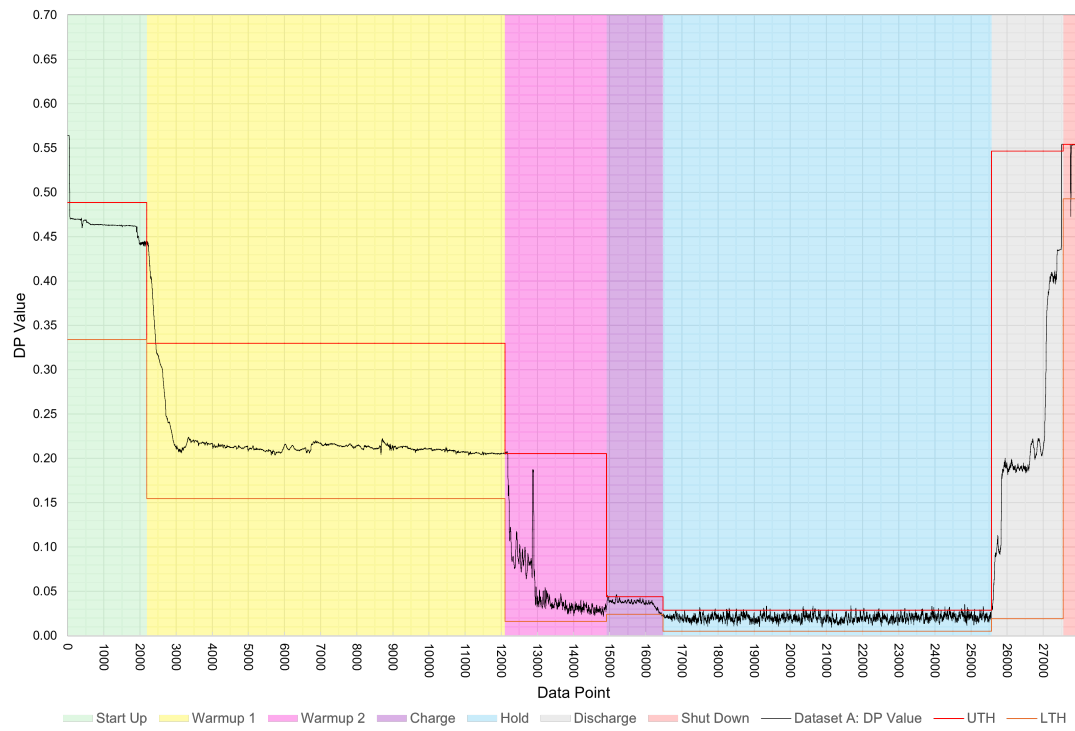


Figure 8.1: Data set A plotted with NOC thresholds - DP approach

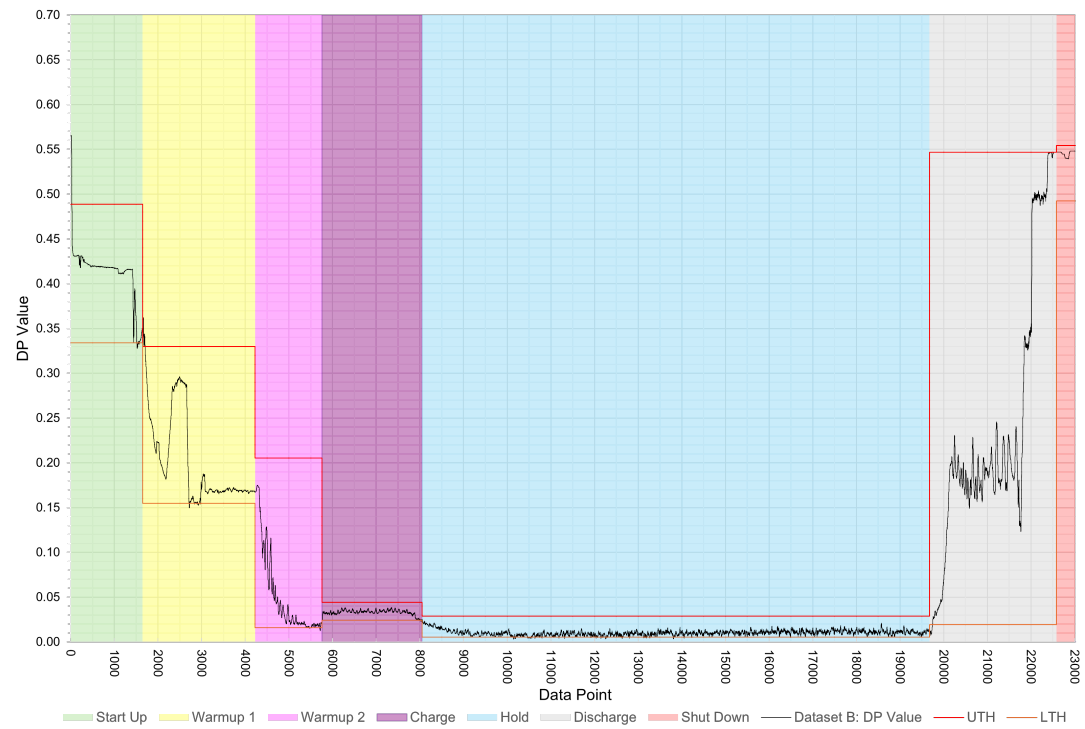


Figure 8.2: Data set B plotted with NOC thresholds - DP approach

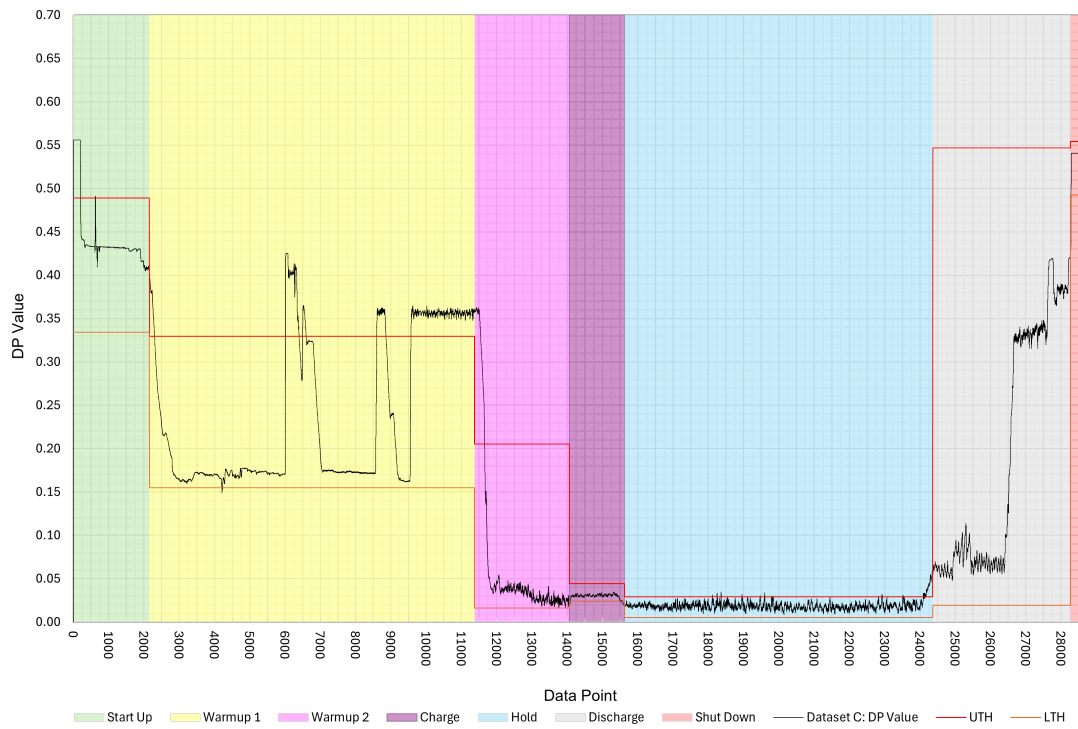


Figure 8.3: Data set C plotted with NOC thresholds - DP approach

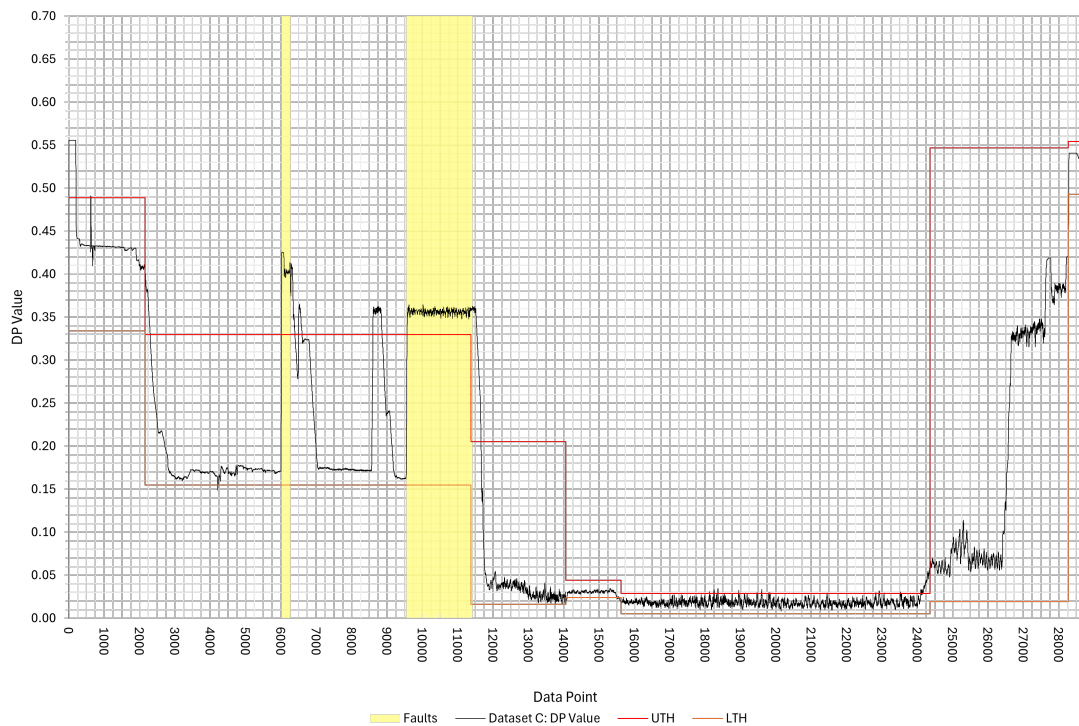


Figure 8.4: Data set C plotted with NOC thresholds and known fault intervals highlighted - DP approach

