

FAULT DIAGNOSIS & FIELD MEASUREMENT
PREDICTION TECHNIQUES FOR A GAS METERING
SYSTEM

SITI ASFARINA BINTI NIZAMUDDIN

13808

AP IR DR IDRIS ISMAIL

ELECTRICAL & ELECTRONICS ENGINEERING
UNIVERSITI TEKNOLOGI PETRONAS
MAY 2014

SITI ASFARINA

B. ENG. (HONS)

ELECTRICAL & ELECTRONICS ENGINEERING

MAY 2014

**Fault Diagnosis & Field Measurement Prediction Techniques for a Gas
Metering System**

by

Siti Asfarina Binti Nizamuddin

13808

Dissertation submitted in partial fulfilment of
the requirements for the
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)

MAY 2014

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

**Fault Diagnosis & Field Measurement Prediction Techniques
for a Gas Metering System**

by

SITI ASFARINA BINTI NIZAMUDDIN

13808

A project dissertation submitted to the Electrical & Electronics Programme
Universiti Teknologi PETRONAS
in partial fulfilment for the requirement for the
BACHELOR OF ENGINEERING (HONS)
(ELECTRICAL & ELECTRONICS ENGINEERING)

Approved by,

(AP Ir Dr Idris Ismail)

Project Supervisor

UNIVERSITI TEKNOLOGI PETRONAS

TRONOH, PERAK

MAY 2014

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

SITI ASFARINA BINTI NIZAMUDDIN

ABSTRACT

This report discusses on research regarding fault diagnosis system for a process plant. In this project, the process studied is Petronas gas metering system to Kapar Power Plant. There are two parts to this project. The first part is focused on proposing a backup fault diagnosis method for this gas metering system. The second part of the project is to propose suitable field measurement prediction techniques, which could be used in the event of a fault or intermediate condition.

In order to achieve the first objective, this report first discusses the potential fault diagnosis methods which can be applied to the metering system. The advantages and disadvantages of each method were evaluated. From evaluation, it was chosen to propose fault diagnosis system using Adaptive Neuro Fuzzy Inference System (ANFIS). In order to carry out fault diagnosis, data is first filtered into fault data and healthy data. The faults filtered in this report include transmitter fault and hang fault for parameters of Temperature, Pressure and Gross Volume. Once healthy data was identified, it was further classified into normal and intermediate categories. This process was done through three different methods, which are the hyperbox model, linear model and ANFIS model. Once these models were analysed, the writer has chosen to proceed with ANFIS model for data classification. Classified data was then grouped into clusters.

The second part of the project is focused on proposing suitable field measurement prediction technique using ANFIS that can be used in the event of fault or intermediate conditions. Six different ANFIS models were developed to estimate parameters Temperature, Pressure and Gross Volume during transmitter and hang fault. Five variables such as ANFIS input, data division, number of epoch for training, type of membership function and randomisation of data were varied in order to develop the best model. ANFIS prediction model for Temperature produced satisfactory results of less than 1% error. ANFIS prediction model for Pressure and Gross Volume on the other hand need to be further developed to meet industrial requirements.

ACKNOWLEDGEMENT

I would like to take this opportunity to express my greatest gratitude and appreciation towards everyone who contributed towards the completion of this project, be it directly or indirectly. I would like to personally thank the following people whom without, this project would not have been successful:

- AP Ir Dr Idris Ismail (project supervisor)
- Nurfatimah Syalwiah Rosli (masters student)
- Nguyen Tuan Hung (graduate assistant of Universiti Teknologi PETRONAS)
- Information Resource Centre of Universiti Teknologi PETRONAS

Not to forget my parents, friends and family members who acted as my pillars of strength through words of encouragement and motivational support throughout the duration of this project. Your assistance and kindness will not be forgotten.

TABLE OF CONTENTS

CERTIFICATION OF APPROVAL	ii
CERTIFICATION OF ORIGINALITY	iii
ABSTRACT	iv
ACKNOWLEDGEMENT	v
LIST OF FIGURES	vii
LIST OF TABLES	viii
LIST OF ABBREVIATIONS	ix
INTRODUCTION	1
1.1 BACKGROUND	1
1.2 PROBLEM STATEMENT	2
1.3 OBJECTIVE AND SCOPE OF STUDY	2
LITERATURE REVIEW AND THEORY	3
2.1 FAULT DIAGNOSIS	3
2.2 QUALITATIVE MODEL BASED	5
2.3 PROCESS HISTORY BASED	5
2.4 QUANTITATIVE MODEL BASED	7
2.5 COMPARISON OF METHODS	7
2.6 FIELD MEASUREMENT PREDICTION	8
METHODOLOGY	9
3.1 PROCESS MODELLING	9
3.2 DATA FILTERING	11
3.3 FAULT DIAGNOSIS	12
3.4 FIELD MEASUREMENT PREDICTION	16
3.5 PROJECT ACTIVITIES	17
3.6 GANTT-CHART AND KEY MILESTONE.	19
RESULTS AND DISCUSSION	20
4.1 DATA FILTERING	21
4.2 PART I: FAULT DIAGNOSIS	23
4.3 PART III: FIELD MEASUREMENT PREDICTION	33
CONCLUSION AND RECOMMENDATION	45
5.1 CONCLUSION	45
5.2 RECOMMENDATION	46
REFERENCES	47
APPENDICES	49

LIST OF FIGURES

Figure 1: Fault Diagnosis Methods	4
Figure 2: Diagram Representing Neural Network	6
Figure 3: Flow Chart for Methodology	10
Figure 4: Training for Parameter Classification.....	15
Figure 5: Figure for Project Activities	18
Figure 6: Bar Chart of Transmitter Fault Frequency Against Parameter.....	21
Figure 7: Bar Chart of Hang Fault Frequency Against Parameter.....	22
Figure 8: Hyperbox Model to Analyse Temperature	23
Figure 9: Hyperbox Model to Analyse Pressure	24
Figure 10: Hyperbox Model to Analyse Gross Volume.	25
Figure 11: Linear Plot of Pressure against Temperature (Graph 1).....	27
Figure 12: Membership Functions for Parameters.....	29
Figure 13: Classification for Parameters Temperature, Pressure and Gross Volume	30
Figure 14: Class Frequency for Parameters Temperature, Pressure and Gross Volume	31
Figure 15: Histogram Representing Frequency of Each Cluster	32
Figure 16: Chosen Input Combinations for ANFIS Model 1 - ANFIS Model 6	38
Figure 17: Actual Energy against Predicted Energy for ANFIS Model 1- ANFIS Model 6	44

LIST OF TABLES

Table 1: Types of Fuzzy Inference Systems	6
Table 2: Observer Activity During Normal and Abnormal Conditions.....	7
Table 3: Comparison of Different Fault Diagnosis Categories.....	8
Table 4: Comparison of Different Fault Diagnostic Methods	8
Table 5: Parameter and Unit	9
Table 6: Types of Faults in Gas Metering System	11
Table 7: Operating Range Set by PGB	12
Table 8: Limits for Parameters to Draw Hyperbox Model	12
Table 9: Pairings for Linear Model	13
Table 10: Conditions for Classification of Data	14
Table 11: Class Limits According to MF Peak Values.....	15
Table 12: Range for Each Parameter Class.....	16
Table 13: Details of Hang Fault for Temperature.....	22
Table 14: Details of Hang Fault for Pressure.....	22
Table 15: 10% Span for Parameter Values	26
Table 16: Equations for Lower Limit, Upper Limit and Line of Best Fit.....	26
Table 17: Classification of Healthy Data using Linear Model	27
Table 18: Parameter Limits for Temperature.....	28
Table 19: Parameter Limits for Pressure.....	28
Table 20: Parameter Limits for Gross Volume	29
Table 21: Conditions for Data Clustering	32
Table 22: ANFIS Model Number and Predicted Field Measurement.....	33
Table 23: Input Combination for ANFIS Model 1.....	35
Table 24: Input Combination for ANFIS Model 2.....	36
Table 25: Input Combination for ANFIS Model 3.....	36
Table 26: Input Combination for ANFIS Model 4.....	37
Table 27: Input Combination for ANFIS Model 5.....	37
Table 28: Input Combination for ANFIS Model 6.....	37
Table 29: Selection of Data Division for ANFIS Model 1-ANFIS Model 6.....	39
Table 30: Selection of Number of Epoch for ANFIS Model 1-6.....	40
Table 31: Selection of MF Type for ANFIS Model 1-6	41
Table 32: MAPE for Randomised and Sequential Data ANFIS Model 1-6	42
Table 33: Developed Models for Field Measurement Prediction	43

LIST OF ABBREVIATIONS

PGB	PETRONAS Gas Berhad
GTS	Group Technical Solutions
P	Pressure
T	Temperature
V_g	Gross volume
CV	Calorific value
S_g	Standard gravity
ANFIS	Adaptive neuro fuzzy inference system
E	Energy
E_h	Energy consumption per hour
V_h	Volume of gas per hour
V_{g(meas)}	Measured gross volume
Z_{meas}	Measured compressibility factor
Z_{base}	Base compressibility
P_{base}	Base pressure
T_{meas}	Measured temperature
T_{base}	Base temperature
P_{meas}	Measures pressure
MAPE	Mean Average Percentage Error
P_(UL)	Higher operating limit for parameter
P_(LL)	Lower operating limit for parameter
FIS	Fuzzy Inference System

CHAPTER 1

INTRODUCTION

1.1 Background

We are surrounded by many different types of systems in our daily lives. Some of these systems include physical systems, engineering systems and organic systems. Under abnormal conditions, such as faults, each of these systems can be subjected to failure. According to Korbiczin, in engineering terms, these abnormalities are referred to as faults. Faults in an engineering system can cause the process to shut down, jeopardising the economic performance of the system. It can also pose a threat towards the safety of the surrounding personnel [1]. The purpose of this paper is to develop a fault diagnosis method for the gas metering system to Kapar power plant along with field measurement prediction techniques.

1.1.1 Fault Diagnosis

Fault diagnosis refers to the prompt identification and analysis of system abnormalities. The early detection of these faults is important to ensure reliability, safety and efficiency of the process [2]. There are many different types of fault diagnoses that have been applied in the industry. These diagnoses are divided into three main models which are quantitative models, qualitative models and process history based models. Each category is further divided into different diagnostic methods [3]. In this project, fault diagnosis is focused on two different types of faults, which are hang fault and transmitter fault. The diagnosis of this project covers three different system parameters, which are Temperature, Pressure and Gross Volume.

1.1.2 Gas Metering System

It is of utmost importance that gas meters are calibrated accurately, especially in the oil and gas industry [4]. The gas metering system to Kapar Power Plant consists of a double piping system with records of two runs for each pipeline. The transmitters used include the temperature transmitter, flow transmitter and pressure transmitter [5]. The measurements are used to compute the energy that is supplied by Petronas. Customers are billed based on the calculation of energy. The possible types of fault identified for this system include transmitter fault, hang fault and drift fault.

1.2 Problem Statement

It is recommended that all engineering processes be equipped with fault diagnosis system for early detection of faults. The metering system to Kapar power plant is currently in need of a backup fault diagnosis system. In the event that a fault happens at this metering system, the system may produce some inaccurate measurements. These measurements may affect the energy consumption calculation and hence, degrade the integrity of the billing system for customers. It is therefore recommended that fault diagnosis method and field measurement prediction techniques be developed for this gas metering system.

1.3 Objective and Scope of Study

The main objective of this paper is divided into four different sections. The divisions are as follows:

- To study potential fault diagnosis methods which can be proposed for metering system to Kapar power plant.
- To evaluate and compare several known fault diagnosis methods which can be applied for the metering system.
- To propose implementation for a trial basis a suitable fault diagnosis method for the metering system.
- To develop model for field measurement prediction which may be used in the event that a fault occurs.

CHAPTER 2

LITERATURE REVIEW AND THEORY

This chapter is divided into four sections which are fault diagnosis, quantitative model, qualitative model and process history based model. The breakdown of the different fault diagnosis methods are explained in Figure 1. As can be seen from this figure, these three sections of diagnoses are further divided into different methods. Each of these diagnosis methods has their respective advantages and disadvantages.

2.1 Fault Diagnosis

This section of the paper is dedicated to giving an overview of fault diagnosis systems, the importance of fault diagnosis and the different types of fault diagnosis techniques that can be applied in engineering systems. According to Merriam Webster's dictionary, fault can be defined as a problem or a defect. Diagnosis is defined as the act of studying something or someone in order to identify a problem. System is an assembly of interconnected components that work with each other [6]. Fault diagnosis also refers to timely detect and diagnose faults [7].

The importance of fault diagnosis systems can be divided into two categories, which is the importance towards complex safety-threatening systems and smaller non-safety-threatening systems [6]. In the former category, fault diagnosis is aimed to prevent failure of system which can jeopardise the safety of humans, create negative impacts towards the environment and cause monetary loss. In the latter category, fault diagnosis aims to improve reliability and efficiency of the system [8].

There are three main components that make up fault diagnosis, which are fault detection, fault isolation and fault identification. Fault detection is sensing the presence of a system abnormality. Fault isolation refers to concluding the position of the fault while fault identification refers to assessing fault type [7]. There are many different types of fault diagnosis techniques that can be used in a system. These

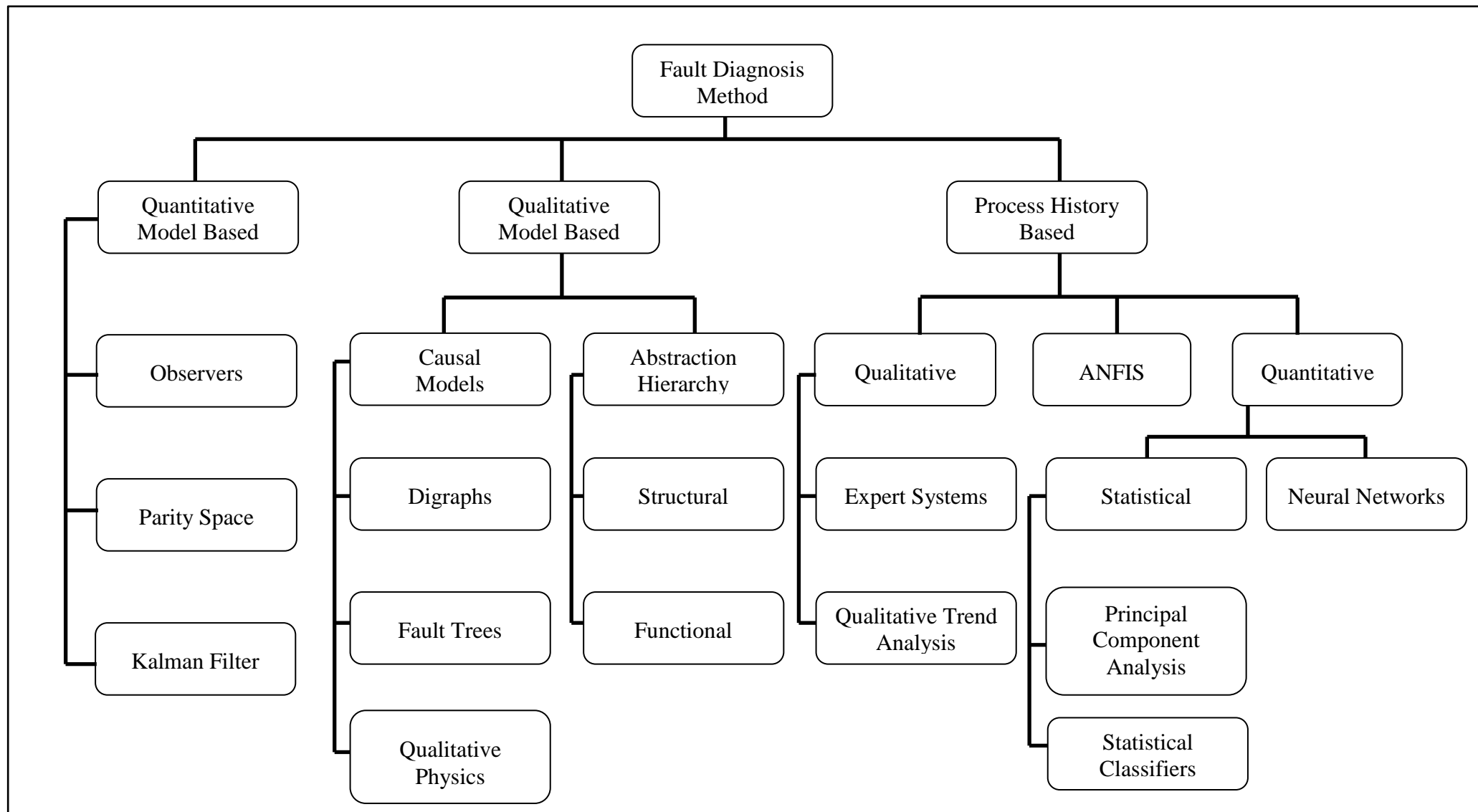


Figure 1: Fault Diagnosis Methods [7]

diagnoses are divided into three different models which are quantitative model based, qualitative model based and process history based [8].