With the fixed threshold method applied in the DP method the FAR and TAR percentages for data set C evaluated with NOC thresholds which produce results as summarised in

Table 8.5

Table 8.5: FDI performance metrics for data set C evaluated with NOC thresholds - DP approach

Total Data Points	28696
Total Alarms	3670
True Alarms	2096
Total False Alarms	1574
FAR	5.49
TAR	99.34

The FDI performance shows a FAR of 5.49% and a TAR of 99.34%. In terms of detection this is a significant percentage with a relatively low FAR considering that the thresholds on the NOC data sets yielded a FAR of 2.67% for data set A and 3.02% for data set B. Overall the approach shows promise for FDI with significant portions of each known fault being classified as a violation of the thresholds.

8.3 FDI on the THERESA system using the SBS SVD method

The fixed threshold based method applied to the SBS singular values where the data includes three selected time series variables referred to as SV 1, SV 2 and SV 3. In this case each SV is analysed independently, thus the approach is duplicated for each SV. Tables 8.6, 8.7 and 8.8 indicate the upper and lower thresholds for each mode for data set A and B for SV1, SV2 and SV3.

Table 8.6: Upper and lower threshold values for each mode of data set A and B - SBS approach for SV1

Mode	UTH-Data set A	LTH-Data set A	UTH-Data set B	LTH-Data set B
Startup	20.824	17.004	20.867	16.577
Warmup 1	19.563	18.834	20.137	18.910
Warmup 2	19.821	19.350	19.789	19.186
Charge	19.420	19.371	19.436	19.388
Hold	19.489	19.378	19.536	19.425
Discharge	20.129	17.937	20.544	18.424
Shutdown	17.497	14.682	18.592	18.226

Table 8.7: Upper and lower threshold values for each mode of data set A and B - SBS approach for SV2

Mode	UTH-Data set A	LTH-Data set A	UTH-Data set B	LTH-Data set B
Startup	6.731	4.510	6.765	3.933
Warmup 1	3.898	3.177	3.979	3.228
Warmup 2	3.626	3.253	3.497	3.324
Charge	3.241	3.193	3.312	3.289
Hold	3.322	3.264	3.359	3.320
Discharge	6.699	3.366	6.310	3.342
Shutdown	7.502	6.367	6.388	6.247

Table 8.8: Upper and lower threshold values for each mode of data set A and B - SBS approach for SV3

Mode	UTH-Data set A	LTH-Data set A	UTH-Data set B	LTH-Data set B
Startup	2.221	1.990	2.221	1.946
Warmup 1	1.905	1.784	2.023	1.764
Warmup 2	2.076	1.810	1.983	1.774
Charge	1.857	1.812	1.775	1.749
Hold	1.915	1.835	1.884	1.793
Discharge	2.090	1.799	2.224	1.791
Shutdown	2.117	2.037	2.109	2.104

Using the thresholds in Tables [8.6](#), [8.7](#) and [8.8](#) a selection process is used to determine the combined thresholds. The combined UTH is chosen as the highest threshold across the 2 data sets in each mode. The combined LTH is chosen as the lowest threshold across the 2 data sets in each mode. This results in the final upper and lower thresholds for each SV as shown in Tables [8.9](#), [8.10](#) and [8.11](#).

Table 8.9: Combined upper and lower threshold values for each mode combined across NOC- SBS approach for SV1

Mode	UTH-NOC	LTH-NOC
Startup	20.867	16.577
Warmup 1	20.137	18.834
Warmup 2	19.821	19.186
Charge	19.436	19.371
Hold	19.536	19.378
Discharge	20.544	17.937
Shutdown	18.592	14.682

Table 8.10: Combined upper and lower threshold values for each mode combined across NOC- SBS approach for SV2

Mode	UTH-NOC	LTH-NOC
Startup	6.765	3.933
Warmup 1	3.979	3.177
Warmup 2	3.626	3.253
Charge	3.312	3.193
Hold	3.359	3.264
Discharge	6.699	3.342
Shutdown	7.502	6.247

Table 8.11: Combined upper and lower threshold values for each mode combined across NOC- SBS approach for SV3

Mode	UTH-NOC	LTH-NOC
Startup	2.221	1.946
Warmup 1	2.023	1.764
Warmup 2	2.076	1.774
Charge	1.857	1.749
Hold	1.915	1.793
Discharge	2.224	1.791
Shutdown	2.117	2.037

With the combined thresholds for NOC developed, a visualisation and analysis of the thresholds to data set A and can be done to evaluate the FAR produced from the combined NOC thresholds for each SV. Figures 8.5, 8.6 and 8.7 illustrate a visual representation of data set A evaluated with the combined NOC thresholds for SV1, SV2 and SV3. For data set A the new FAR determined with the combined thresholds for SV1,SV2 and SV3 are: 2.91%, 3.55% and 1.79% respectively.

The same procedure is applied for data set B where the combined thresholds for NOC is used to evaluate the FAR produced from the combined thresholds for each SV. Figures 8.8, 8.9 and 8.10 illustrate a visual representation of the thresholds. For data set B the new FAR determined with the combined thresholds for SV1,SV2 and SV3 are: 2.76%, 2.53% and 3.25% respectively.

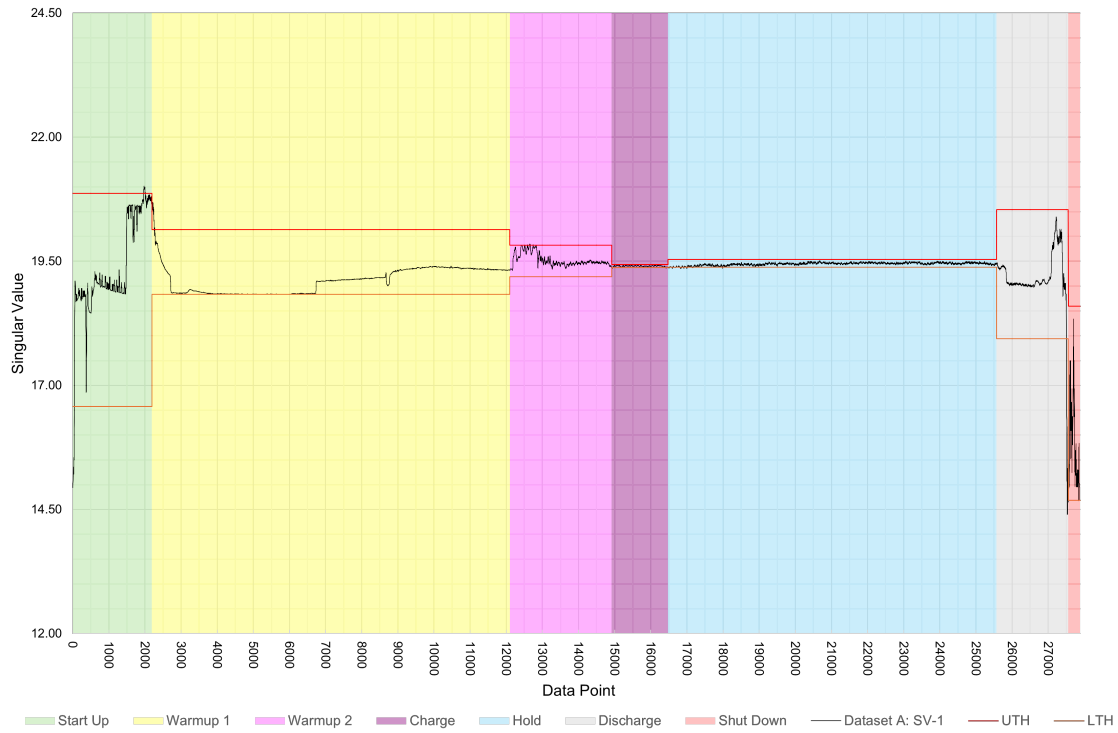


Figure 8.5: Data set A plotted with combined NOC thresholds - SBS SVD approach for SV1