2.2 Qualitative Model Based

In this modelling method, the inputs and outputs of the system are expressed in terms of qualitative functions. According to Isermann, this model utilises static and dynamic relations within system variables to explain system performance in qualitative expressions [9]. Physical or chemical information regarding the process is required to develop these functions. There are two methods to obtain the priori knowledge which are through causal models or abstraction hierarchy [10]. In this paper, diagnosis of digraphs (causal model) is discussed.

Signed Digraphs (SDG) is a causal model for fault diagnosis. This diagnosis uses cause-effect method to extract data. In this method, a mathematical model is first formed. It is then represented as a graph using a system of arcs [11]. The graph explains the path of fault.

2.3 Process History Based

This modelling method differs from qualitative model and quantitative model in the sense that instead of relying on priori knowledge, this diagnosis is based on historical process data. Through a process called feature extraction, the extracted process data is modified to become priori knowledge [13]. Isermann described this modelling method as relations which connect symptoms (input variables) to faults (output variables) [9]. In this paper, the diagnosis of artificial neural networks which is a quantitative, non-statistical approach is discussed. The next method discussed is adaptive neuro fuzzy inference system (ANFIS), which has a combination of qualitative and quantitative properties.

Artificial neural networks are systems that are similar to the connection in the brain. It is a learning system, constructed from processing elements (PE). This method is able to estimate non-linear relations [2]. According to Principe in his book titled *Neural and Adaptive Systems*, PEs are connected and transmit data to one another, similar to neurons in the brain. Topology of the system is determined by the connection of the PEs. Weights, w_{ij} along the connections are able to scale the transmitted signals. PEs analyse the received signals and generate functions, f . These functions are transmitted to other PEs or are translated as the system output [14].

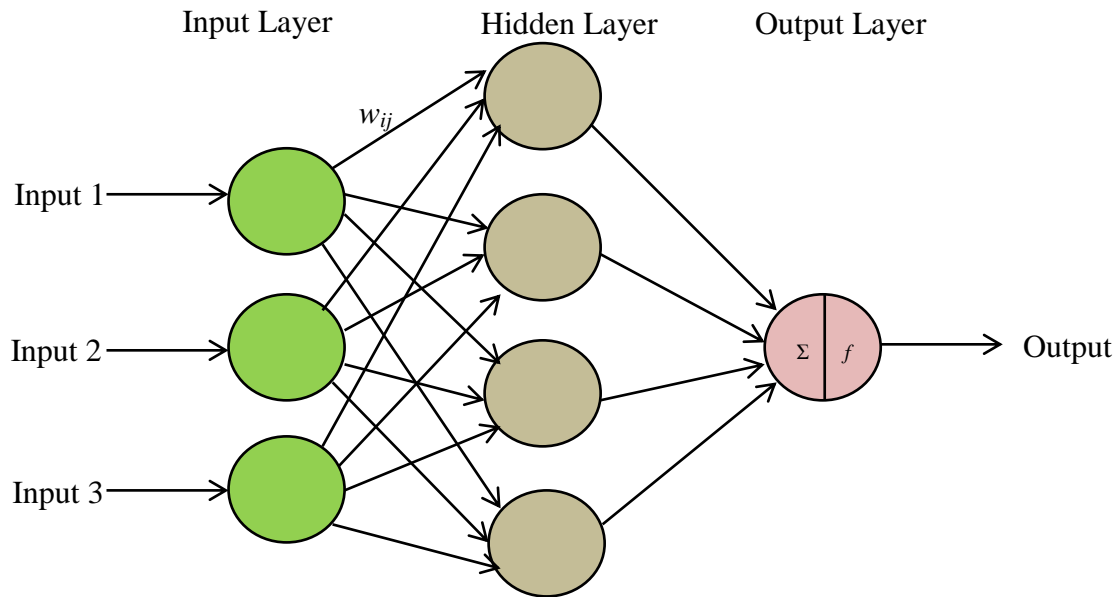


Figure 2: Diagram Representing Neural Network [14]

The second method that was researched is the adaptive neuro fuzzy inference system, (ANFIS). ANFIS is a hybrid type of artificial intelligence that combines properties of both fuzzy inference system (FIS) with adaptive neural networks. There are three different types of fuzzy logics which can be used, which are Type 1, Type 2 and Type 3. A description of these three inference systems are given in Table 1.

Table 1: Types of Fuzzy Inference Systems [15]

FIS	FIS Operation
Type 1	<ul style="list-style-type: none"> - To calculate output, mean of output from rules and output from MF is calculated. - Monotonic function used as output MF.
Type 2	<ul style="list-style-type: none"> - To calculate output, minimum firing strength and output MF is analysed
Type 3	<ul style="list-style-type: none"> - Developed by using Sugeno's rules. - Output is calculated as the mean of each rules' output

An adaptive network operates using the principle of gradient descent. System learning depends on training data, in which during training, the parameters of each node are varied to achieve the smallest error. [15]

2.4 Quantitative Model Based

Quantitative models are models in which the inputs and outputs of the system are expressed in mathematical equations. The models are developed based on priori knowledge of the system. This method consists of two main parts which are residual generation and decision making. Residuals are signals generated based on the input and output signals [2]. These signals test for the presence of faults. In the next step which is the decision making step, an analysis is then conducted in the residuals to decide whether a fault is present [7]. The qualitative method that is analysed and compared in this paper is the observer method.

In the observer method, robust residuals which are able to distinguish faults are developed. There are two types of observers which can be applied, which are state observers and output observers [2]. Observers which are specific in their response towards different faults are developed. The behaviour of the observers and residuals indicate the presence of fault [7]. This behaviour is detailed within Table 2.

Table 2: Observer Activity During Normal and Abnormal Conditions

Condition	Observer activity
Normal condition	Process is followed by observer and residual magnitude is small.
Abnormal condition (fault)	Value of residual increases and observer is unable to follow the process.

2.5 Comparison of Methods

Within Table 3, the quantitative model based diagnosis, qualitative model based diagnosis and process history based diagnosis are compared. From this table of comparison, the writer has chosen to use the process history based diagnosis. This is because it depends on historical data rather than priori knowledge. It does not require heavy modelling, making it easy to implement. Besides that, this method of diagnosis is the most commonly used in the industry. Although the model is unable to adapt to different conditions, it is able to respond to a set of data classified as faults. Therefore, the scope of this report is limited to the sample training data.

Table 3: Comparison of Different Fault Diagnosis Categories

Diagnosis Algorithm	Advantages	Reference	Disadvantages	Reference
Quantitative Model Based	Fault isolation possible	[9]	Complex mathematical model	[7]
	Able to detect multiple faults		Constricted to linear models	
Qualitative Model Based	Robust	[10]	Unable to diagnose multiple faults	[10]
	Able to detect fault path		Unable to adapt to different conditions	
Process History Model	Robust	[13]	Unable to diagnose multiple faults	[15]
	Little modelling required (easy to implement)		Unable to adapt to different conditions outside sample data	
	Require little priori knowledge			

There are many different criteria that indicate a good fault diagnosis method. In Table 4, the observer method, digraphs method and ANFIS method are compared in terms of favourable diagnostic method criteria. The 'x' represents that the diagnosis method meets the respective diagnosis criteria.

Table 4: Comparison of Different Fault Diagnostic Methods [13]

Diagnosis Criteria	Observer	Digraphs	ANFIS
Early detection and diagnosis	x		X
Isolable	x		X
Robust	x	x	X
Novelty identifiability		x	X
Adaptable		x	
Multiple fault analysis	x	x	

2.6 Field Measurement Prediction

Field measurement refers to values obtained from instruments on site. ANFIS is one of the methods which can be used to predict field measurements. According to a research conducted in Marmara University Turkey, ANFIS model was used to forecast weather. In order to achieve this, parameters of air pressure, temperature and wind speed were used as ANFIS input to predict weather [16]. Another research conducted by University of Medea Algeria used ANFIS model of inputs of mean sunshine duration and air temperature to predict solar radiation [17]. Similarly, this report explores the suitability of ANFIS model to predict field measurements of Temperature, Pressure and Gross Volume for a gas metering system.

CHAPTER 3

METHODOLOGY

In this section of the report, the methodology to conducting this project is explained. It is divided into four stages which are process modelling, data filtering, fault diagnosis and field measurement prediction. The flow chart to represent project methodology is demonstrated in Figure 3.

3.1 Process Modelling

In process modelling, data is collected from the gas metering system. Data was collected from Petronas Gas Berhad (PGB) ranging from 24th May 2013 – 20th September 2013. As many as 2880 data points were collected on an hourly basis, comprising of the parameters detailed within Table 5.

Table 5:Parameter and Unit

Parameter	Unit
Date	-
Time	Hour
Temperature (T)	°C
Pressure (P)	kPag
Calorific Value (CV)	MJ/Sm ³
Standard Gravity (sg)	-
Gross Volume (Vg)	m ³
Standard Volume	Sm ³
Energy (E)	GJ

Next, an energy production model from a previous research by Maryam Jamela [18] was applied. The inputs to this model are temperature, pressure, gross volume and calorific value. Using these input values, the model will then produce an output of energy.

The model is as follows is demonstrated in Eq. 1 and Eq. 2.

$$V_h = \sum_h \frac{V_{g(meas)} \times P_{meas} \times T_{base} \times Z_{base}}{P_{base} \times T_{meas} \times Z_{meas}} \quad \text{Eq. 1}$$

$$E_h = \sum_h \frac{V_h \times CV}{1000} \quad \text{Eq. 2}$$

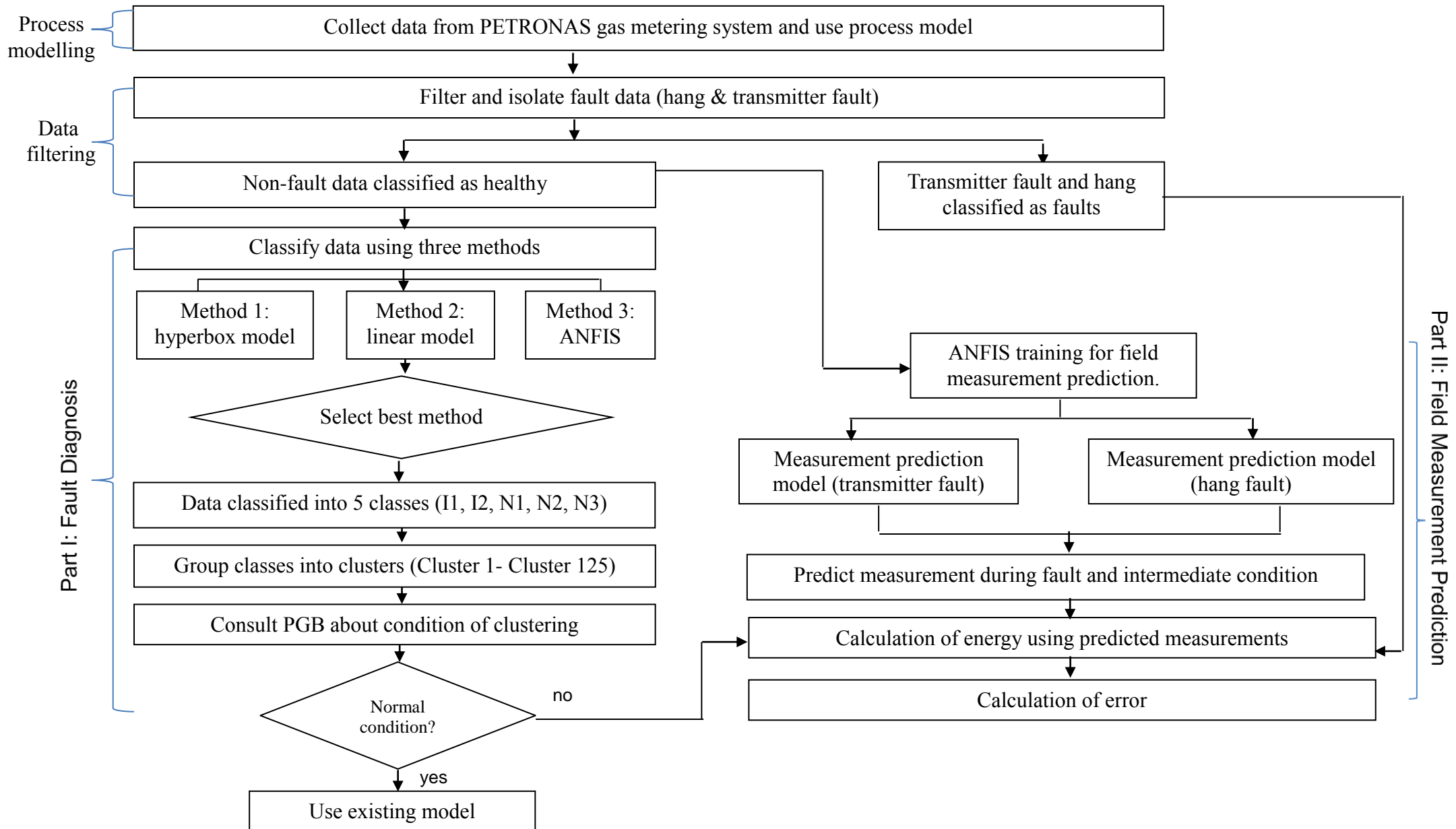
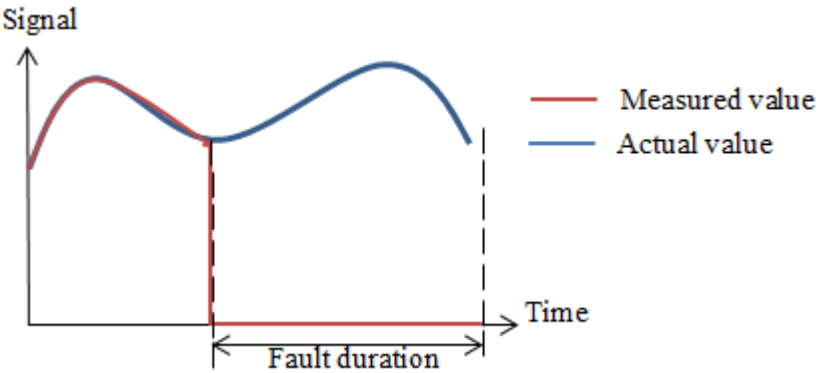
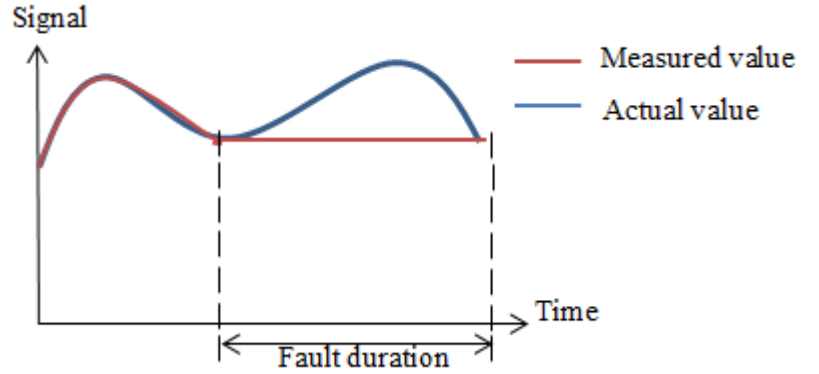
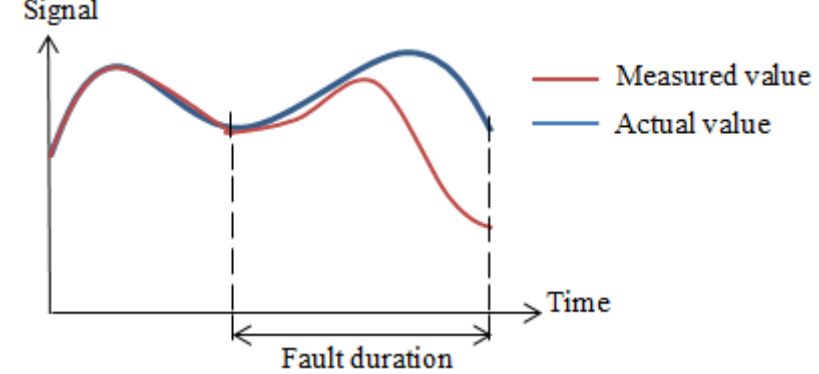


Figure 3: Flow Chart for Methodology

3.2 Data Filtering

The data is first filtered into healthy and fault data. Fault data is data which falls within the transmitter fault and hang region while healthy data is the remaining data. Transmitter fault refers to the condition when transmitter reads zero continuously while hang refers to the condition when the reading of the transmitter displays the same value continuously. In this report, repetitions of two or more (continuous reading for more than one hour) are considered as hang fault. A detailed explanation of these faults is presented in Table 6. Future works can be done to cater for drift faults.