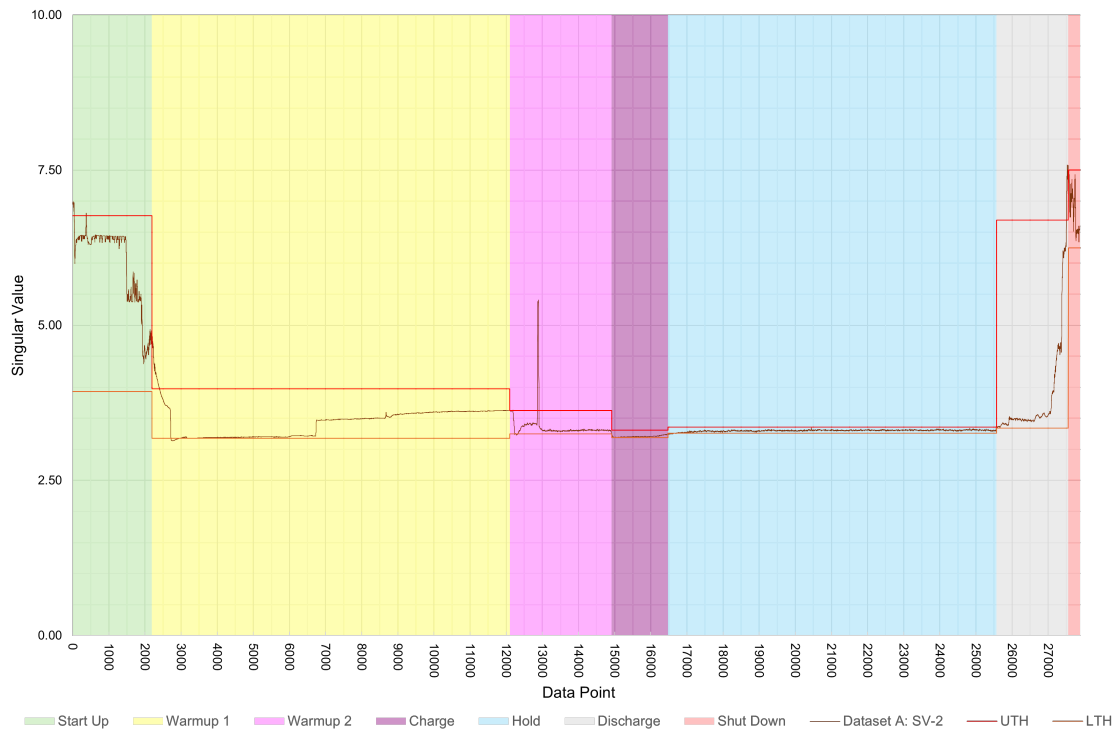


Figure 8.6: Data set A plotted with NOC thresholds - SBS SVD approach for SV2

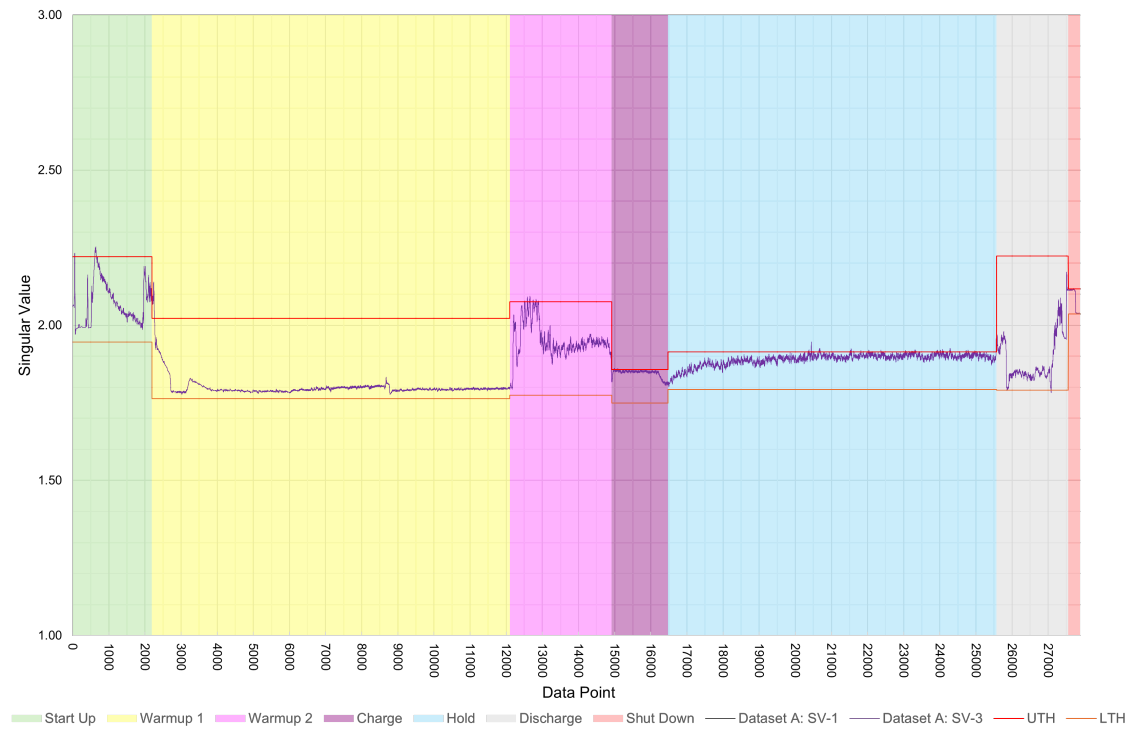


Figure 8.7: Data set A plotted with NOC thresholds - SBS SVD approach for SV3

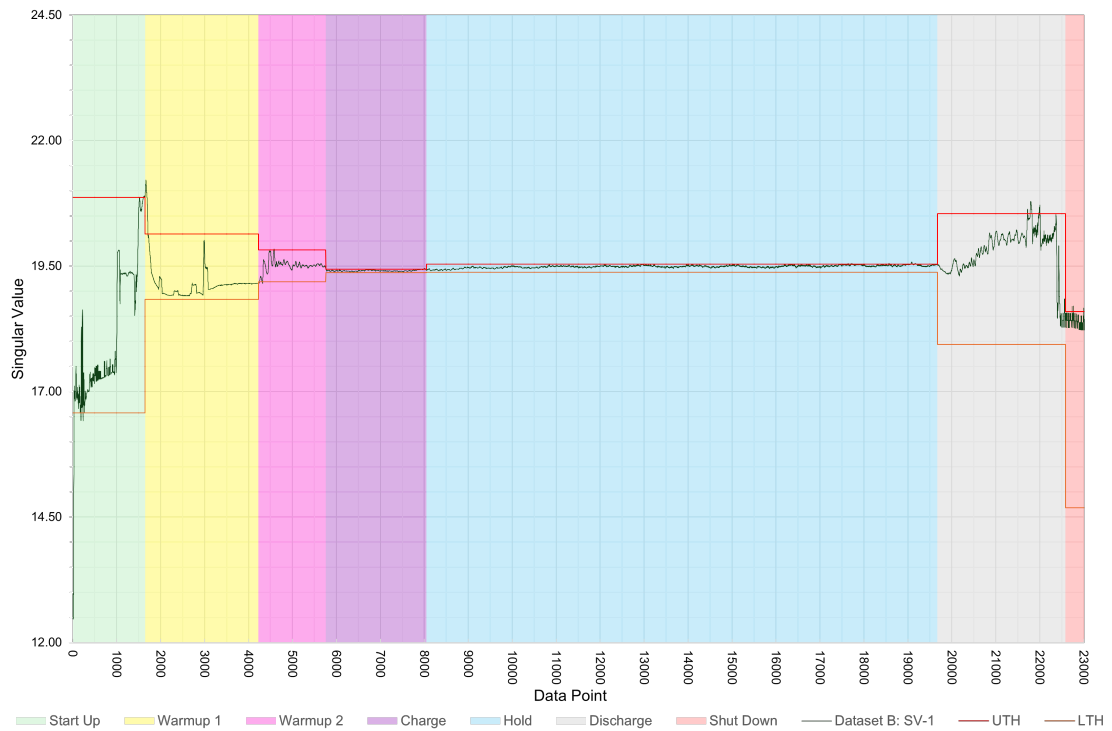


Figure 8.8: Data set B plotted with NOC thresholds - SBS SVD approach for SV1

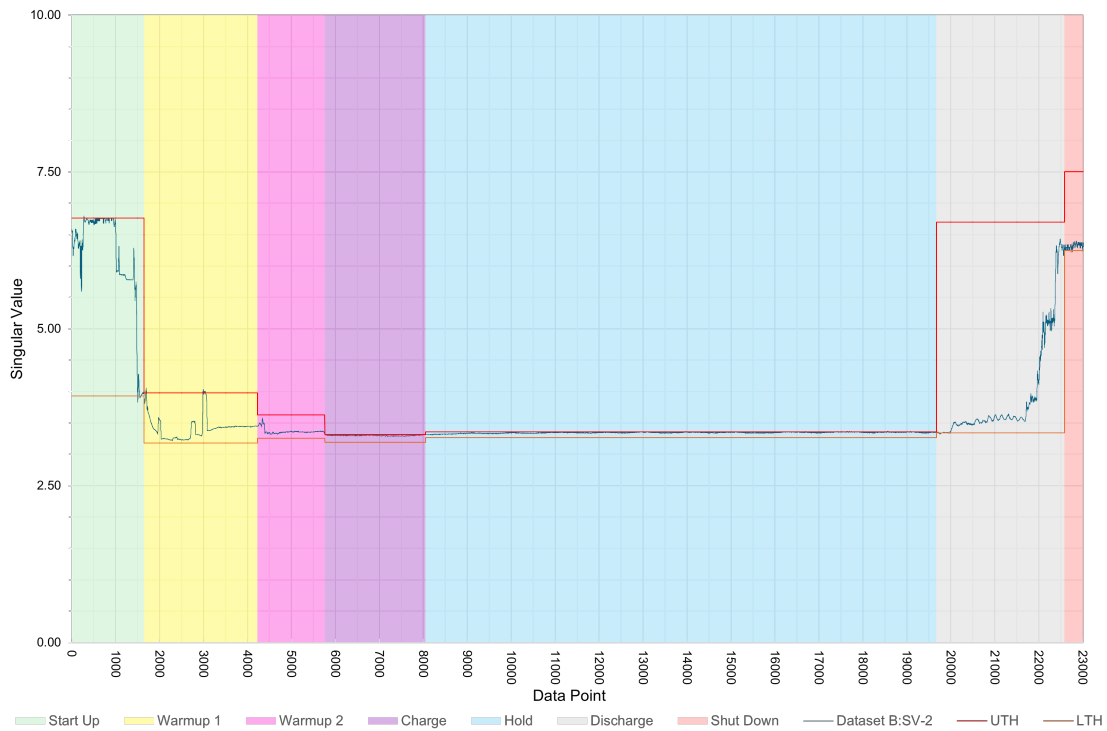


Figure 8.9: Data set A plotted with NOC thresholds - SBS SVD approach for SV2

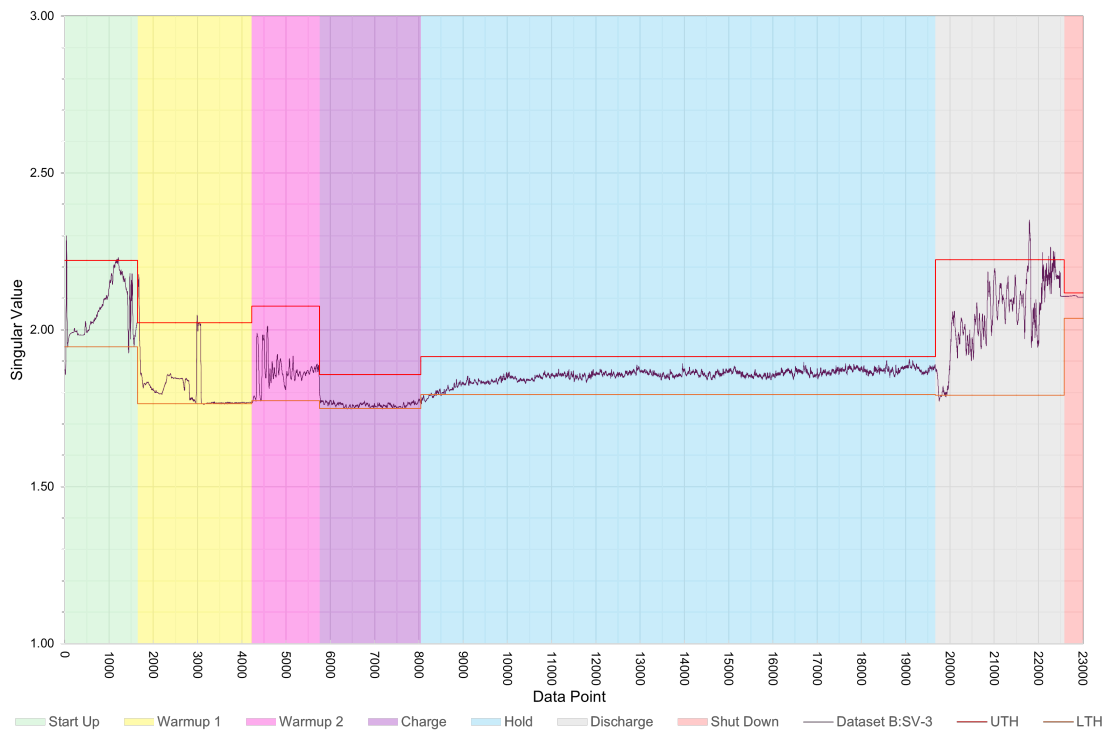


Figure 8.10: Data set A plotted with NOC thresholds - SBS SVD approach for SV3

To evaluate the approach in terms of FD performance metrics, a third data set containing faults is introduced and evaluated with the combined NOC thresholds. Figures [8.11](#), [8.13](#) and [8.15](#) plots data set C evaluated with the combined NOC thresholds. It is expected