Table 6: Types of Faults in Gas Metering System [5]

Fault	Trend of Fault
<p>Transmitter Fault: Zero reading is displayed by transmitter.</p>	 <p>The graph plots Signal on the y-axis and Time on the x-axis. A blue line represents the 'Actual value', which fluctuates normally. A red line represents the 'Measured value', which follows the actual value until it drops to zero at the start of the 'Fault duration' (indicated by a vertical dashed line). It remains at zero until the end of the 'Fault duration' (indicated by another vertical dashed line), after which it returns to the actual value.</p>
<p>Hang Fault: System hang causing transmitter to display constant reading.</p>	 <p>The graph plots Signal on the y-axis and Time on the x-axis. A blue line represents the 'Actual value', which fluctuates normally. A red line represents the 'Measured value', which follows the actual value until it becomes constant at a non-zero value at the start of the 'Fault duration' (indicated by a vertical dashed line). It remains constant until the end of the 'Fault duration' (indicated by another vertical dashed line), after which it returns to the actual value.</p>
<p>Drift Fault: Reading increasingly deviates from the true value. The error increases with increasing time.</p>	 <p>The graph plots Signal on the y-axis and Time on the x-axis. A blue line represents the 'Actual value', which fluctuates normally. A red line represents the 'Measured value', which follows the actual value until the start of the 'Fault duration' (indicated by a vertical dashed line). During the fault duration, the measured value increasingly deviates from the actual value, showing a downward trend. It returns to the actual value at the end of the 'Fault duration' (indicated by another vertical dashed line).</p>

Data was also filtered according to operating range set by PGB. Data outside of this operating range was classified as outliers. The operating range is detailed in Table 7.

Table 7: Operating Range Set by PGB

Parameter	Lower Operating Limit $P_{(LL)}$	Higher Operating Limit $P_{(UL)}$
T (°C)	20	35
P (kPag)	2000	6500
CV (MJ/Sm ³)	35	42
Sg	0.5	0.75
Vg (m ³)	0	4500
E (GJ)	0	6500

3.3 Fault Diagnosis

Once the data has been filtered, it is then classified as normal and intermediate conditions. This was done using three methods, which are the hyperbox model, the linear model and ANFIS.

3.3.1 Methodology for Hyperbox Model

The hyperbox model was developed by setting the box limits to conform to a set hyperbox limit for each parameter, as demonstrated in Table 8. These values were chosen based on analysis of training data. The data points within the hyperbox limit are classified as normal while data points out of the hyperbox limit are classified as outliers. These limits are shown as blue lines within Figure 8 to Figure 10.

Table 8: Limits for Parameters to Draw Hyperbox Model

Parameter	Lower Limit	Upper Limit
Temperature (°C)	25	30
Gross Volume (m ³)	300	600
Pressure (kPag)	3000	6500

Before the values are plotted, they must first be normalised. Normal conditions fall between the normalised range of 1 to -1. In order to achieve this, the limits in Table 8 were scaled down. The lower limits were scaled down to -1 while the upper limit was scaled down to 1. The equations used for the normalisation process are shown as Eq. 3 to Eq. 5.

Range of Parameter Value	Scaling Equation	Eq.
Lower Limit – Upper Limit	$1 - \left(\frac{\text{Upper Limit} - \text{parameter value}}{\text{Upper Limit} - \text{Lower Limit}} \times 2 \right)$	3
\geq Upper Limit	$1 + \left(\frac{\text{parameter value} - \text{Upper Limit}}{\text{Upper Limit} - \text{Lower Limit}} \times 2 \right)$	4
\leq Lower Limit	$-1 - \left(\frac{\text{Lower Limit} - \text{parameter value}}{\text{Upper Limit} - \text{Lower Limit}} \times 2 \right)$	5

In order to plot the points in the hyperbox, the normalised data points are first scanned to determine which points fall within the acceptable range and which points are in intermediate range. The acceptable range for normalised data is from -1 to 1. Data points outside of the acceptable range are considered as outliers.

3.3.2 Methodology for Linear Model

The second method used for fault diagnosis was to classify data into normal and intermediate conditions using the linear model. From this model, the data was classified into four different categories, which were Normal 1 (N1), Normal 2 (N2), Intermediate 1 (I1) and Intermediate 2 (I2). The first step in order to construct this model was to pair the five parameters to be plotted on an x-y coordinate. The pairings were chosen as demonstrated in Table 9.

Table 9: Pairings for Linear Model

Graph	Y-Axis	X-Axis
1	Pressure (kPag)	Temperature (°C)
2	Calorific Value (mJ/Sm ³)	Temperature (°C)
3	Gross Volume (m ³)	Temperature (°C)
4	Standard Volume (Sm ³)	Temperature(°C)
5	Calorific Value (mJ/Sm ³)	Pressure (kPag)
6	Gross Volume (m ³)	Pressure (kPag)
7	Standard Volume (Sm ³)	Pressure (kPag)
8	Calorific Value (mJ/Sm ³)	Gross Volume (m ³)
9	Calorific Value (mJ/Sm ³)	Standard Volume (Sm ³)
10	Gross Volume (m ³)	Standard Volume (Sm ³)

Once the pairings were chosen, the graphs were plotted with line of best fit displayed on the plot.

The next step was to find the lower and upper limit from the line of best fit. This was done by finding a variation of 10% from the range of data. The 10% from range values were used to determine the lower limit and upper limit of the plot. The equations used to determine the lower limit and upper limit are demonstrated as Eq. 6 and Eq. 7:

Limit Line	Equation of Line	Eq.
Lower Limit Line	Line of best fit + 10% of Range	6
Upper Limit Line	Line of best fit – 10% of Range	7

Lower Limit Line, Upper Limit Line, $P_{(LL)}$ and $P_{(UL)}$ were the limits used to classify data into four different classes. The classification conditions are demonstrated in Table 10. Using this classification method, each data point was classified as either as I1, I2, N1 or N2.

Table 10: Conditions for Classification of Data

Classification	Lower Limit	Upper Limit
Intermediate 1 (I1)	$P_{(LL)}$	Lower Limit Line
Normal 1 (N1)	Lower Limit Line	Line of Best Fit
Normal 2 (N2)	Line of Best Fit	Upper Limit Line
Intermediate 2 (I2)	Upper Limit Line	$P_{(UL)}$

3.3.3 Methodology for ANFIS

The final method used for data classification was the ANFIS method. Using this method, the ANFIS model was trained using previous parameter values as input and current parameter values as output. Figure 4 shows the input-output relationship used for training of ANFIS model. The membership function was then obtained by setting the number of membership functions to 4. The FIS was trained using hybrid method of 60 epochs. Using ANFIS, data was classified into five categories, which are Normal 1 (N1), Normal 2 (N2), Normal 3 (N3), Intermediate 1 (I1) and Intermediate 2 (I2).

Before the data are input into the ANFIS training model, they are first normalised between the limits of zero and one. The normalization process was done using Eq. 8.

$$Data(n)_{(normalised)} = \frac{Data(n) - P_{(LL)}}{P_{(UL)} - P_{(LL)}} \quad \text{Eq. 8}$$

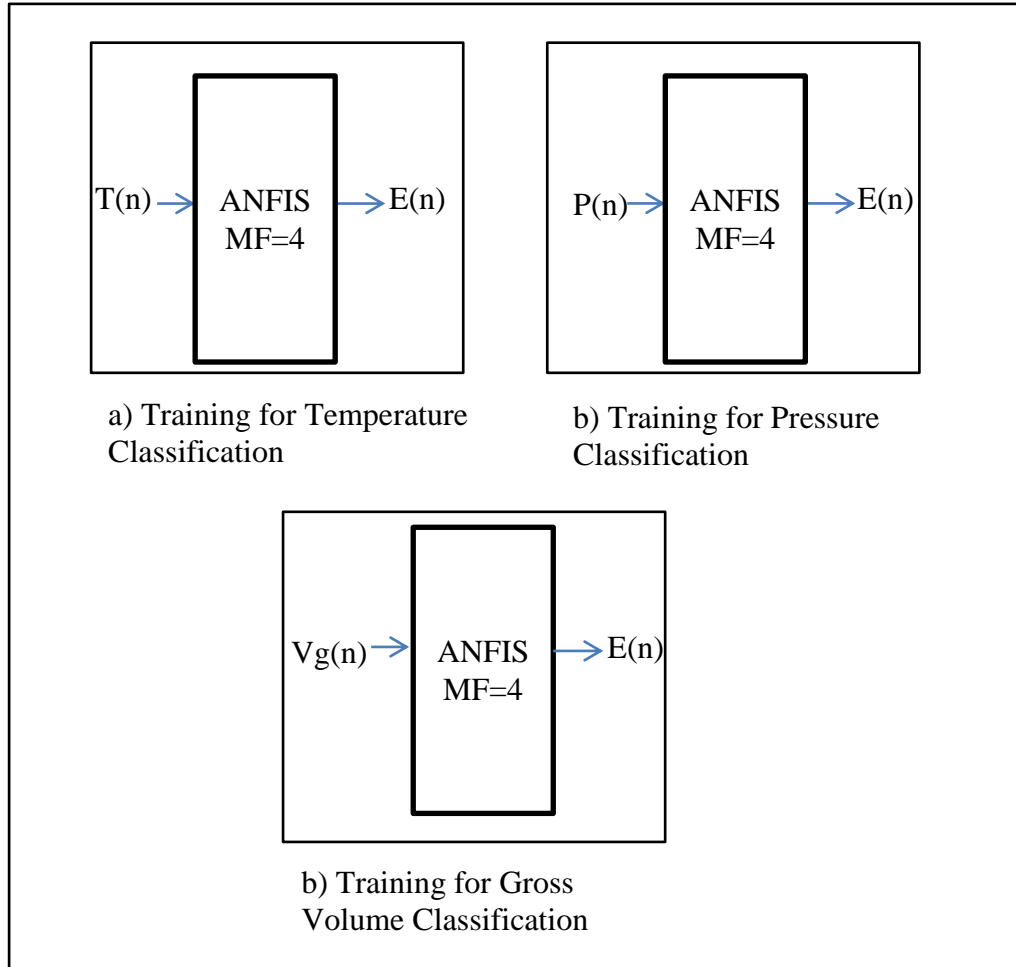


Figure 4: Training for Parameter Classification

Four membership functions were generated for each of the three ANFIS training models. The normalised peaks of each membership function were noted and denormalised back to normal peak values according to Eq. 9:

$$Peak = Parameter_{LL} + Peak_{(normalised)} \times (Parameter_{UL} - Parameter_{LL}) \quad \text{Eq. 9}$$

The class limits were then assigned as peak values according to Table 11.

Table 11: Class Limits According to MF Peak Values

Class Limit	MF Peak
Class Limit 1	1 st peak
Class Limit 2	2 nd peak
Class Limit 3	3 rd peak
Class Limit 4	4 th peak

The final step for data classification is to assign parameter classes according to the limit. The parameter classes were set according to the following ranges in Table 12.

Table 12: Range for Each Parameter Class

Parameter Class	Lower Limit	Upper Limit
I1	$P_{(LL)}$	Class Limit 1
N1	Class Limit 1	Class Limit 2
N2	Class Limit 2	Class Limit 3
N3	Class Limit 3	Class Limit 4
I2	Class Limit 4	$P_{(UL)}$

Parameters of Temperature, Pressure and Gross Volume are classified into classes I1, I2, N1, N2 and N3 following conditions set in Table 12.

These three methods are then compared and the best method of data classification is chosen. Once data has been classified, it is further grouped into clusters. The condition for clustering is demonstrated in Appendix B. The condition for each cluster is then identified as either Normal or Intermediate. This project proposes that ANFIS parameter prediction model be used for clusters that fall within the Intermediate range.

3.4 Field Measurement Prediction

In order to predict field measurements, the artificial intelligence ANFIS was used. ANFIS model was developed to predict parameters of Temperature, Gross Volume and Pressure during fault conditions.

Data filtered as healthy is used in this step. This data is first normalised from 0 to 1. The reason for normalisation is because data of different ranges and units are input into ANFIS. In order to develop accurate membership functions, the data should have a standardised unit of values. The normalisation was done according to Eq. 8.

It is then divided into training data and checking data. Training data is used to train the ANFIS model while checking data is presented as input to the model in order to calculate the percentage error. Two ANFIS models were trained for each parameter, which are the ANFIS model for parameter prediction during hang fault and the ANFIS model for parameter prediction during transmitter fault.

In order to determine the best model for each parameter, five variables were manipulated which are input, percentage of data division into training and checking, epoch number for training, type of membership function and randomisation of data before input into ANFIS. The model with the lowest mean average percentage error (MAPE) was selected as the best model. The equation used in order to calculate the MAPE is shown as Eq.10. These models were developed in order to predict field

measurements during fault and intermediate conditions. The predicted parameters values are then used to calculate energy using Eq. 1 and Eq. 2.

$$MAPE = \sum_h \frac{Value_{(predicted)} - Value_{(measured)}}{Value_{(measured)}} \times 100 \quad \text{Eq. 10}$$

3.5 Project Activities

Project activities refer to the sequencing of tasks that need to be completed in order to achieve the project objectives. Figure 5 represents the project activities carried out. It starts with studying on various fault diagnosis methods and ends with calculation of error for the developed ANFIS model.

The first activity which was to study various fault diagnosis methods was conducted using books, thesis papers and journals from credible sources. The various methods were studied and compared in order to choose the most suitable method for this project. The next step was to collect data from the gas metering system. Data to be collected include temperature ($^{\circ}\text{C}$), pressure (kPa), calorific value (MJ/Sm^3), standard gravity, gross volume (m^3), and energy (GJ). Data was then filtered into fault and healthy data. Fault data such as transmitter fault and hang were filtered as faults while the remaining data was considered as healthy data.

Once the data has been filtered, the healthy data was analysed and classified. In order to carry out this activity, healthy data was classified as normal and intermediate conditions. This was done using three methods which were the hyperbox model, linear model and ANFIS. These three methods were evaluated in order to determine the most suitable model.

Once the data has been classified into normal and intermediate conditions, they are further clustered according to conditions of parameters occurring simultaneously. Each cluster is then analysed to determine if it is a normal or intermediate cluster. This project proposes that intermediate and fault measurements be predicted using ANFIS field measurement prediction model.

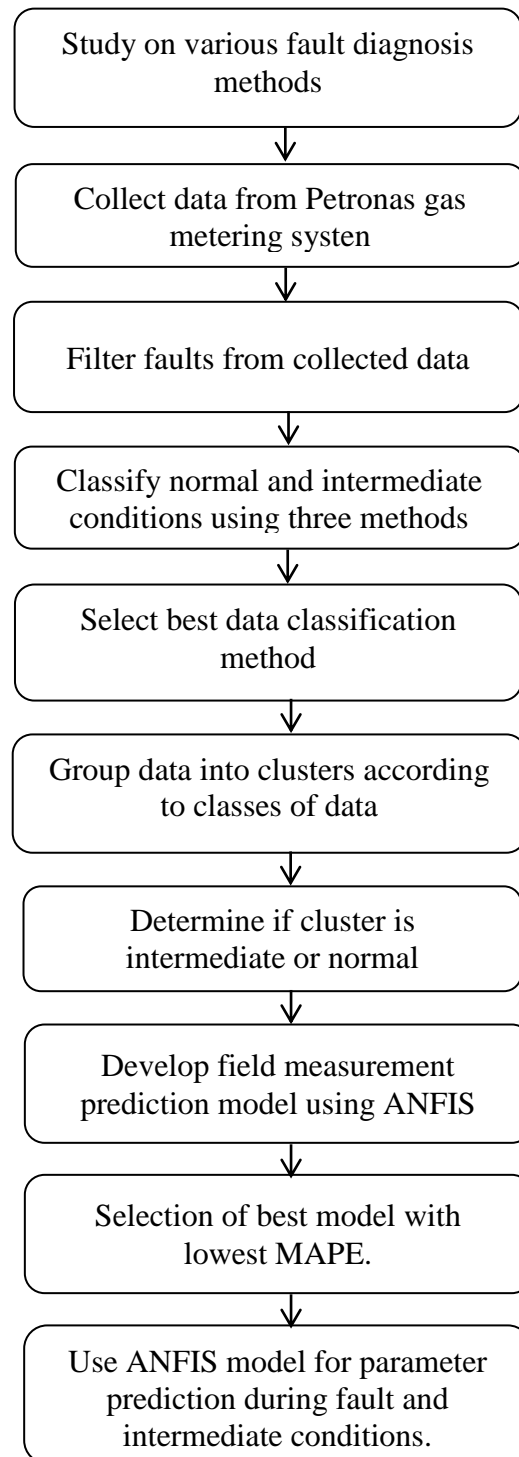


Figure 5: Figure for Project Activities

The next step was to develop a field measurement prediction model. A total of six ANFIS evaluation models were developed to predict field measurements during transmitter and hang faults. The parameters that were predicted are Temperature, Pressure and Gross Volume during transmitter and hang faults. The parameters predicted are then presented into the energy calculation equation in order to determine the best model. The ANFIS model which registered the lowest MAPE is considered to be the best model.

In this project, it is proposed that fault data and data which falls within the intermediate range are predicted using the field measurement prediction model. This model can be used as a backup in case of the above mentioned conditions.

3.6 Gantt-Chart and Key Milestone.

The gantt-chart and key milestones have been included in the appendices section. Table A.1 represents the activities during FYP 1 while Table A.2 represents the activities to be conducted during FYP 2.

CHAPTER 4

RESULTS AND DISCUSSION

This section of the report presents and discusses the findings of this project. The findings are divided into three sections which are data filtering, fault diagnosis, and field measurement prediction. In the data filtering section, the results for filtering of transmitter fault and hang fault are discussed. An explanation of transmitter faults, hang faults and drift faults is given in Table 6. The frequency and time of fault occurrence are analysed in Table 13 and Table 14.

Within the fault diagnosis section, results for the three methods of data classification which are hyperbox model, linear model and ANFIS are presented. From the results obtained, ANFIS was chosen as the most suitable data classification method. Five classes were generated for each parameter, which are Normal 1 (N1), Normal 2 (N2), Normal 3 (N3), Intermediate 1 (I1), and Intermediate 2 (I2).

The next step taken was to cluster the data according to class. This was done by analysing the class of parameters Temperature, Pressure and Gross Volume simultaneously. A total of 125 clusters were generated from this conditioning. Each cluster was identified as either a normal or intermediate cluster. Analysis of clustering results is presented in this chapter.

The final result presented in this section of report is the results for field measurement prediction. This process was conducted using ANFIS prediction model. In order to determine the best model, five variables were manipulated, which are ANFIS input, percentage of data division into training and checking, number of epoch for training, type of membership function and randomization of data. The best model was chosen based on the model which registers the lowest MAPE.

4.1 Data filtering

Data filtering is the step conducted in order to divide the data into fault and healthy data. The first step conducted was to filter the data according to operating range set by PGB. Data outside of this operating range was classified as outliers.

The next step conducted was to filter the data into healthy and fault data. Parameters Temperature, Pressure and Gross Volume were scanned for fault data. The remaining data was considered as healthy data. From the sample of 2880 data points, 2501 data points were healthy.

In order to isolate and filter transmitter fault, the five parameter values were scanned for a value of zero. It was observed that the percentage of transmitter faults identified was very small. Temperature, Pressure and Calorific Value recorded two transmitter fault data values, which corresponds to 0.069 %. Gross Volume and Standard Gravity on the other hand recorded nine transmitter faults, which corresponds to 0.313%. These data are represented within Figure 6.

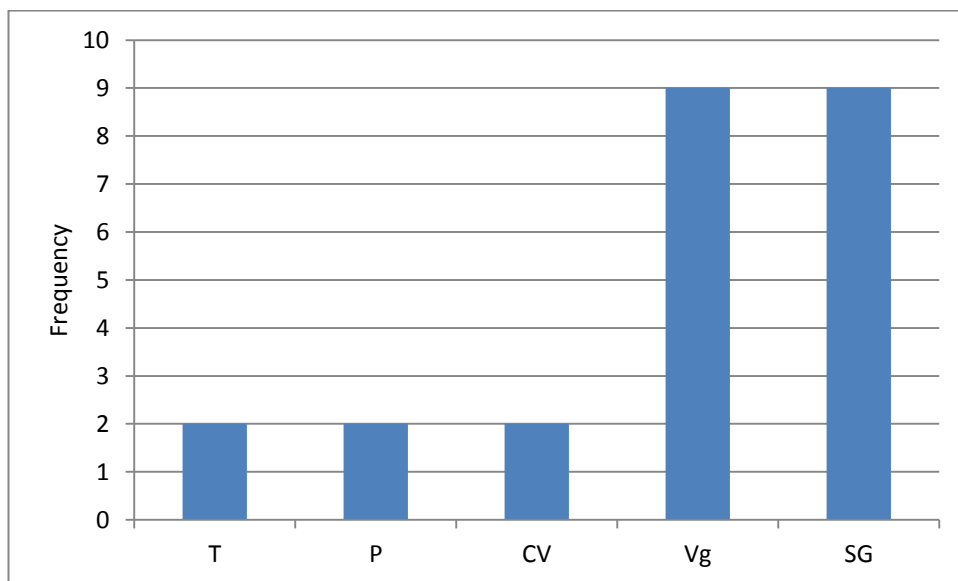


Figure 6: Bar Chart of Transmitter Fault Frequency Against Parameter

After the transmitter fault has been filtered, hang faults were filtered. In order to filter hang faults, parameters Pressure, Temperature and Gross Volume were scanned for repetitive readings. From the collected data, the number of hang faults identified for temperature and pressure parameters are 359. It was observed that both these data experienced hang fault simultaneously. This amounts to 12.47% of the data. There were no hang faults detected within Gross Volume parameter. Figure 7 represents the results collected for hang fault.

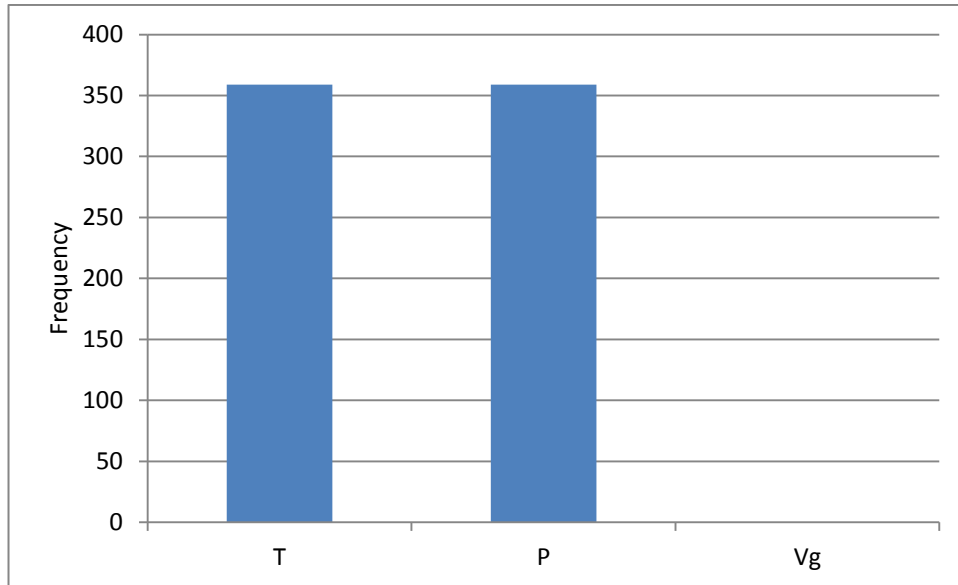


Figure 7: Bar Chart of Hang Fault Frequency Against Parameter

From parameters Pressure and Temperature, there were four sets of data repetition found. These data are represented in Table 13 and Table 14.

Table 13: Details of Hang Fault for Temperature

Set	No. of repetitions	Parameter value (°C)	Start time of Hang Fault	Stop time of Hang fault
1	3	28.39897	12:00, 15/08/2013	02:00, 08/15/2013
2	5	28.3506	01:00, 17/08/2013	05:00, 17/08/2013
3	47	28.25176	07:00, 18/08/2013	05:00, 19/08/2013
4	267	28.25177	07:00, 20/08/2013	09:00, 31/08/2013

Table 14: Details of Hang Fault for Pressure

Set	No. of repetitions	Parameter value (kPag)	Start time of Hang Fault	Stop time of Hang fault
1	3	5578.379	12:00, 15/08/2013	02:00, 08/15/2013
2	5	5703.219	01:00, 17/08/2013	05:00, 17/08/2013
3	24	5437.455	07:00, 18/08/2013	06:00, 18/08/2013
4	290	28.25177	07:00, 18/08/2013	09:00, 31/08/2013

4.2 Part I: Fault Diagnosis

Once the data has been filtered as fault and healthy data, the next step is to classify the data into normal and intermediate conditions. Three methods were used for this classification, which are the hyperbox model, linear model and ANFIS model. ANFIS model was chosen as the best classification method. The classified data are then further clustered into 125 clusters in data clustering step. Each cluster is identified as either a normal or intermediate cluster.

4.2.1 Hyperbox Model

Using hyperbox model, data is classified into normal and outliers. The first step conducted to draw the hyperbox was to set the box limits. These limits are stated in Table 8 and are shown as blue lines within Figures 8 to Figure 10. Data points outside of the acceptable range are considered as outliers. In this paper, normal conditions are plotted as a green 'x' while outliers are plotted as a red 'x'.

4.2.1.1 Results for Hyperbox Model (Temperature Analysis)

Figure 8 represents the results for analysis on temperature. This plot functions to filter out temperature outliers. Therefore, with respect to Gross Volume and Pressure, all data points are plotted within the range of 1 to -1. With respect to temperature, most of the data points fall within the normalised range of 0-0.8. This corresponds to a temperature value of 27.5°C-29.5°C. The mode value of these data is 28.25°C. According to the results, 0.28% of the Temperature data points were classified as abnormal conditions. Only one transmitter fault data point was detected at a normalised value of -13.

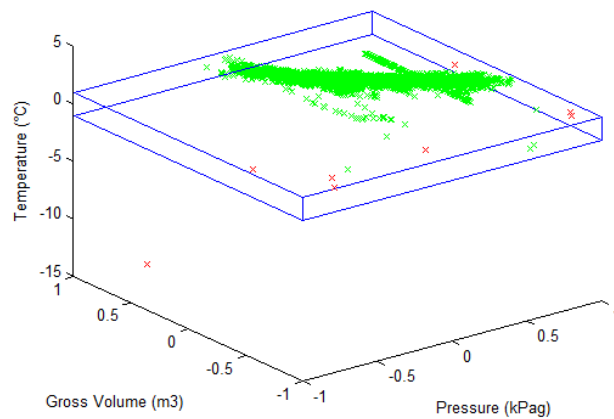


Figure 8: Hyperbox Model to Analyse Temperature

4.2.1.2 Results for Hyperbox Model (Pressure Analysis)

Figure 9 represents the results for analysis on Pressure. According to the results, 0.567% of the values were categorised as abnormal conditions. As for the mode value, a normalised value of 0.3928 was recorded. This corresponds to a value of 5437.4 kPag. Most of the data points fall within the normalised range of -0.8 to 0.75. This corresponds to values of 3350 kPag to 6062 kPag. One transmitter fault was detected from this data at a normalised value of -2.71.

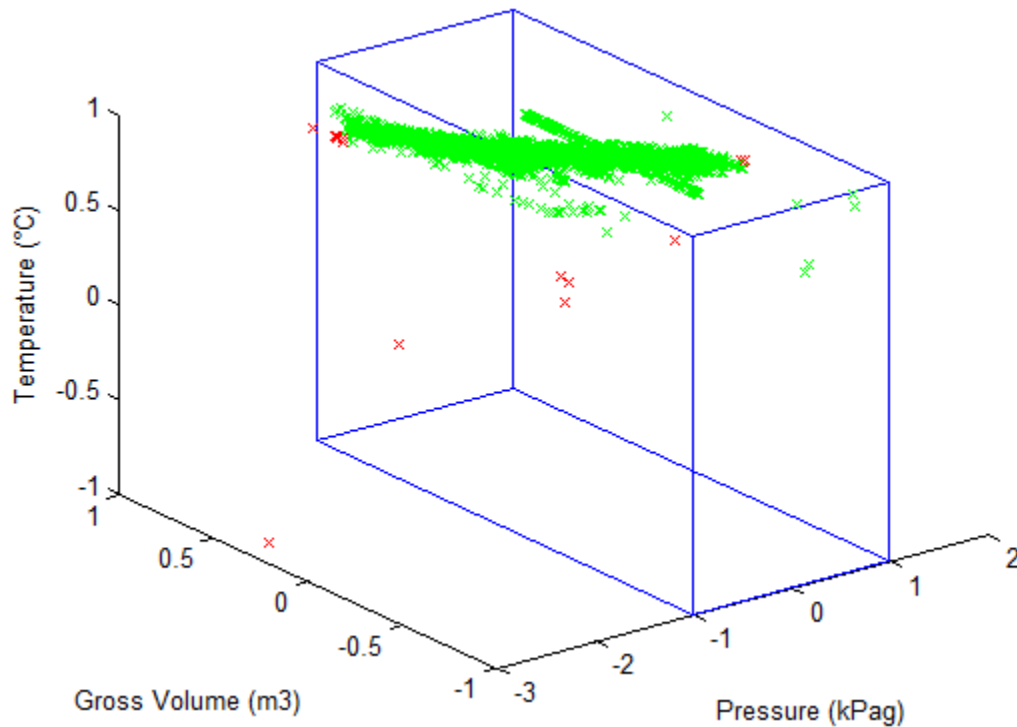


Figure 9: Hyperbox Model to Analyse Pressure

4.2.1.3 Results for Hyperbox Model (Gross Volume Analysis)

Figure 10 represents the results for analysis on Gross Volume. According to the results, most of the data falls within a normalised range of 0.5 to 1.5. This corresponds to a value of 525 m³ to 675 m³. From the data, 27.9% of the values were categorised as abnormal conditions. As for the mode value, a normalised value of -3 was recorded. This corresponds to a Gross Volume of 0 m³. The mode value of 0 m³ represents transmitter fault. This results show that the frequency for transmitter fault in terms of Gross Volume is high.

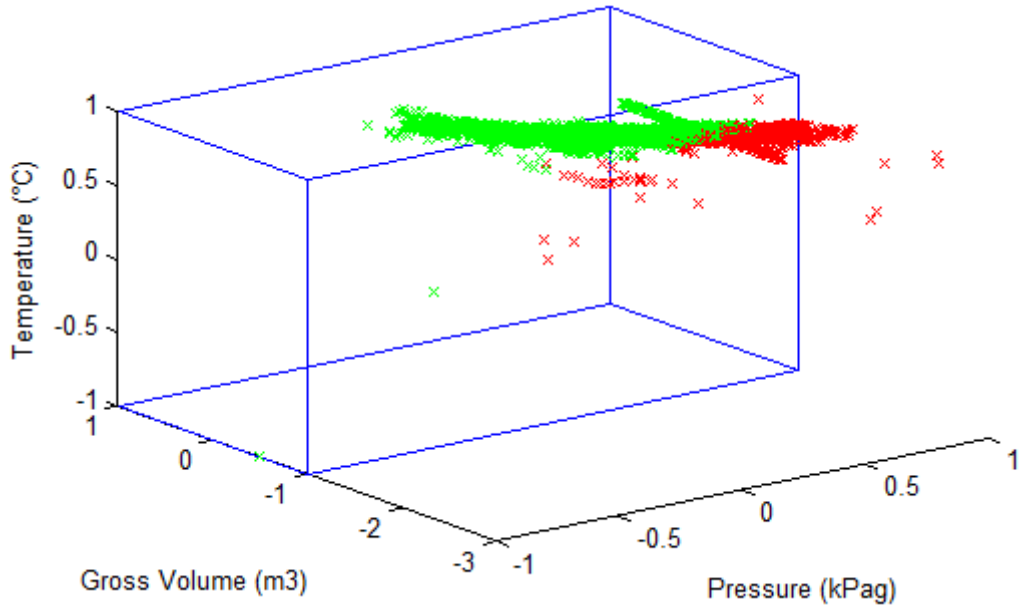


Figure 10: Hyperbox Model to Analyse Gross Volume.