that faults within the system should elicit a response from the system such that either of the singular values indicate the presence of a fault by violating the thresholds. This could be either through fault persistence or through system compensation.

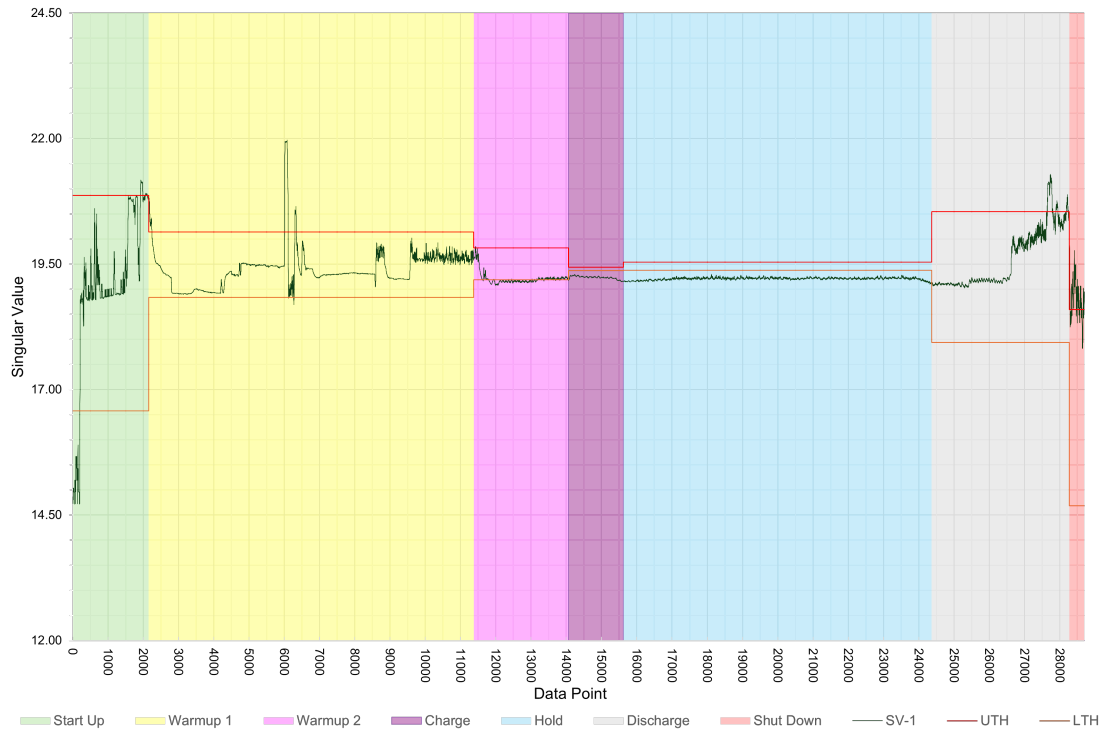


Figure 8.11: Data set C plotted with NOC thresholds - SBS SVD approach for SV1

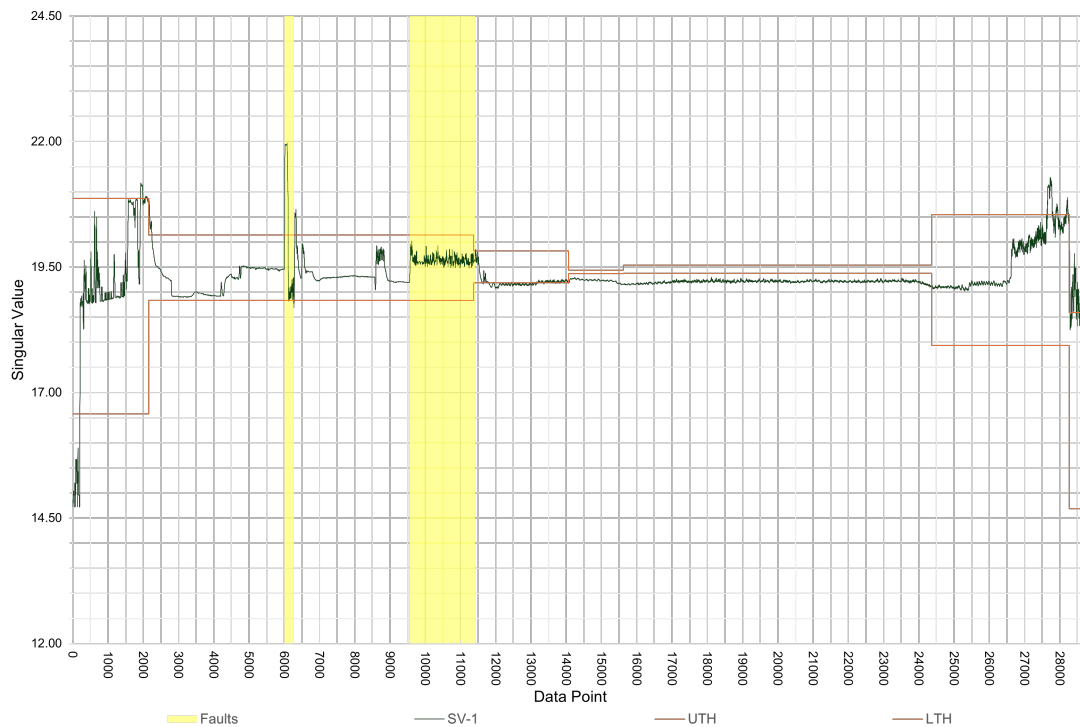


Figure 8.12: Data set C plotted with NOC thresholds and known fault intervals highlighted - SBS SVD approach for SV1

In Figure 8.11 the expected result would be that the system operates between the defined thresholds for a majority of the samples, but that faults cause the data to violate these thresholds. As seen in Figure 8.12 the known faults within the system are highlighted and no faults are present in the charging and holding modes, however in these modes the data are violating the thresholds for a majority of the samples. This contributes significantly to false alarms within the system.