4.2.1.4 Comparison of Hyperboxes

In comparison of all three hyperboxes, it can be seen that the most number of outliers lies within the Gross Volume box. This variable also registered the highest number of transmitter faults. The general shape of the Gross Volume box and Pressure box are similar. However, the Temperature box is seen to be very thin. The reason for this is because the acceptable range for temperature is small, that is 25°C-30°C. This range is small compared to the acceptable range of Pressure, 3000 kPag-6500 kPag and Gross Volume, 300m³ - 600m³. Therefore, the furthest outlier which is 0°C registers a large normalised value of -13. The furthest outlier for pressure is 0 kPag, which registered a normalised value of -2.714. Gross Volume on the other hand has its furthest outlier of 0m³ at a normalised value of -3. It can be seen that the furthest outlier for all three parameters is the transmitter fault.

4.2.2 Linear Model

The second method used to classify data into normal and intermediate conditions was the linear model. From this model, the data was classified into four different categories, which were Normal 1 (N1), Normal 2 (N2), Intermediate 1 (I1) and Intermediate 2 (I2).

The data points were plotted according to pairings shown in Table 9. Once the data points were plotted, 10% range from span of parameters was calculated. The results of this calculation are shown in Table 15.

Table 15: 10% Span for Parameter Values

Paramater	Minimum Value	Maximum Value	Range	10% of Range
Temperature (°C)	27.45	32.91	5.46	0.546
Pressure (kPag)	1582.38	6435.59	4853.21	485.321
Calorific Value (mJ/Sm ³)	21.92	39.5538	17.6338	1.76338
Gross Volume (m ³)	64.8	549.9	485.1	48.51
Standard Volume (Sm ³)	1057.3	19645.5	18588.2	1858.82

From the results in Table 15, Lower Limit Line Equation and Upper Limit Line Equation were calculated using Eq. 6 and Eq. 7. The line of best fit and equation for line of best fit were obtained automatically from simulation. These results are shown in Table 16.

Table 16: Equations for Lower Limit, Upper Limit and Line of Best Fit

Graph	Line of Best Fit Equation	Lower Limit Line Equation	Upper Limit Line Equation
1	$y = 962.13x - 2749$	$y = 962.13x - 22234.321$	$y = 962.13x - 22263.279$
2	$y = -0.0567x + 39.969$	$y = -0.0567x + 38.20562$	$y = -0.0567x + 41.73238$
3	$y = -89.61x + 2905.5$	$y = -89.61x + 2954.01$	$y = -89.61x + 2856.99$
4	$y = -483.21x + 30539$	$y = -483.21x + 28680.18$	$y = -483.21x + 32397.82$
5	$y = 0.00008x + 37.96$	$y = 0.00008x + 36.20262$	$y = 0.00008x + 39.72938$
6	$y = -0.0762x + 708.22$	$y = -0.0762x + 659.71$	$y = -0.0762x + 756.73$
7	$y = 0.2267x + 15741$	$y = 0.2267x + 13882.18$	$y = 0.2267x + 17599.82$
8	$y = -0.0002x + 38.402$	$y = -0.0002x + 36.63862$	$y = -0.0002x + 40.16538$
9	$y = 0.00008x + 36.92$	$y = 0.00008x + 35.16162$	$y = 0.00008x + 38.68838$
10	$y = 0.0158x + 89.763$	$y = 0.0158x + 41.253$	$y = 0.0158x + 138.273$

Ten graphs were plotted using parameters demonstrated in Table 9 and equations calculated in Table 16. Figure 11 shows a sample plot using pairings of Graph 1. The line in black represents line of best fit, the line in red represents the upper limit line while the line in green represents the lower limit line. Graphs for each pairing were plotted as Figure C1 and are attached within the appendix section of this paper.

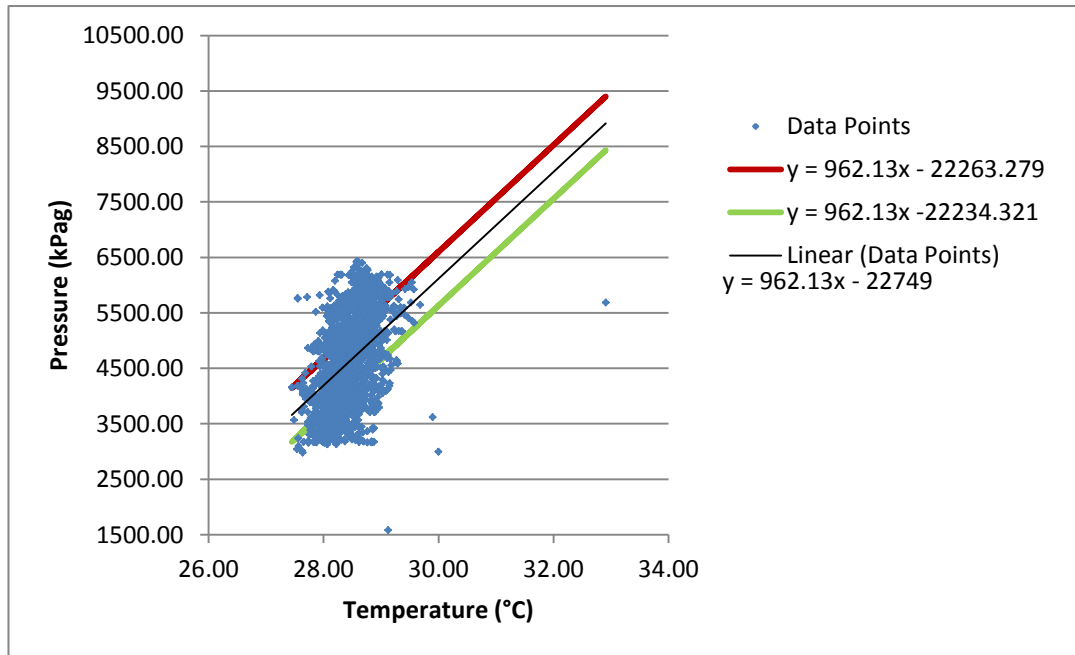


Figure 11: Linear Plot of Pressure against Temperature (Graph 1)

Using the results shown in Table 16, data was classified into four categories. The conditions for classification are shown in Table 10. The total number of healthy data points was 2501. The percentage of data points for each classification is represented in Table 17.

Table 17: Classification of Healthy Data using Linear Model

Graph	I1 (%)	N1 (%)	N2 (%)	I2 (%)
1	30.73	21.34	18.61	29.31
2	0.06	55.15	44.79	0.00
3	24.47	28.57	23.73	23.22
4	4.78	22.08	68.30	4.84
5	0.06	54.52	45.42	0.00
6	3.07	44.79	49.00	3.13
7	5.24	42.06	47.13	5.58
8	0.06	65.74	34.21	0.00
9	0.06	44.68	55.26	0.00
10	29.37	21.86	23.22	25.55

From the results displayed in Table 17, it can be seen that some plots have an even distribution across categories while other plots are not distributed as evenly. Graphs 1, 3 and 10 have an even distribution across the four classes of data. The remaining graphs have s most of the data is classified within the N1 and N2 conditions. Graphs 2, 8 and 9 registered less than 1% for data in the I1 and I2 range. The reason for this result is because these three graphs have the parameter Calorific Value plotted on the y-axis. Measurements collected for this parameter is very consistent, with small variation. A percentage variation of 10% for linear model is too large to capture the variations in Calorific Value.

4.2.3 ANFIS Classification

ANFIS model of membership function equal to four was trained for data classification. Three models were trained separately, for parameters Temperature, Pressure and Gross Volume. ANFIS has been designed such that the user may define input-output relationship rules. However, the writer did not have enough priori knowledge of the system and therefore relied on software computation to compute input-output relationship rules. The membership functions generated are displayed in Figure 12.

ANFIS membership functions were used to determine class limits for parameters as shown in Table 11 and Table 12. The class limits determined for parameters Temperature, Pressure and Gross Volume are shown in Table 18 to Table 20.

Table 18: Parameter Limits for Temperature

Parameter Class	Lower Limit (°C)	Upper Limit (°C)
Intermediate 1 (I1)	20.00	27.75
Normal 1 (N1)	27.75	29.21
Normal 2 (N2)	29.21	30.53
Normal 3 (N3)	30.53	32.91
Intermediate 2 (I2)	32.91	35.00

Table 19: Parameter Limits for Pressure

Parameter Class	Lower Limit (kPag)	Upper Limit (kPag)
Intermediate 1 (I1)	2000	2891
Normal 1 (N1)	2891	3729
Normal 2 (N2)	3729	4526
Normal 3 (N3)	4526	6345
Intermediate 2 (I2)	6345	6500

Table 20: Parameter Limits for Gross Volume

Parameter Class	Lower Limit (m ³)	Upper Limit (m ³)
Intermediate 1 (I1)	0.000	87.24
Normal 1 (N1)	87.24	164.2
Normal 2 (N2)	164.2	261.8
Normal 3 (N3)	261.8	556.0
Intermediate 2 (I2)	556.0	4500

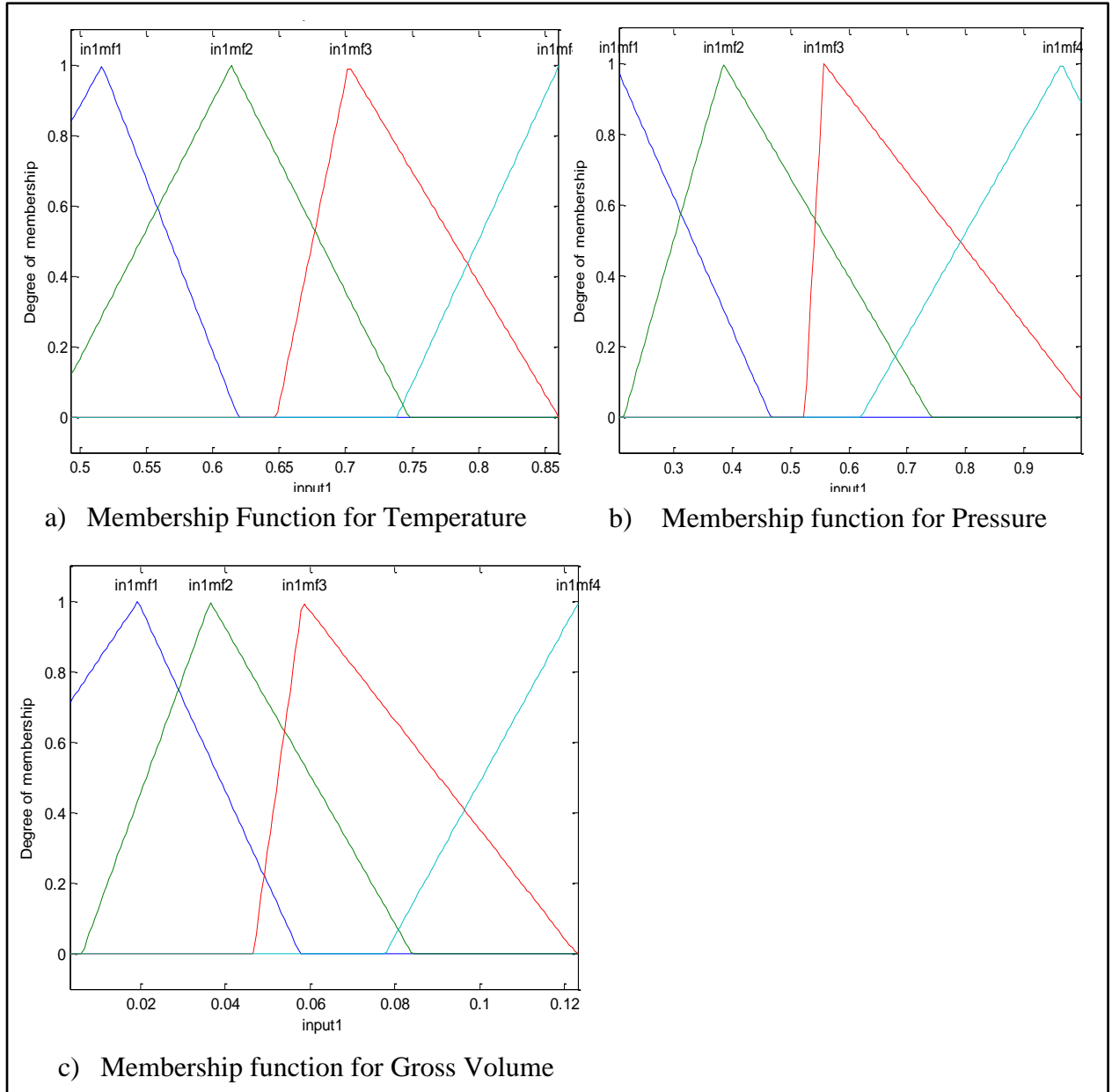


Figure 12: Membership Functions for Parameters

From a comparison of hyperbox model, linear model and ANFIS, ANFIS was chosen as the best data classification method. This is because the limits are flexible and the model is able to learn information from the training data. This is not the case for hyperbox model, which is more rigid as the limits are fixed. ANFIS is also better

than linear model because it is more easily automated in MATLAB. The reason for this is because linear method classifies data according to two parameters at a time while ANFIS classifies according to one parameter at a time. Besides that, ANFIS is able to automatically generate class limits from membership functions while the user needs to determine percentage variation for linear model.

The next step was to divide the parameters into five different classes according to the ranges provided in Table 18, Table 19 and Table 20. Figure 13 demonstrates the division of parameters into classes. The limit lines in red represent $P_{(LL)}$ and $P_{(UL)}$ while the limit lines in green represent Class Limit 1 - Class Limit 4.

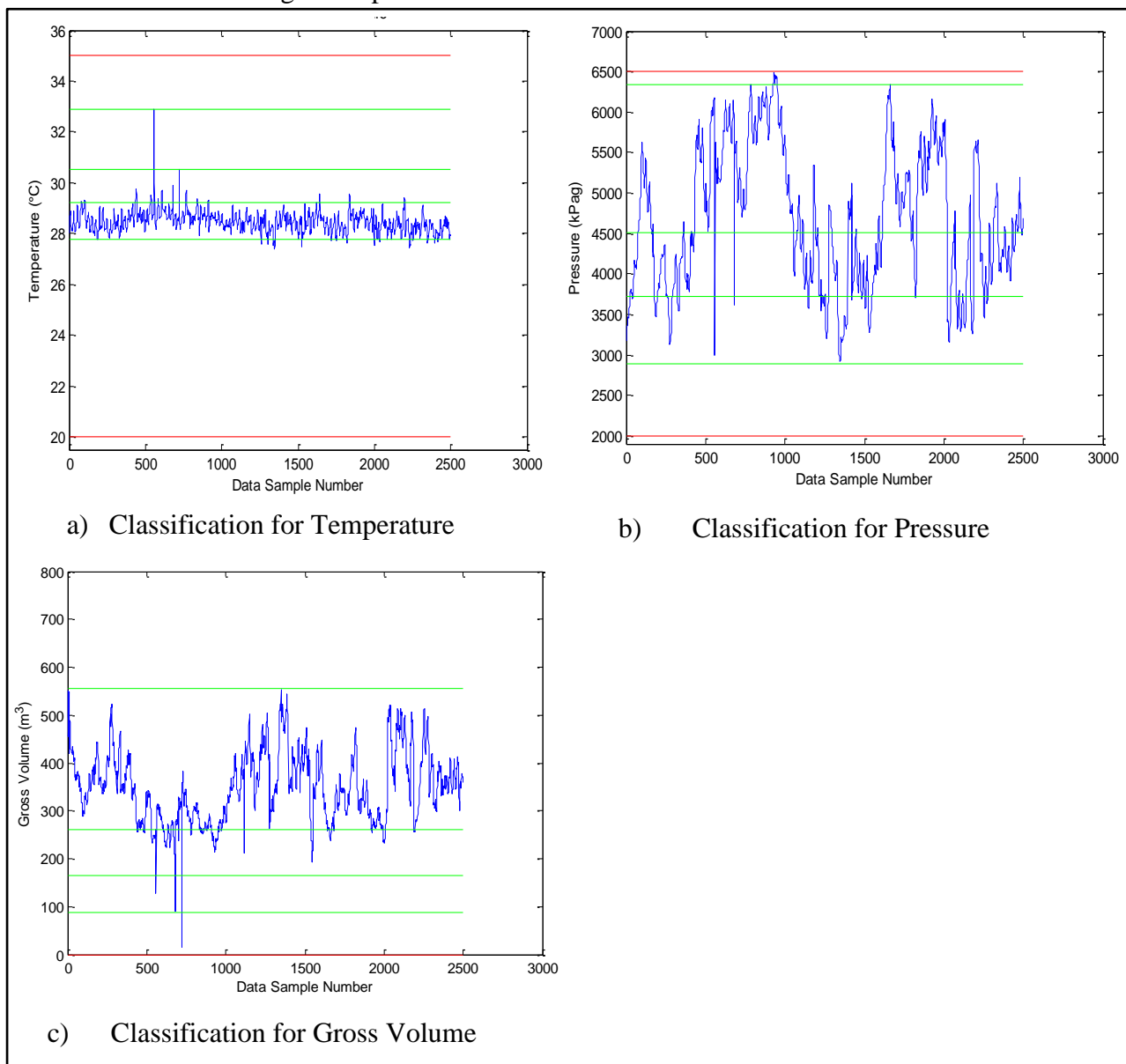


Figure 13: Classification for Parameters Temperature, Pressure and Gross Volume

The frequency of each class for parameters Pressure, Temperature and Gross Volume were counted and represented in Figure 14.

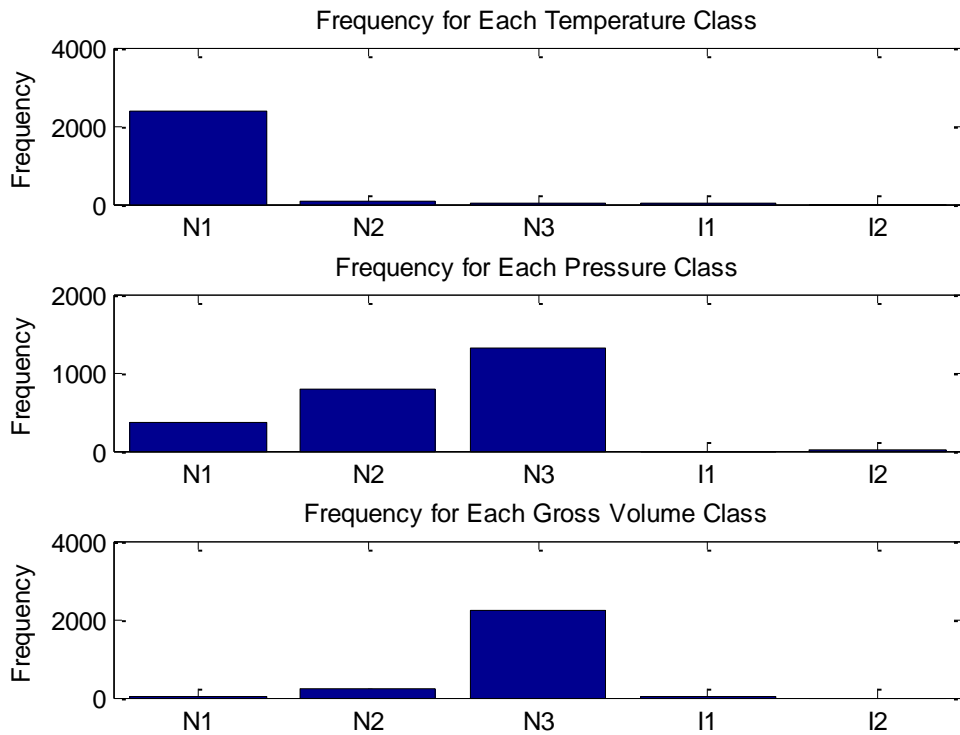


Figure 14: Class Frequency for Parameters Temperature, Pressure and Gross Volume

From the histogram of Temperature, it can be seen that most of the data points falls within the N1 class. N2, N3 and I1 have a small number of data points while I2 does not have any data points. The histogram of pressure shows that most of the pressure data lies within the normal class, N3. N2 and N1 also have a considerable number of data points. I2 has a very small number of data points while no data points were registered in the I1 class. Gross Volume consists of mode data within the N3 region. N2, I1 and N1 do not have many data points while no data points fall within the I2 region.

4.2.4 Data Clustering

In this part, data was clustered according to classes of parameters Temperature, Pressure and Gross Volume occurring at the same time. The reason for data clustering is to analyse the performance of metering system according to the class of parameters. By categorising the clusters as normal and intermediate, operators may be alerted to utilise the metering system with caution in the event that the system operates under intermediate conditions. The data was clustered according to

conditions occurring simultaneously shown in Table 21. Table 21 does not show all conditions for clustering. The complete table is shown in Appendix B

Table 21: Conditions for Data Clustering

Cluster	T	P	Vg
Cluster 1	N1	N1	N1
Cluster 2	N1	N1	N2
Cluster 3	N1	N1	N3
Cluster 4	N1	N1	I1
Cluster 5	N1	N1	I2
Cluster 6	N1	N2	N1
Cluster 7	N1	N2	N2
Cluster 8	N1	N2	N3
Cluster 9	N1	N2	I1
Cluster 10	N1	N2	I2
Cluster 11
Cluster 12
Cluster 120	I2	I1	I2
Cluster 123	I2	I2	N3
Cluster 124	I2	I2	I1
Cluster 125	X`	I2	I2

Each class was then grouped together into clusters. These clusters were formed through the condition of parameter classes occurring simultaneously. The conditions for clustering are represented in Appendix B. The frequency of each cluster was then calculated and represented in a histogram, shown by Figure 15. Clusters which registered a frequency of zero are not shown in the histogram.

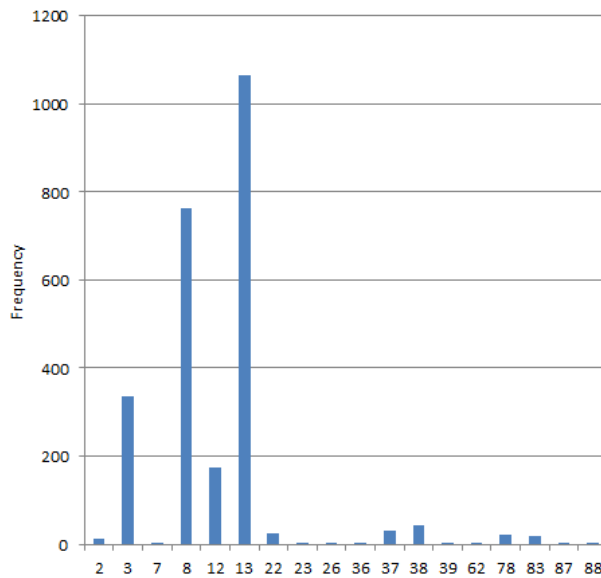


Figure 15: Histogram Representing Frequency of Each Cluster

The reason why many categories were registered as zero frequency was because for the classification of temperature parameter, most of the data points fall within the N1 region. N1 region for Temperature corresponds to Cluster 1 - Cluster 25. Therefore, the highest frequency of data points will be recorded among these clusters, as demonstrated by the histogram in Figure 15. The cluster which recorded the highest number of data points is cluster 13. The reason for this trend is because this cluster has conditions for mode of all three parameters which is N1 from Temperature, N3 from Pressure and N3 from Gross Volume.

4.3 Part III: Field Measurement Prediction

This section of the report is dedicated towards discussing ANFIS model development for field measurement prediction. In order to identify the best model, variables of ANFIS input, percentage of data division for training and checking, number of epoch for model training, type of membership function and randomisation of data were varied. The models were developed for parameters Temperature, Pressure and Gross Volume. Two models were developed for each parameter, which are for transmitter and hang faults. A total of six models were developed. The details of each model are shown in Table 22.

Table 22: ANFIS Model Number and Predicted Field Measurement

ANFIS Model Number	Field Measurement Prediction
ANFIS Model 1	Temperature prediction for transmitter fault
ANFIS Model 2	Pressure prediction for transmitter fault
ANFIS Model 3	Gross volume prediction for transmitter fault
ANFIS Model 4	Temperature prediction for hang fault
ANFIS Model 5	Pressure prediction for hang fault
ANFIS Model 6	Gross volume prediction for hang fault

Once data has been normalised, the five manipulated variables were tested in order to determine the best model for each parameter. The five manipulated variables are ANFIS input, data division, number of epoch for training, type of membership function, and randomisation of data. The testing for five different manipulated variables are presented in this section of report.

4.3.1 Manipulated Variable 1: ANFIS input

The most important variable when training an ANFIS model is the training inputs to the model. The inputs to the ANFIS model were varied in order to find the best model. The best model is identified as the model that exhibits lower parameter and energy MAPE. MAPE is calculated according to Eq.10. The energy MAPE on the other hand is calculated by using the value of predicted parameter in Eq.1 and Eq.2. MAPE was then once again calculated for energy using Eq.10.

Table 23 to Table 25 shows the results for manipulation of ANFIS input for transmitter faults. In order to predict current field parameter value during hang fault condition, past values of other parameters are analysed as ANFIS input. It was chosen to take values of the past seven hours because from the analysed data, the longest period for transmitter fault was seven hours.

The chosen combination of inputs is highlighted in green and demonstrated within Figure 16. From the combination of inputs for ANFIS model 1, although five inputs gave the lowest error, four inputs were chosen. This is because four inputs are able to achieve a low MAPE of 0.1842%. This already satisfies the criteria set by PGB, which was to achieve MAPE for energy of less than 1%. As for ANFIS model 2, three inputs were chosen as the lowest MAPE was achieved, of 7.0262%. These parameters were Temperature, Gross Volume, and Pressure. Calorific Value, Standard Gravity and Energy were not included because a higher MAPE was recorded. The same process was repeated for ANFIS model 3. Three inputs of Temperature, Pressure, and Gross Volume were chosen as the best model, with an MAPE of 8.7913%.

A similar process was repeated in developing model for parameter prediction during hang fault. Table 26 to Table 28 shows the results for manipulation of ANFIS input for hang faults. The difference is that instead of analysing past parameter values, current values of other parameters are analysed as ANFIS input. The chosen combination of inputs is highlighted in green and demonstrated in Table 23 to Table 28. In order to predict Temperature, four inputs were chosen as the most suitable model of MAPE 0.2114%. As for ANFIS model 5, two inputs were chosen as the lowest MAPE was achieved, of 4.3030%. Two inputs of Temperature and Pressure were also chosen for ANFIS model 6, resulting in MAPE of 6.0701. It was observed

that Calorific Value and Standard Gravity as inputs resulted in higher MAPE for both ANFIS model 5 and 6.

Table 23: Input Combination for ANFIS Model 1

Input	Parameter MAPE	Energy MAPE
P(n-7)	3.7842%	2.1754%
Vg(n-7)		
P(n-7)	2.6432%	0.1842%
Vg(n-7)		
CV(n-7)		
SG(n-7)		
P(n-7)	2.1826%	0.1754%
Vg(n-7)		
CV(n-7)		
SG(n-7)		
Hour(n-7)		

Table 24: Input Combination for ANFIS Model 2

Input	Parameter MAPE	Energy MAPE
T(n-7)	7.7587	7.7587
Vg(n-7)		
T(n-7)	7.0262	7.0262
Vg(n-7)		
P(n-7)		
T(n-7)	8.473	8.473
Vg(n-7)		
CV(n-7)		
T(n-7)	10.0356	10.0356
Vg(n-7)		
SG(n-7)		
T(n-7)	7.2635	7.2635
Vg(n-7)		
E(n-7)	7.6764	7.6764
T(n-7)		
Vg(n-7)		
P(n-7)		
E(n-7)		

Table 25: Input Combination for ANFIS Model 3

Input	Parameter MAPE	Energy MAPE
T(n-7)	9.5255	9.5255
P(n-7)		
T(n-7)	8.7913	8.7913
P(n-7)		
Vg(n-7)		
T(n-7)	9.0232	9.0232
P(n-7)		
Vg(n-7)		
CV(n-7)		
T(n-7)	9.0209	9.0209
P(n-7)		
Vg(n-7)		
SG(n-7)	9.0239	9.0239
T(n-7)		
P(n-7)	9.021	9.021
Vg(n-7)		
SG(n-7)		
E(n-7)		
E(n-7)		

Table 26: Input Combination for ANFIS Model 4

Input	Parameter MAPE	Energy MAPE
P(n)	4.2381	1.3254
Vg(n)		
P(n)	2.2419	0.2114
Vg(n)		
CV(n)		
SG(n)		
P(n)	1.7263	0.2032
Vg(n)		
CV(n)		
SG(n)		
Hour(n)		

Table 27: Input Combination for ANFIS Model 5

Input	Parameter MAPE	Energy MAPE
T(n)	4.3030	4.3029
Vg(n)		
T(n)	5.1104	5.1104
Vg(n)		
CV(n)		
T(n)	6.0508	6.0508
Vg(n)		
SG(n)		
T(n)	8.0748	8.0748
Vg(n)		
CV(n)		
SG(n)		

Table 28: Input Combination for ANFIS Model 6

Input	Parameter MAPE	Energy MAPE
T(n)	6.0701	6.0701
P(n)		
T(n)	6.3636	6.3636
P(n)		
CV(n)		
T(n)	9.1100	9.1100
P(n)		
SG(n)		
T(n)	9.6258	9.6258
P(n)		
CV(n)		
SG(n)		

Figure 16 represents the chosen input combinations for each model.

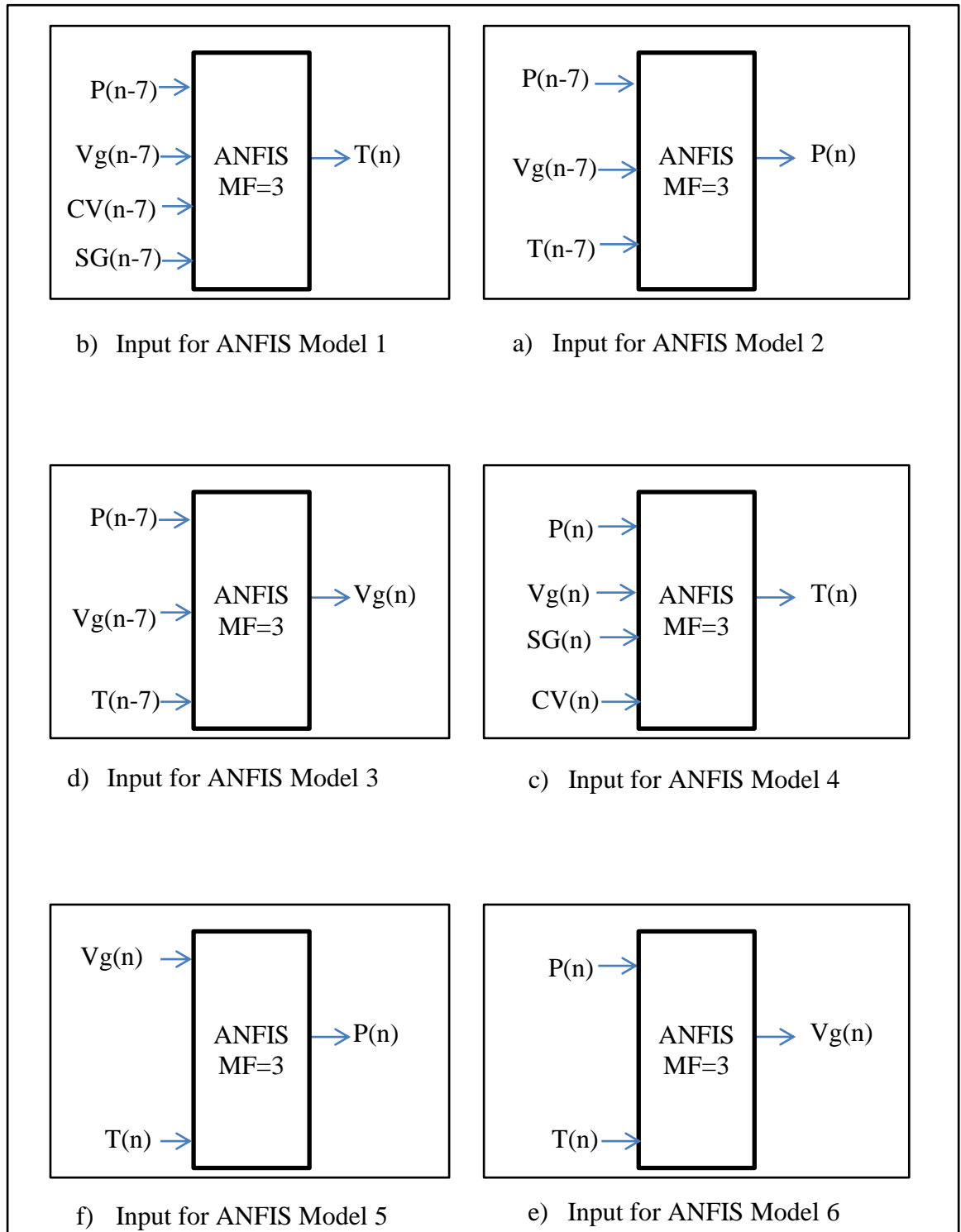


Figure 16: Chosen Input Combinations for ANFIS Model 1 - ANFIS Model 6

4.3.2 Manipulated Variable 2: Data Division

In this section of the project, the division of data into training and checking was manipulated. Three different sets of data were applied to each ANFIS model. The three sets are 50-50 (50% training and 50% checking), 25-75 (25% training and 75% checking) and 75-25 (75% training and 25 % checking). The results for this test are shown in Table 29.