There are many possible explanations that can account for this occurrence, it is however postulated that the charging and holding modes are very dependent on the experimental parameters under which the system was set to operate. The data sets that have been acquired for this study are not consistent in terms of set points for various parameters during an experiment, the violation of these thresholds by data set C may indicate that the method is sensitive to operational point.

Examining the known fault intervals it is evident that fault 2 is not detected and only samples in the fault interval of fault 1 violate the thresholds, thus the fault is detected, but missed detections are relatively high. Figures 8.13 and 8.14 do not show the same degree of false alarms as the first SV. In this case the system seems to be operating within the expected range as set by the thresholds.

Examining the known fault intervals, compared to the first SV all points classified under fault 1 violate the thresholds and some of the points classified as fault 2, thus fault 1 is detected with greater accuracy than fault 2.

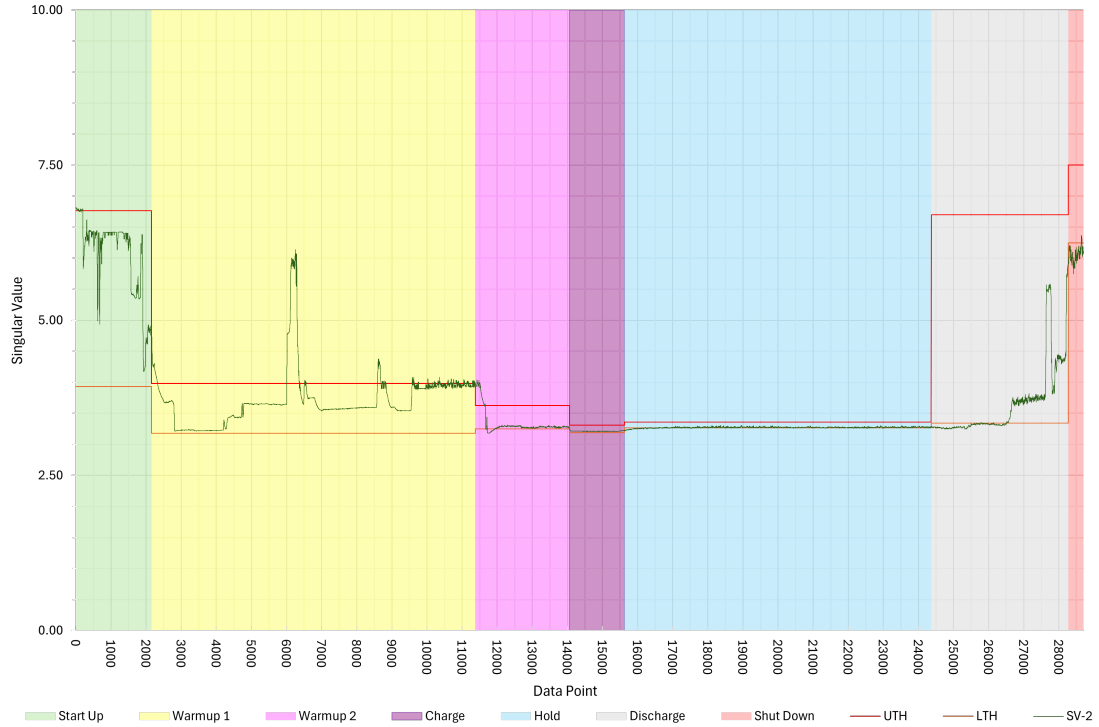


Figure 8.13: Data set C plotted with NOC thresholds - SBS SVD approach for SV2

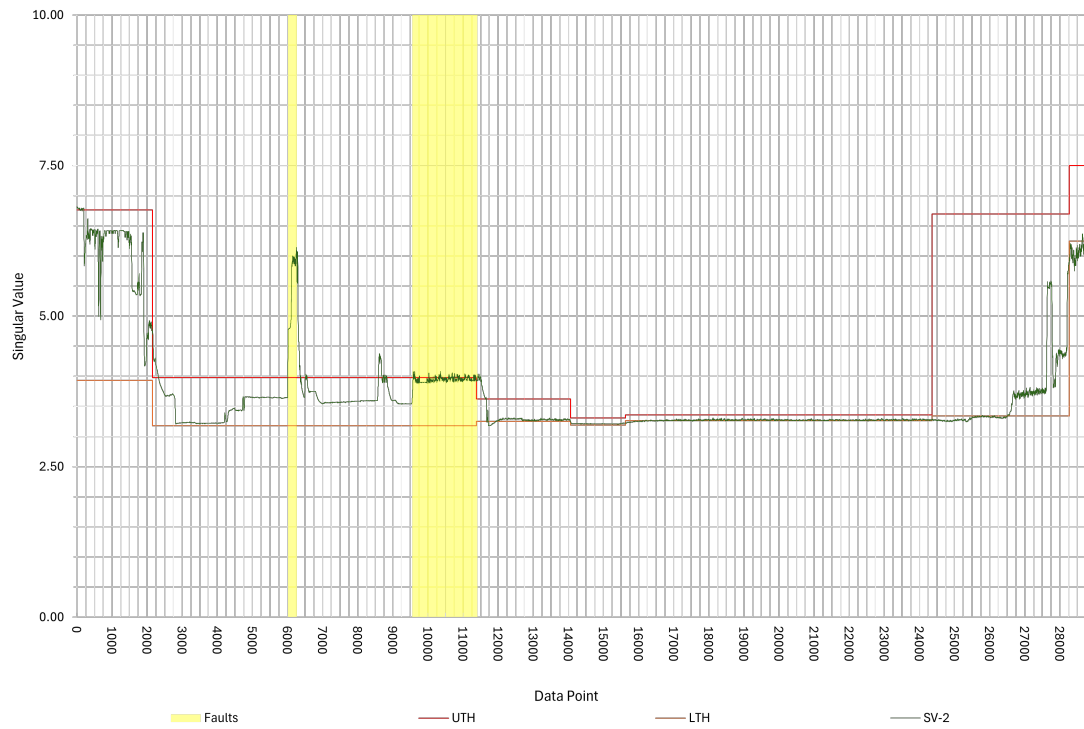


Figure 8.14: -Data set C plotted with NOC thresholds and known fault intervals highlighted - SBS SVD approach for SV2