Table 29: Selection of Data Division for ANFIS Model 1- ANFIS Model 6

a) ANFIS Model 1			b) ANFIS Model 4		
Data Division	Parameter MAPE	Energy MAPE	Data Division	Parameter MAPE	Energy MAPE
50-50	0.8022	0.7865	50-50	0.7559	0.6254
25-75	1.9876	1.2453	25-75	1.9693	1.7865
75-25	0.7389	0.6754	75-25	0.6075	0.5642

c) ANFIS Model 2			d) ANFIS Model 5		
Data Division	Parameter MAPE	Energy MAPE	Data Division	Parameter MAPE	Energy MAPE
50-50	7.6453	7.6453	50-50	7.2852	7.2852
25-75	8.4054	8.4054	25-75	27.96	27.96
75-25	7.0262	7.0262	75-25	4.303	4.303

e) ANFIS Model 3			f) ANFIS Model 6		
Data Division	Parameter MAPE	Energy MAPE	Data Division	Parameter MAPE	Energy MAPE
50-50	9.1727	9.1727	50-50	7.3818	7.3818
25-75	9.5345	9.5345	25-75	8.6124	8.6124
75-25	9.0209	9.0209	75-25	6.0701	6.0701

The same trend for all six models was observed, that is data division of 75% for training and 25% for checking resulted in the smallest MAPE. This data division set was chosen for all six models.

4.3.3 Manipulated Variable 3: Number of Epoch

In this section of the report, the number of epoch is varied for each ANFIS Model. The variation is set as increasing steps of 20 iterations, from 20 epoch until 100 epoch. The epoch with the lowest MAPE was selected as the best epoch for the ANFIS model. The results for varying number of epoch is shown in Table 30.

Table 30: Selection of Number of Epoch for ANFIS Model 1-6

a) ANFIS Model 1			b) ANFIS Model 4		
Epoch No	Parameter MAPE	Energy MAPE	Epoch No	Parameter MAPE	Energy MAPE
20	0.7711	0.7613	20	0.6124	0.432
40	0.762	0.6634	40	0.6106	0.4298
60	0.7484	0.6621	60	0.6088	0.4288
80	0.7444	0.662	80	0.6075	0.4278
100	0.7389	0.6489	100	0.6061	0.4263

c) ANFIS Model 2			d) ANFIS Model 5		
Epoch No	Parameter MAPE	Energy MAPE	Epoch No	Parameter MAPE	Energy MAPE
20	7.1852	7.1852	20	4.2854	4.2854
40	7.1666	7.1666	40	4.303	4.3029
60	7.0262	7.0262	60	4.303	4.3029
80	7.0283	7.0283	80	4.303	4.3029
100	7.0101	7.0101	100	4.303	4.3029

e) ANFIS Model 3			f) ANFIS Model 6		
Epoch No	Parameter MAPE	Energy MAPE	Epoch No	Parameter MAPE	Energy MAPE
20	8.7935	8.7935	20	6.0126	6.0126
40	8.7007	8.7007	40	6.0559	6.0559
60	8.6097	8.6097	60	6.0701	6.0701
80	8.5931	8.5931	80	6.065	6.065
100	8.5852	8.5852	100	6.0588	6.0588

From the results obtained, the writer was unable to determine a pattern to find the best number of epoch. This could be due to the phenomena of overfitting. Overfitting occurs when the ANFIS model is over trained, such that when new data is presented to the model, inaccurate results are obtained. ANFIS model 1, 2, 3 and 4 show similar trends where the accuracy of ANFIS model increases with increasing epoch number. Thus, epoch number of 100 was chosen for each of these models. The trend

for ANFIS model 5 and 6 was such that lower epoch number gave lower MAPE. Therefore, epoch number of 20 was chosen for both these models.

4.3.4 Manipulated Variable 4: ANFIS membership function type

In this section of the report, the influence of membership function (MF) type on ANFIS model accuracy was tested. Six different MF types were tested which are trimf, gaussmf, trapmf, gbellmf, pimf, and disgmf. The results for this test are presented in Table 31.

Table 31: Selection of MF Type for ANFIS Model 1-6

b) ANFIS Model 1			a) ANFIS Model 4		
MF Type	Parameter MAPE	Energy MAPE	MF Type	Parameter MAPE	Energy MAPE
trimf	0.7486	0.7567	trimf	0.6061	0.5676
gaussmf	0.7389	0.6479	gaussmf	0.5902	0.4986
trapmf	0.8336	0.789	trapmf	0.7424	0.5873
gbellmf	0.7163	0.6324	gbellmf	0.5567	0.4367
pimf	0.8293	0.7678	pimf	0.731	0.5787
disgmf	0.7373	0.6487	disgmf	0.5763	0.4876

c) ANFIS Model 2			d) ANFIS Model 5		
MF Type	Parameter MAPE	Energy MAPE	MF Type	Parameter MAPE	Energy MAPE
trimf	6.8221	6.8221	trimf	4.2569	4.2569
gaussmf	7.0101	7.0101	gaussmf	4.2854	4.2854
trapmf	7.2483	7.2483	trapmf	4.5319	4.5319
gbellmf	7.0663	7.0663	gbellmf	4.2969	4.2969
pimf	7.3847	7.3847	pimf	4.2736	4.2736
disgmf	7.2037	7.2037	disgmf	4.7341	4.7341

e) ANFIS Model 3			f) ANFIS Model 6		
MF Type	Parameter MAPE	Energy MAPE	MF Type	Parameter MAPE	Energy MAPE
trimf	8.789	8.789	trimf	5.6836	5.6836
gaussmf	8.8423	8.8423	gaussmf	6.0126	6.0126
trapmf	8.6529	8.6529	trapmf	5.915	5.915
gbellmf	8.5852	8.5852	gbellmf	6.0346	6.0346
pimf	8.6535	8.6535	pimf	5.8597	5.8597
disgmf	8.7216	8.7216	disgmf	6.0552	6.0552

From the results obtained, there was no significant trend identified when the type of MF was varied. MF type of gbellmf was chosen for ANFIS model 1, 3 and 4. MF type of trimf was chosen for ANFIS model 2, 5 and 6.

4.3.5 Manipulated Variable 5: Randomization of Data

In this section of the report, the effect of randomisation of data was tested. Each ANFIS model was tested with input data in sequence and input data that has been randomised according to rows. The MAPE was then compared. The results of this test are presented in Table 32.

Table 32: MAPE for Randomised and Sequential Data ANFIS Model 1-6

ANFIS Model	Randomised Data MAPE (%)	Sequential Data MAPE (%)
ANFIS Model 1	0.6126	2.7542
ANFIS Model 2	5.6586	6.8221
ANFIS Model 3	10.6729	8.5852
ANFIS Model 4	0.5567	1.6432
ANFIS Model 5	4.9776	4.2569
ANFIS Model 6	5.5025	5.6836

Out of the six models, four models which are ANFIS model 1, 2, 4 and 6 produced lower MAPE when data was randomised. ANFIS model 3 and 4 produced better results when data was input sequentially. Randomisation therefore applied for ANFIS models which produced better results with randomised input.

4.3.6 Models for Parameter Prediction During Transmitter and Hang Fault

Once testing for the five manipulated variables was conducted, an ANFIS model was developed comprising the best of each variable. The parameters of the six developed models are shown in Table 33. From the results shown in the Table, ANFIS Model 1 and ANFIS Model 4 exhibit low MAPE of less than 1% which is required by PGB. Model for Pressure and Gross Volume which are ANFIS Models 2, 3, 5 and 6 have a MAPE of 4.3% to 8.6%. These results are unable meet the industrial requirements by PGB.

Table 33: Developed Models for Field Measurement Prediction

ANFIS MODEL	Model Input	Data Division	Epoch Number	MF Type	Data Randomisation	Parameter MAPE	Energy MAPE
ANFIS Model 1	P(n-7), Vg(n-7), CV(n-7), SG(n-7), T(n-7)	75-25	100	Gbellmf	Yes	0.6126%	0.2126%
ANFIS Model 2	P(n-7), Vg(n-7), T(n-7)	75-25	100	Trimf	Yes	5.6586%	5.6586%
ANFIS Model 3	P(n-7), Vg(n-7), T(n-7)	75-25	100	Gbellmf	No	8.5852%	8.5852%
ANFIS Model 4	P(n), Vg(n), CV(n), SG(n)	75-25	100	Gbellmf	Yes	0.5567%	0.3567%
ANFIS Model 5	T(n), Vg(n),	75-25	20	Trimf	No	4.2569%	4.2569%
ANFIS Model 6	T(n), P(n),	75-25	20	trimf	Yes	5.5025 %	5.5025%

The six models detailed within Table 33 show the parameters of prediction models to calculate and predict energy. The energy was predicted using Eq. 1 and Eq. 2. The plot of predicted energy against actual energy for the six different models is shown in Figure 17. As can be seen in Figure 17, ANFIS Model 1 and ANFIS Model 4 produced very accurate results, with actual energy value almost overlapping with predicted energy value. ANFIS Model 2, 3, 5 and 6 on the other hand did not produce such accurate results. The predicted energy is unable to accurately follow the pattern of actual energy, and some spikes in the plot were noted. In order for these four models to be used practically, it must be further improved in order to reduce the MAPE to a value of less than 1%. The plot of actual energy against predicted energy for ANFIS model 1-6 is demonstrated in Figure 17.

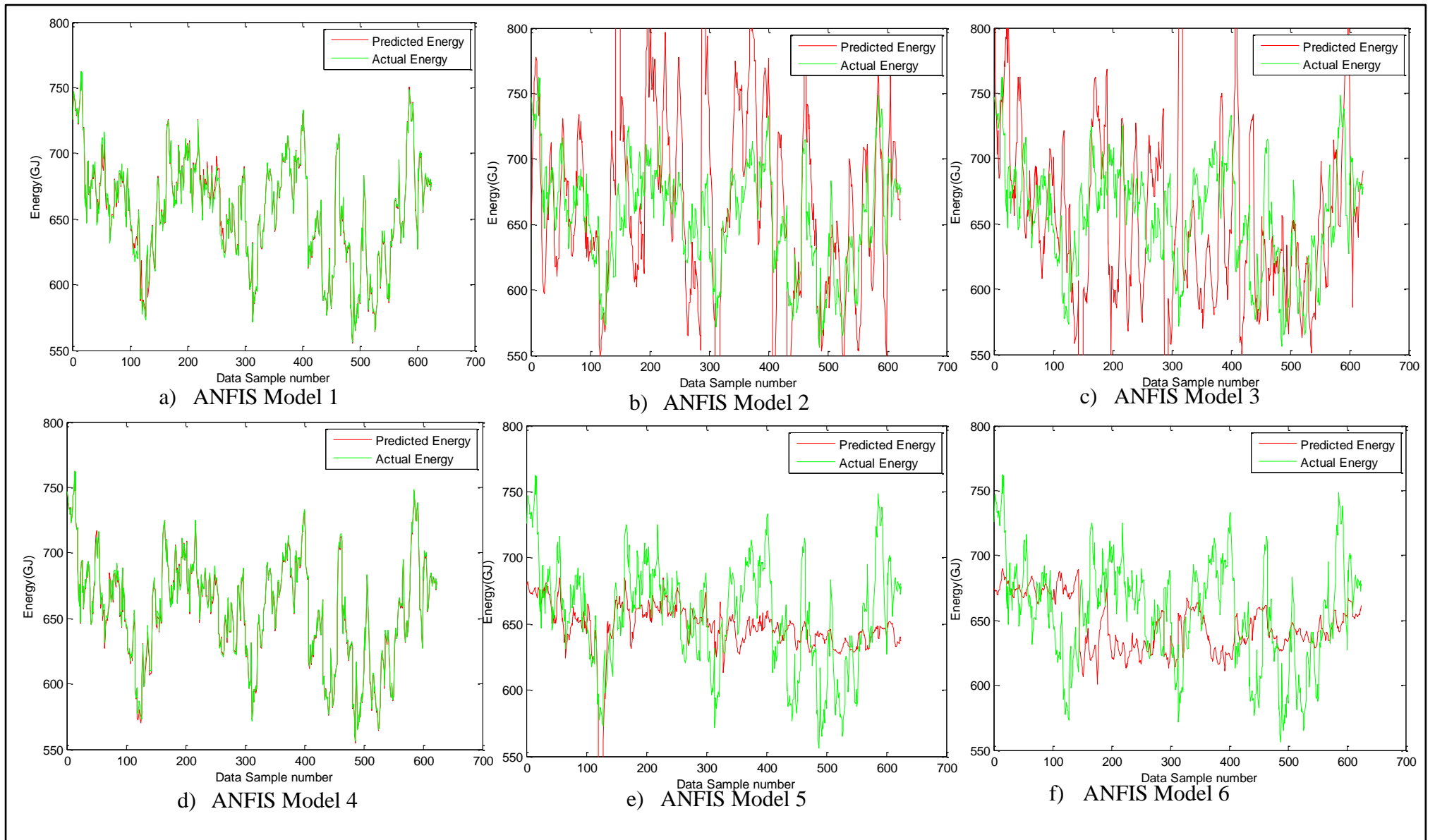


Figure 17: Actual Energy against Predicted Energy for ANFIS Model 1- ANFIS Model 6

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

To conclude this report, the objectives of the project have been achieved. The writer has studied potential fault diagnosis methods which can be applied at the gas metering system to Kapar power plant. The writer then compared and evaluated the different types of diagnosis methods. From the evaluation, the writer has chosen to use the ANFIS model. A fault diagnosis for trial basis was then proposed. Lastly, the writer proposed field measurement prediction technique using ANFIS which could be used as a backup measurement in case of fault condition.

In order to develop fault diagnosis method, the writer first filtered out fault data, including hang fault and transmitter fault. Healthy data was then classified using three methods which are hyperbox model, linear model and ANFIS. The three methods were compared and ANFIS was chosen as the best. The reason for choosing ANFIS is that it is more flexible than hyperbox model in the sense it is able to learn from data to set class limits. It is also simpler to implement in MATLAB compared to linear model. Besides that, the class limits are automatically generated through MATLAB. This eliminates the problem of having to define percentage variation as needed to be done in linear model. ANFIS model was used to classify the data into five different classes (N1, N2, N3, I1, I2). Next, the data was clustered into 125 clusters according to conditions occurring simultaneously. The writer recommends that field measurement prediction techniques be applied for any cluster within the intermediate range.

In order to predict field measurements, six ANFIS models were developed to cater for Temperature, Pressure and Gross Volume parameters during hang and transmitter faults. Five variables which are ANFIS input, data division, number of epoch, type of membership function and randomization of data were tested in order to determine the best models. Models 1 and 4 have recorded an MAPE of 0.2126%

and 0.3567% respectively. This satisfies the industrial requirements set by PGB, which is MAPE of less than 1%. Models 2, 3, 5 and 6 need to be further developed to meet industrial requirements, as the current MAPE is above 1%.