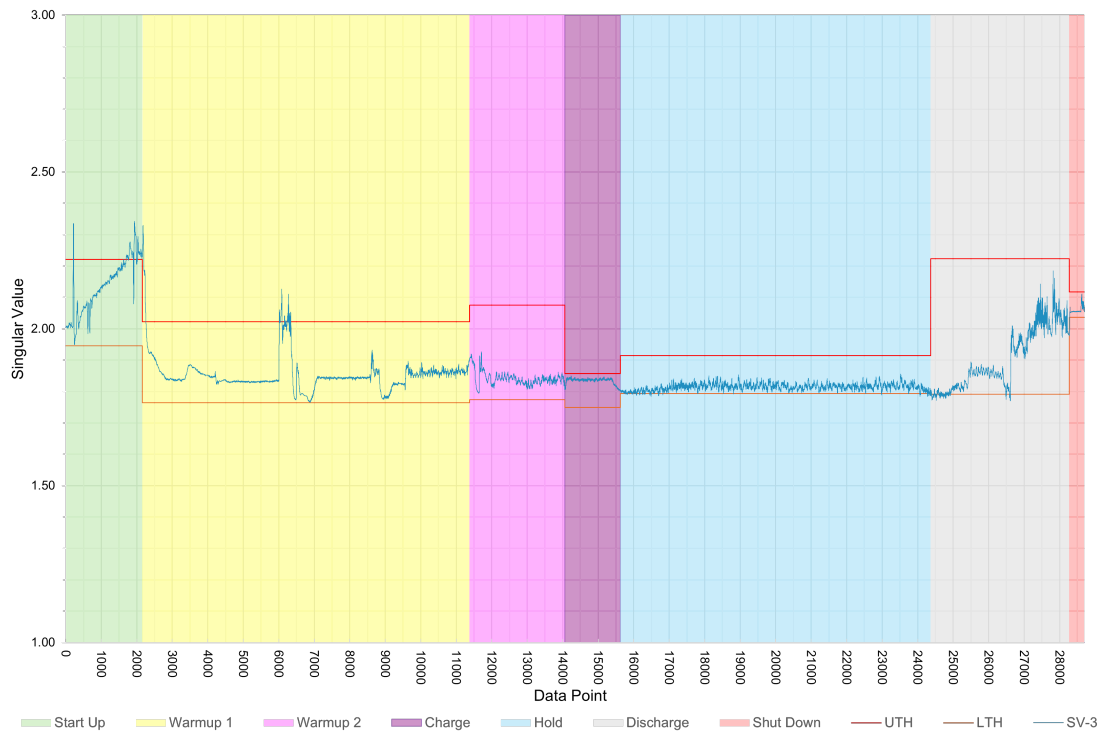


Figure 8.15: Data set C plotted with NOC thresholds - SBS SVD approach for SV3

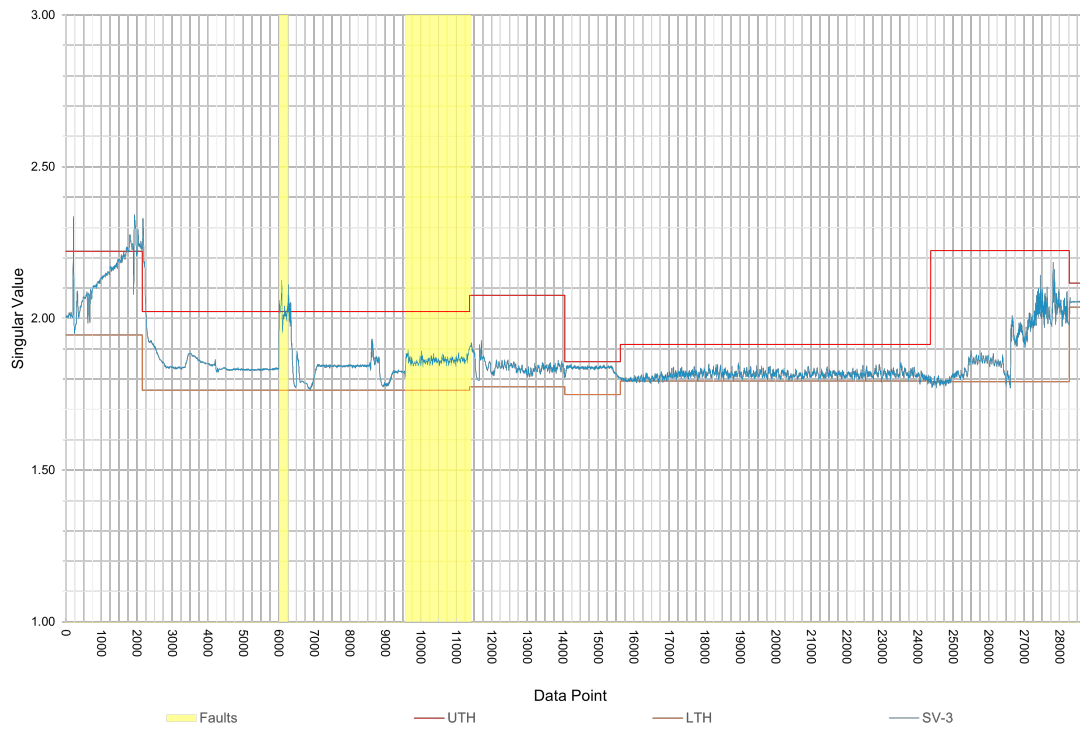


Figure 8.16: Data set C plotted with NOC thresholds and known fault intervals highlighted - SBS SVD approach for SV3

The third SV observed for data set C in Figures 8.15 and 8.16 indicates that the set thresholds account for the process expected limits with false alarms mainly occurring around mode transitions.

Examining the known fault intervals in Figure 8.16 it is clear that compared to the second SV, fault 1 is less pronounced with partial detection of the samples and fault 2 has 0 samples detected indicating that the overall performance in terms of detection is less in SV3 than in SV1 and SV2.

In this chapter the key performance metrics used to determine the FDI characteristics of the SBS SVD approach is the FAR and TAR percentages for data set C evaluated with NOC thresholds the parameters for the analysis is shown in Tables 8.12, 8.12 and 8.12.

Table 8.12: FDI performance metrics for data set C evaluated with NOC thresholds - SBS SVD approach - SV1

Total Data Points	28695
Total Alarms	12854
True Alarms	96
Total False Alarms	12758
FAR	44.46
TAR	4.55

Table 8.13: FDI performance metrics for data set C evaluated with NOC thresholds - SBS SVD approach - SV2

Total Data Points	28695
Total Alarms	5510
True Alarms	795
Total False Alarms	4715
FAR	16.43
TAR	37.68

Table 8.14: FDI performance metrics for data set C evaluated with NOC thresholds - SBS SVD approach - SV3

Total Data Points	28695
Total Alarms	1184
True Alarms	111
Total False Alarms	1073
FAR	3.74
TAR	5.26

Overall across SV1, SV2 and SV3 the FDI performance varies. In terms of FAR, SV1 achieved the highest value of 44.46% followed by SV2 that achieved a value of 16.43% and the lowest value of 3.74% is achieved by SV3.

Further, in terms of TAR, SV2 achieved the highest value of 37.68% followed by SV3 at a significantly lower 5.26% and the lowest value is achieved by SV1 with 4.55%.

It is clear from this analysis that between fault 1 and fault 2, fault 1 showed better detection rates across all three singular values, where fault 2's detection rates are substantially lower.

It is also evident that contrary to what is expected, the FAR was the highest in the first SV and the lowest in the third SV, however this is attributed to the charge and hold modes thresholds that were introducing false alarms for a significant duration of the process. From this it is deduced that the first SV is sensitive to operational point.

8.4 Conclusion

The objective of this chapter was to provide an extended analysis of the application of EGBV for FDI on the THERESA system using two EGBV analysis approaches. This chapter has provided results that can be used to draw conclusions on the characteristics of the FDI performance using the fixed reference DP approach or the SBS SVD approach. Each of these methods have been used together with threshold development to determine

the FAR and TAR percentages for the methods.

For the DP approach the following results are obtained: The FDI performance shows a FAR of 5.49% and a TAR of 99.34%. In terms of detection this is a significant percentage with a relatively low FAR considering that the thresholds on the NOC data sets yielded a FAR of 2.67% for data set A and 3.02% for data set B. Overall the approach shows promise for FDI with significant portions of each known fault being classified as a violation of the thresholds.

For the SBS SVD approach the following results were obtained: Overall across SV1, SV2 and SV3 the FDI performance varies. In terms of FAR, SV1 achieved the highest value of 44.46% followed by SV2 that achieved a value of 16.43% and the lowest value of 3.74% is achieved by SV3.

Further, in terms of TAR, SV2 achieved the highest value of 37.68% followed by SV3 at a significantly lower 5.26% and the lowest value is achieved by SV1 with 4.55%.

It is clear from this analysis that between fault 1 and fault 2, fault 1 showed better detection rates across all three singular values, where fault 2's detection rates are substantially lower.

It is also evident that contrary to what is expected, the FAR was the highest in the first SV and the lowest in the third SV, however this is attributed to the charge and hold modes thresholds that were introducing false alarms for a significant duration of the process. From this it is deduced that the first SV is sensitive to operational point.

Chapter 9

Conclusion and recommendations

9.1 Introduction

The culmination of this research on EGBV applied on the THERESA thermal power plant system with integrated heat storage capabilities has provided valuable insights into the potential for future FDI applications employing similar methods. This chapter summarises the key findings of the study, highlighting the effectiveness and limitations of the methods employed. It also offers recommendations for future research and practical applications, based on the insights gained from the application of the EGBV method on the THERESA system and the outcomes of the sensitivity analysis using AR as a metric. These conclusions and recommendations aim to guide further exploration and implementation of EGBV in large-scale industrial systems, contributing to the advancement of energy-based FDI.

9.2 Reflection on research objectives

The conclusion of this study requires evaluation of the research objectives to ensure that the research question is adequately addressed. The subsequent subsections discuss how each objective is addressed.

9.2.1 THERESA system overview and data acquisition

The objective of this section is to examine the THERESA system in terms of its operation and data acquisition. Chapter 3 addresses this objective by offering a comprehensive analysis of the system's functionality, components, operational modes, and data. This chapter establishes the context for the subsequent chapters, providing essential background for the discussions that follow.

9.2.2 Application of EGBV on the THERESA system

The application of EGBV on the THERESA system is addressed in Chapter 4 where the process is discussed in depth guided by the methods as discussed in Chapter 2. In this chapter, the THERESA system graph is constructed and system attributes are determined. The final aspect of the objective is addressed through the determination of the energy attributed graph that forms the basis for the node signature matrix.

9.2.3 Verification of EGBV applied on the THERESA system

The objective of verifying the output node signature matrices is addressed in Chapter 5 where the node signature matrices are evaluated to determine if the results obtained from further analysis are realistic.

9.2.4 EGBV analysis

The EGBV analysis objective is addressed in Chapter 6. In this chapter two methods are explored. The first is the fixed reference DP method and the second is the SBS SVD method. Further, the outputs of each method are visualised for the THERESA system, and analysis of these investigations extracts what is deemed as characteristic information for each method applied on the THERESA system.

9.2.5 Sensitivity analysis

The investigations of the EGBV analysis methods, discussed in Chapter 6, are accompanied by a comparative sensitivity analysis which determines the sensitivity of each method to changes in the node signature matrix. These sensitivity analyses are done based on matrix perturbations which impart a fixed change in node signature matrix entries to determine if these changes cause a significant change in the AR compared to the unperturbed reference set.

9.2.6 The application of EGBV for FDI purposes

Chapter 8 has demonstrated the application of FDI on the THERESA system through the use of fixed thresholds developed from two NOC data sets for each mode of the THERESA system. The chapter evaluated the performance metrics of two approaches, the fixed reference DP approach and the SBS approach. The objective is addressed and remains open for future research to explore and refine.

9.3 Recommendations for future work

9.3.1 Increase information within the node-signature matrix

Through the application of the current EGBV methodology on the THERESA system it is evident that the method becomes increasingly more complex and computationally demanding as information within the system and subsequently the node signature matrix increases.

However, from this study, it is recommended that the construction of the node signature matrix is re-evaluated specifically to ensure that systems with storage capabilities are better described.

By inserting a new column to the right of the first column (rate of exergy change), the new proposed node signature matrix will have a rate of energy stored component, which should account for stored energy. Also, this term may help determine losses within the system as the equation, $\dot{E}_{in} - \dot{E}_{stored} - \dot{E}_{out} - \dot{E}_{losses} = 0$, should only have one unknown variable, $\dot{E}_{losses}$. The reason for this suggestion is to make the method's verification procedure within a node signature matrix set more straight forward as currently verification of the energy stored in a component over time requires tedious computations. The possibility that the added information could change the output of the approaches is limited as this property should inherently be captured in the data of the current matrix structure.

To elaborate, the current node signature matrix takes into account the rate of exergy change and rate of energy change at various points within the system. However, the system is not well described in terms of the node signature matrix for storage components. In this case, the node signature matrix inherently accounts for energy into nodes and out of nodes, however, in an instance where these values do not balance, a specific component, considering there is no fault present, is storing energy or losing energy to the environment, or possibly a component where such energy interaction is not accounted for.

9.3.2 Utilisation of structural information

The main benefit of utilising EGBV for FDI is considered to be the conservation of system structural information by nodes and links in graphs that define a specific system. However, the current analysis methods used for FDI do not leverage this property, and during the analysis the link between the data and the structural information is lost, leading to possible complications in determining the location of the fault or discrepancy. Using a method that would maintain the link between the output data and the original fault data (a specific process variable within a node) will be beneficial not only in FDI but also in troubleshooting the method during development.

9.3.3 EGBV generalisation

The EGBV method is still in its developing phases and still depend on the type of system it is applied on. Application of the EGBV method has been investigated on various systems, these systems share a majority of characteristics that is used within the EGBV application, however for each system it is necessary to make assumptions and decisions that are specific to the system. It would be beneficial if the method can be developed in such a way to be more generally applicable. Another benefit will be to devise a generalised methodology of generating a system graph.

9.3.4 Reference generation in mode varying systems

The investigation of an optimal fixed reference node signature matrix is a topic which may hold valuable information for further study as this system is the first system studied in EGBV where a reference determination is not straightforward.

Also, further research in reference determination methods where the reference node signature is not fixed throughout the evaluation of data, as used in the SBS approach could give more insight into the ideal analysis approaches of these outputs.

In the THERESA system, determining the system mode is a tedious process involving multiple process variables to be evaluated and knowledge about the system to be incorporated in decision making. Further effort in establishing a framework for confident mode determination in complex mode-varying systems should reduce the time input required for analysis of these systems and has merit for further investigation.

9.4 Further research on the THERESA system

This study has applied the EGBV methodology on the THERESA system with accessible information and data, however, the THERESA system holds potential for future research. It is recommended that a process model of the THERESA system be compiled as this model will be valuable for validation and verification of further research. Further analysis of FDI on the THERESA system holds merit. This is based on the assumption that certain experiments can be executed where faults are induced on the system in a controlled manner to generate fault operational data for analysis.

9.5 Closure

The EGBV method as applied in this study characterises the THERESA system as a large scale industrial process with heat storage capabilities. The approaches of analysis, namely the fixed reference DP approach and the SBS approach show that in terms of the sensitivity analysis the SBS approach is more sensitive than the DP approach. The applicability of the EGBV method was evaluated by means of a standard threshold approach. Also, two EGBV approaches, the DP and SBS approaches were evaluated. The DP approach showed superior detection performance. Applying EGBV on the THERESA system demonstrates promise for FDI in large-scale industrial systems with heat storage capabilities. The characteristics of the EGBV analysis methods applied in this study could prove valuable for future advancements in EGBV for FDI.

Bibliography

- [1] 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.
- [2] W. Wolmarans, “A comparison of PCA- and energy-based fault detection and isolation in a physical heated twotank process,” Thesis, North-West University (South Africa), 2022. [Online]. Available: <https://repository.nwu.ac.za/handle/10394/41379>
- [3] J. Smith, “Energy-based fault detection and isolation in an industrial steam turbine system,” Master’s thesis, North-West University (South Africa), 2021. [Online]. Available: <http://hdl.handle.net/10394/37755>
- [4] K. R. Uren, G. van Schoor, and L. Auret, “An energy-attributed graph approach for the purposes of FDI in a heated two-tank system,” *IFAC-PapersOnLine*, vol. 52, no. 14, pp. 159–164, 2019.
- [5] G. Van Schoor and K. R. Uren, “A vision of energy-based visualisation of large industrial systems for the purposes of condition monitoring,” in *Conference on Condition Monitoring and Diagnostic Engineering Management (COMADEM)*., North-West University, South Africa., 2018.
- [6] Hochschule Zittau/Görlitz, “THERESA – Thermal Energy Storage Test Facility,” Jun. 2022, accessed: Jan. 22, 2024. [Online]. Available: <https://theresa.hszg.de/en/theresa>
- [7] T. Klette, T. Gubsch, C. Vogel, D. Kratzsch, S. Braun, S. Härtelt, and A. Kratzsch, “ATHLET simulation code: Model validation of a thermal high-performance storage system,” *International Journal of Smart Grid and Clean Energy*, pp. 1–10, 2018.
- [8] Native Dynamics, “Volume and Wetted Area of Partially Filled Horizontal Vessels,” 2013, accessed: Oct. 17, 2023. [Online]. Available: <https://neutrium.net/equipment/volume-and-wetted-area-of-partially-filled-horizontal-vessels/>
- [9] K. R. Uren and G. Van Schoor, “Reflection on the energy graph-based visualisation approach to FDI of large-scale industrial systems,” *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 4094–4107, 2023. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2405896323021493>

- [10] Y.-J. Park, S.-K. S. Fan, and C.-Y. Hsu, "A Review on Fault Detection and Process Diagnostics in Industrial Processes," *Processes*, vol. 8, no. 9, p. 1123, Sep. 2020. [Online]. Available: <https://www.mdpi.com/2227-9717/8/9/1123>
- [11] R. Isermann, *Fault-Diagnosis Applications: Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems*. Springer Science & Business Media, Apr. 2011, google-Books-ID: v2EOBm7bTJgC.
- [12] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "A review of process fault detection and diagnosis: Part I: Quantitative model-based methods," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 293–311, Mar. 2003. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0098135402001606>
- [13] S. Greyling, G. Van Schoor, K. R. Uren, and H. Marais, "An energy graph eigendecomposition approach to fault detection and isolation applied to a gas-to-liquids process," *Computers & Chemical Engineering*, vol. 168, p. 108040, Dec. 2022. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0098135422003726>
- [14] JJA de Bruin, "Energy-based visualisation of a transcritical CO₂ heat pump cycle," Master's thesis, North-West University (South Africa), 2019.
- [15] H. Naser, G. Van Schoor, K. R. Uren, North-West University (South Africa). Potchefstroom Campus, and South Africa. M-Tech Industrial (Pty) Ltd, "Energy based visualisation of a Brayton cycle power conversion unit for the purpose of condition monitoring," Ph.D. dissertation, North-West University (South Africa), 2019. [Online]. Available: <http://hdl.handle.net/10394/34078>
- [16] Janos J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*, 1st ed. CRC Press, Nov. 2017. [Online]. Available: <https://www.taylorfrancis.com/books/9781351448796>
- [17] Steven X. Ding, *Model-based Fault Diagnosis Techniques*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. [Online]. Available: <http://link.springer.com/10.1007/978-3-540-76304-8>
- [18] R. Isermann and P. Ballé, "Trends in the application of model-based fault detection and diagnosis of technical processes," *Control Engineering Practice*, vol. 5, no. 5, pp. 709–719, 1997.
- [19] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, "A review of process fault detection and diagnosis: Part III: Process history based methods," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 327–346, Mar. 2003. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S009813540200162X>
- [20] T. A. Lemma, *A Hybrid Approach for Power Plant Fault Diagnostics*. Cham: Springer International Publishing, 2018, vol. 743.
- [21] C. Du Rand, G. Van Schoor, and C. Nieuwoudt, "Enthalpy–entropy graph approach for the classification of faults in the main power system of a closed Brayton cycle

- HTGR,” *Annals of Nuclear Energy*, vol. 36, no. 6, pp. 703–711, Jun. 2009. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0306454909000735>
- [22] 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, Feb. 2012. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S030645491100380X>
- [23] H. Naser, G. Van Schoor, and K. Uren, “Energy-based fault detection and isolation of a Brayton cycle-based HTGR power conversion unit – A comparative study,” *Annals of Nuclear Energy*, vol. 164, p. 108616, Dec. 2021. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0306454921004928>
- [24] H.-J. Marais, G. V. Schoor, and K. R. Uren, “An Energy-based approach to condition monitoring of industrial processes,” *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 772–777, 2015. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2405896315017498>
- [25] 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.
- [26] S. Van Graan, “Graph matching as a means to energy visualisation of a counter-flow heat exchanger,” Master’s thesis, North-West University (South Africa), 2017.
- [27] L. B. Fouché, K. R. Uren, and G. v. Schoor, “Energy-based visualisation of an axial-flow compressor system for the purposes of Fault Detection and Diagnosis,” *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 314–319, Jan. 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2405896316305183>
- [28] H.-J. Marais, “Energy-based fault detection for an auto-thermal reformer,” Ph.D. dissertation, North-West University (South Africa), 2016. [Online]. Available: <http://hdl.handle.net/10394/18945>
- [29] 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, Feb. 2019. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S095915241730207X>
- [30] C. Martinson, G. Van Schoor, and K. Uren, “Energy and exergy analyses of a subcritical pulverised coal-fired boiler based on the effects of slagging and fouling,” *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2017–2022, Jul. 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2405896317307590>
- [31] De Bruin, J., Uren, K., Van Schoor, G., and Van Eldik, M., “A thermodynamic cycle model of a transcritical CO₂ heat pump for energy-visualisation.” in *Conference on Condition Monitoring and Diagnostic Engineering Management (COMADEM)*, 2018.

- [32] K. R. Uren, G. van Schoor, M. van Eldik, and J. J. A. de Bruin, “An Energy Graph-Based Approach to Fault Diagnosis of a Transcritical CO₂ Heat Pump,” *Energies*, vol. 13, no. 7, p. 1783, 2020.
- [33] 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 (Basel, Switzerland)*, vol. 21, no. 6, 2019.
- [34] S. Greyling, G. V. Schoor, K. R. Uren, and H. Marais, “Exergy graph-based fault detection and isolation of a gas-to-liquids process,” *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 13 674–13 681, 2020. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2405896320312027>
- [35] J. Vosloo, K. R. Uren, G. Van Schoor, L. Auret, and H. Marais, “Exergy-based fault detection on the Tennessee Eastman process,” *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 13 713–13 720, 2020.
- [36] G. Van Schoor, K. R. Uren, M. A. Van Wyk, P. A. Van Vuuren, and C. P. D. Rand, “An energy perspective on modelling, supervision, and control of large-scale industrial systems: Survey and framework,” *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 6692–6703, 2014. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1474667016426649>
- [37] 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,” *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2842–2847, Jul. 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2405896317310212>
- [38] S. Jouili and S. Tabbone, “Attributed Graph Matching Using Local Descriptions,” in *Advanced Concepts for Intelligent Vision Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, vol. 5807, pp. 89–99, series Title: Lecture Notes in Computer Science. [Online]. Available: http://link.springer.com/10.1007/978-3-642-04697-1_9
- [39] K. R. Uren, J. J. A. De Bruin, G. Van Schoor, M. Van Eldik, and North-West University (South Africa). Potchefstroom Campus, “Energy-based visualisation of a transcritical CO₂ heat pump cycle,” *Energies*, 2020. [Online]. Available: <http://hdl.handle.net/10394/34156>
- [40] R. Styger, K. R. Uren, G. Van Schoor, North-West University (South Africa). Potchefstroom Campus, and National Research Foundation (South Africa), “Graph reduction techniques for exergy-based FDI on the Tennessee Eastman process,” Master’s thesis, North-West University (South Africa), 2022. [Online]. Available: <http://orcid.org/0000-0003-1796-6672>
- [41] W. Wolmarans, G. van Schoor, and K. R. Uren, “Improved energy graph-based visualisation fault detection and isolation — A spectral theorem approach,” *Computers & Chemical Engineering*, vol. 177, p. 108326, Sep. 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0098135423001965>

- [42] S. Greyling, G. Van Schoor, K. R. Uren, H.-J. Marais, and North-West University (South Africa). Potchefstroom Campus, “Graph-based fault detection for a gas-to-liquids process: an exergy approach,” Ph.D. dissertation, North-West University (South Africa), 2021. [Online]. Available: <https://orcid.org/0000-0002-2163-3611>
- [43] N. Sayeekumar, Shafeeqe Ahmed K., S. P. Karthikeyan, S. K. Sahoo, and I. J. Raglend, “Graph theory and its applications in power systems - a review,” in *2015 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*. Kumaracoil, India: IEEE, Dec. 2015, pp. 154–157. [Online]. Available: <https://ieeexplore.ieee.org/document/7475267>
- [44] B. R. Munson, Ed., *Fundamentals of fluid mechanics*, 7th ed. Hoboken, NJ: Wiley, 2013.
- [45] F. M. White, *Fluid mechanics*, 4th ed., ser. Series in Mechanical Engineering. McGraw-Hill Education, 1998. [Online]. Available: <https://www.abebooks.com/Fluid-Mechanics-Mcgraw-Hill-Series-Mechanical-Engineering/31103411797/bd>
- [46] T. J. Kotas, *The exergy method of thermal plant analysis*. Malabar, Florida: Krieger Publishing Company, 1995.
- [47] M. A. Rosen and I. Dincer, “A study of industrial steam process heating through exergy analysis,” *International Journal of Energy Research*, vol. 28, no. 10, pp. 917–930, Aug. 2004. [Online]. Available: <https://onlinelibrary.wiley.com/doi/10.1002/er.1005>
- [48] Y. A. Çengel and M. A. Boles, *Thermodynamics: an engineering approach*, 8th ed. New York, NY: McGraw-Hill Education, 2015.
- [49] M. Mehdizadeh-Fard, F. Pourfayaz, and A. Maleki, “Exergy analysis of multiple heat exchanger networks: An approach based on the irreversibility distribution ratio,” *Energy Reports*, vol. 7, pp. 174–193, Nov. 2021. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2352484720315912>
- [50] B. Wiencke, “Computing the partial volume of pressure vessels,” *International Journal of Refrigeration*, vol. 33, no. 4, pp. 868–876, Jun. 2010. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0140700709002734>
- [51] D. R. Wilson and T. R. Martinez, “Improved Heterogeneous Distance Functions,” *Journal of Artificial Intelligence Research*, vol. 6, pp. 1–34, 1997.
- [52] Clemens Schneider, T. Klette, S. Braun, S. Hartelt, and A. Kratzsch, “Development of Integration Methods for Thermal Energy Storages Into Power Plant Processes,” in *Proceedings of the ASME 2016 10th International Conference on Energy Sustainability*, Charlotte, North Carolina, 2016.
- [53] I. Sarbu, “Solar Heating and Cooling Systems,” in *Advances in Building Services Engineering*. Cham: Springer International Publishing, 2021, pp. 329–445. [Online]. Available: https://link.springer.com/10.1007/978-3-030-64781-0_5

- [54] I. Sarbu and C. Sebarchievici, “A Comprehensive Review of Thermal Energy Storage,” *Sustainability*, vol. 10, no. 1, p. 191, 2018. [Online]. Available: <https://doi.org/10.3390/su10010191>
- [55] T. Klette, S. Braun, A. Kratzsch, and C. Schneider, “Model Development and Dynamic Simulation of a Thermal High-Performance Storage System with the Simulation Code ATHLET to Increase the Flexibility of T,” *International Journal of Electrical and Electronic Engineering & Telecommunications*, pp. 43–50, 2018.
- [56] M. Holmgren, “X Steam for Matlab,” <http://www.x-eng.com>, Jan. 2006, accessed: Jan. 22, 2024.
- [57] Antonio Valiente Barderas, “How to calculate the volumes of partially full tanks,” *International Journal of Research in Engineering and Technology*, vol. 05, no. 04, pp. 1–7, Apr. 2016. [Online]. Available: <https://ijret.org/volumes/2016v05/i04/IJRET20160504001.pdf>
- [58] W. Wolmarans, G. Van Schoor, and K. R. Uren, “Spectral conditioning within the energy graph-based visualisation fault detection method,” *IFAC-PapersOnLine*, vol. 58, no. 4, pp. 330–335, 2024. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2405896324003239>
- [59] A. Turan and A. Muğan, “Structural and sensitivity reanalyses based on singular value decomposition,” *Structural and Multidisciplinary Optimization*, vol. 48, no. 2, pp. 327–337, Aug. 2013. [Online]. Available: <http://link.springer.com/10.1007/s00158-013-0900-2>
- [60] M. Bask, “Dynamic threshold generators for robust fault detection,” Ph.D. dissertation, Luleå tekniska universitet, Luleå, 2005. [Online]. Available: <https://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-26689>
- [61] E. L. Russell, L. H. Chiang, and R. D. Braatz, *Data-driven Methods for Fault Detection and Diagnosis in Chemical Processes*, ser. Advances in Industrial Control, M. J. Grimble and M. A. Johnson, Eds. London: Springer London, 2000. [Online]. Available: <http://link.springer.com/10.1007/978-1-4471-0409-4>