5.2 Recommendation

This project can be further modified and improved in many aspects. In this section of the report, recommendations for future improvement of the project are detailed. The different sections in which improvements can be made include fault isolation, data classification, fault diagnosis and field measurement prediction. In terms of fault isolation, the developed filtering method does not cater for drift faults. For future modification works, the filtering method could be further modified to cater for drift faults.

In terms of data classification, three methods have been tested which are the hyperbox model, linear model and ANFIS. After comparison of these three methods, ANFIS was chosen as the best method of data classification for fault diagnosis. The concern of using ANFIS for data classification is that once the class limits have been set, the limits are inflexible and unable to respond to changes in data. This issue may be addressed by introducing principle component analysis (PCA) in order to generate a vigilance parameter. The vigilance parameter will make the class limits more flexible and respond to new changes in data.

Fault diagnosis for this project is currently limited to singular fault. Development to include diagnosis of multiple faults simultaneously will help this project to advance by milestones. The next recommendation for fault diagnosis is to use a different diagnosis method, such as adaptive resonance theory (ART). The advantage of ART over ANFIS is that the generated class limits are flexible as ART it is able to learn new information without overwriting existing information (high stability plasticity). This will allow the model to respond to changes in new data while maintaining information learned from training data to set suitable class limits. Another method that may be implemented for fault diagnosis is time-based analysis. Using this method, deviation of data with respect to time may be analysed in order to determine if a fault may occur.

The next recommendation is proposed for field measurement prediction. Six different models for field measurement prediction have been proposed, as detailed in Table 22. Currently, Models 1 and 4 satisfy industrial requirements. However, Models 2, 3, 5 and 6 need to be developed further to reduce MAPE. This can be done by dividing the models further according time. For example, models could be separated according to weekdays and weekends or daytime and nighttime. Another method is to create separate models according to class of data, such as N1, N2, N3, I1 and I2. These methods may increase the accuracy of the prediction model. Besides that, different artificial intelligent methods such as neural networks and fuzzy logics may be explored.

REFERENCES

- [1] Korbicz, J., Koscielny, J. M., Kowalczyk, Z., & Cholewa, W. (2004). *Fault Diagnosis Models, Artificial Intelligence, Applications*. New York: Springer.
- [2] Isermann, R. (2005). *Model-based Fault-detection and Diagnosis – Status and Applications*. *Annual Reviews in Control*, 29, 71-85
- [3] Palade, V., Bocaniala, C. D., & Jain, L. C. (2006). *Computational Intelligence in Fault Diagnosis*. London: Springer.
- [4] Mohsin, R., & Nasri, N. S. (1995). *Calibration concept and equipment drifting factors: its application in gas metering facilities*. Proceedings of The Eleventh Symposium of Malaysia Chemical Engineers. C15-1.
- [5] Nguyen, T. H. (2013). *Intelligent Fault Diagnostic System for Instruments in Gas Transportation System*, Degree Project Proposal, Universiti Teknologi Petronas.
- [6] Merriam-Webster (2014). Retrieved 15, February, 2014 from <http://www.merriam-webster.com/>
- [7] Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2003). *A Review of Process Fault Detection and Diagnosis Part 1: Quantitative Model-based Methods*. *Computers and Chemical Engineering*, 27, 293-311.
- [8] Chen, J., & Patton R. J. (1999). *Robust Model-Based Fault Diagnosis For Dynamic Systems*. Netherlands: Kluwer Academic Publishers.
- [9] Isermann, R. & Balle, P. (1997). *Trends in the Application of Model-Based Fault Detection and Diagnosis of Technical Processes*. *Control Engineering Practice*, 5(5), 709-719.

- [10] Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2003). *A Review of Process Fault Detection and Diagnosis Part II: Qualitative Models and Search Strategies*. *Computers and Chemical Engineering*, 27, 313-326.
- [11] Kolluri, S., Bajaj, I., & Bushan, M. (2013). *Sensor Network Design for Efficient Fault Diagnosis and Signed Digraph Update*. *IFAC DYCOPS*, 821-826.
- [12] Chang, C., & Chen, C. Y. (2011). *Fault Diagnosis with Automata Generated Languages*. *Computers & Chemical Engineering*, 35(2), 329-341.
- [13] Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2003). *A Review of Process Fault Detection and Diagnosis Part III: Process History Based Methods*. *Computers and Chemical Engineering*, 27, 327-346.
- [14] Principe, J. C., Euliano, N., R., & Lefebvre, W. C. (1999). *Neural and Adaptive Systems*. New York: John Wiley.
- [15] Shing, J., & Roherm J. (1993). *ANFIS: Adaptive-Network-Based Fuzzy Inference System*. *Systems, Man and Cybernetics*, 23, 665-667.
- [16] Mehmet, T. (2010). *Weather Forecasting using ANFIS and ARIMA Models*. *Environmental Research, Engineering and Management*, 1(51), 5-10.
- [17] Mellit, A., Hadj Arab, A., Khorissi, N., & Salhi, H. (2007). *An ANFIS-based Forecasting for Solar Radiation Data from Sunshine Duration and Ambient Temperature*. *Power Engineering Society General Meeting, IEEE*, 1-6.
- [18] Maryam Jamela, I. (2011). *Neural Network prediction Model of Energy Consumption & Billing Verification*. Master of Science Thesis, Universiti Teknologi Petronas

Appendix A- Gantt Chart and Key Milestone

Table A1: Gantt Chart and Key Milestone (FYP 1)

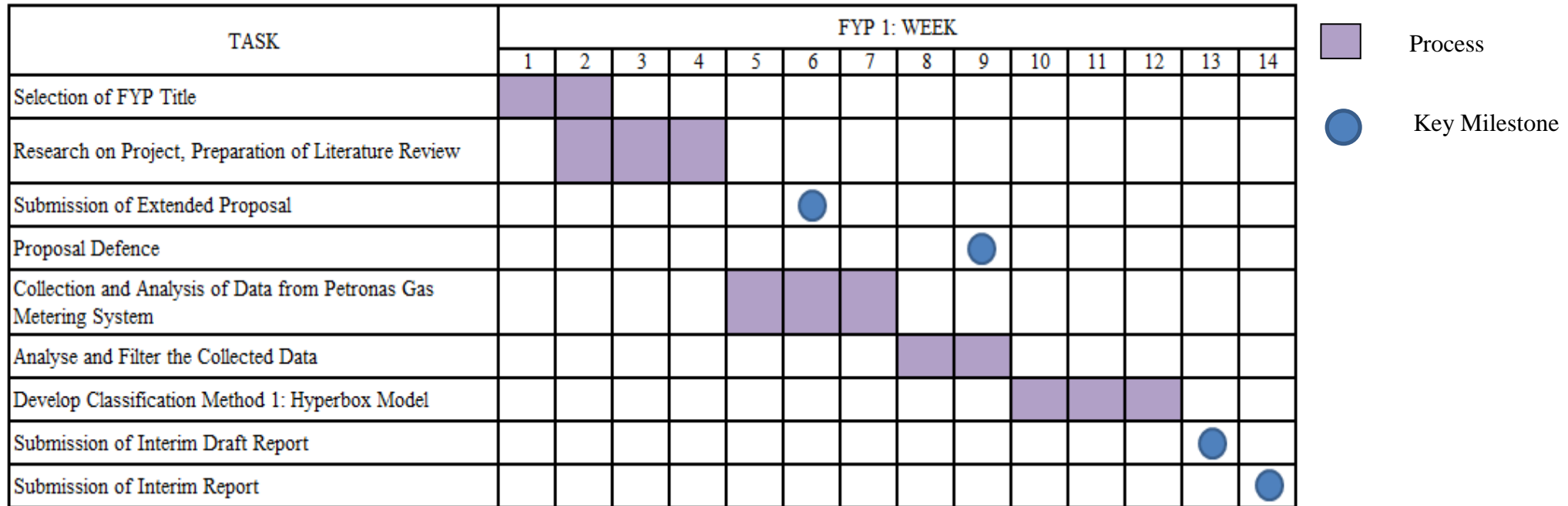


Table A2: Gantt Chart and Key Milestone (FYP 2)

TASK	FYP 2: WEEK													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Develop Classification Method 2: Linear Model	■	■	■											
Submission of Progress Report							●							
Develop Classification Method 3: ANFIS				■	■	■								
Selection of Best Classification Method							■							
Grouping of Data into Clusters								■	■					
Pre-SEDEX										●				
Develop ANFIS Field Measurement Prediction Model										■	■	■		
Submission of Draft Final Report											●			
Submission of Dissertation (Soft Bound)												●		
Submission of Technical Paper												●		
Viva													●	
Submission of Project Dissertation (Hard Bound)														●



Appendix B- Conditions for Data Clustering

Cluster	T	P	Vg
Cluster 1	N1	N1	N1
Cluster 2	N1	N1	N2
Cluster 3	N1	N1	N3
Cluster 4	N1	N1	I1
Cluster 5	N1	N1	I2
Cluster 6	N1	N2	N1
Cluster 7	N1	N2	N2
Cluster 8	N1	N2	N3
Cluster 9	N1	N2	I1
Cluster 10	N1	N2	I2
Cluster 11	N1	N3	N1
Cluster 12	N1	N3	N2
Cluster 13	N1	N3	N3
Cluster 14	N1	N3	I1
Cluster 15	N1	N3	I2
Cluster 16	N1	I1	N1
Cluster 17	N1	I1	N2
Cluster 18	N1	I1	N3
Cluster 19	N1	I1	I1
Cluster 20	N1	I1	I2
Cluster 21	N1	I2	N1
Cluster 22	N1	I2	N2
Cluster 23	N1	I2	N3
Cluster 24	N1	I2	I1

Cluster	T	P	Vg
Cluster 25	N1	I2	I2
Cluster 26	N2	N1	N1
Cluster 27	N2	N1	N2
Cluster 28	N2	N1	N3
Cluster 29	N2	N1	I1
Cluster 30	N2	N1	I2
Cluster 31	N2	N2	N1
Cluster 32	N2	N2	N2
Cluster 33	N2	N2	N3
Cluster 34	N2	N2	I1
Cluster 35	N2	N2	I2
Cluster 36	N2	N3	N1
Cluster 37	N2	N3	N2
Cluster 38	N2	N3	N3
Cluster 39	N2	N3	I1
Cluster 40	N2	N3	I2
Cluster 41	N2	I1	N1
Cluster 42	N2	I1	N2
Cluster 43	N2	I1	N3
Cluster 44	N2	I1	I1
Cluster 45	N2	I1	I2
Cluster 46	N2	I2	N1
Cluster 47	N2	I2	N2
Cluster 48	N2	I2	N3

Cluster	T	P	Vg
Cluster 49	N2	I2	I1
Cluster 50	N2	I2	I2
Cluster 51	N3	N1	N1
Cluster 52	N3	N1	N2
Cluster 53	N3	N1	N3
Cluster 54	N3	N1	I1
Cluster 55	N3	N1	I2
Cluster 56	N3	N2	N1
Cluster 57	N3	N2	N2
Cluster 58	N3	N2	N3
Cluster 59	N3	N2	I1
Cluster 60	N3	N2	I2
Cluster 61	N3	N3	N1
Cluster 62	N3	N3	N2
Cluster 63	N3	N3	N3
Cluster 64	N3	N3	I1
Cluster 65	N3	N3	I2
Cluster 66	N3	I1	N1
Cluster 67	N3	I1	N2
Cluster 68	N3	I1	N3
Cluster 69	N3	I1	I1
Cluster 70	N3	I1	I2
Cluster 71	N3	I2	N1
Cluster 72	N3	I2	N2

Cluster	T	P	Vg
Cluster 73	N3	I2	N3
Cluster 74	N3	I2	I1
Cluster 75	N3	I2	I2
Cluster 76	I1	N1	N1
Cluster 77	I1	N1	N2
Cluster 78	I1	N1	N3
Cluster 79	I1	N1	I1
Cluster 80	I1	N1	I2
Cluster 81	I1	N2	N1
Cluster 82	I1	N2	N2
Cluster 83	I1	N2	N3
Cluster 84	I1	N2	I1
Cluster 85	I1	N2	I2
Cluster 86	I1	N3	N1
Cluster 87	I1	N3	N2
Cluster 88	I1	N3	N3
Cluster 89	I1	N3	I1
Cluster 90	I1	N3	I2
Cluster 91	I1	I1	N1
Cluster 92	I1	I1	N2
Cluster 93	I1	I1	N3
Cluster 94	I1	I1	I1
Cluster 95	I1	I1	I2
Cluster 96	I1	I2	N1

Cluster	T	P	Vg
Cluster 97	I1	I2	N2
Cluster 98	I1	I2	N3
Cluster 99	I1	I2	I1
Cluster 100	I1	I2	I2
Cluster 101	I2	N1	N1
Cluster 102	I2	N1	N2
Cluster 103	I2	N1	N3
Cluster 104	I2	N1	I1
Cluster 105	I2	N1	I2
Cluster 106	I2	N2	N1
Cluster 107	I2	N2	N2
Cluster 108	I2	N2	N3
Cluster 109	I2	N2	I1
Cluster 110	I2	N2	I2
Cluster 111	I2	N3	N1
Cluster 112	I2	N3	N2
Cluster 113	I2	N3	N3
Cluster 114	I2	N3	I1
Cluster 115	I2	N3	I2
Cluster 116	I2	I1	N1
Cluster 117	I2	I1	N2
Cluster 118	I2	I1	N3
Cluster 119	I2	I1	I1
Cluster 120	I2	I1	I2

Cluster	T	P	Vg
Cluster 121	I2	I2	N1
Cluster 122	I2	I2	N2
Cluster 123	I2	I2	N3
Cluster 124	I2	I2	I1
Cluster 125	I2	I2	I2

Appendix C- Graphs for Linear Model

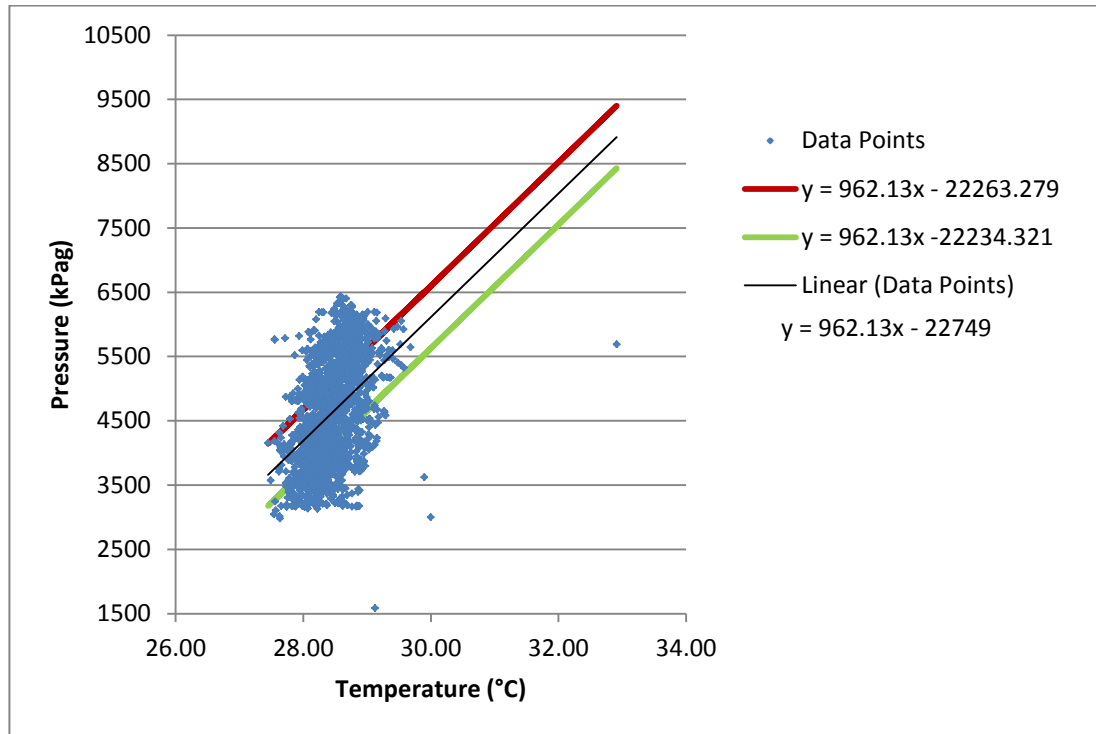


Figure C1: Graph of Pressure (kPag) against Temperature (°C)

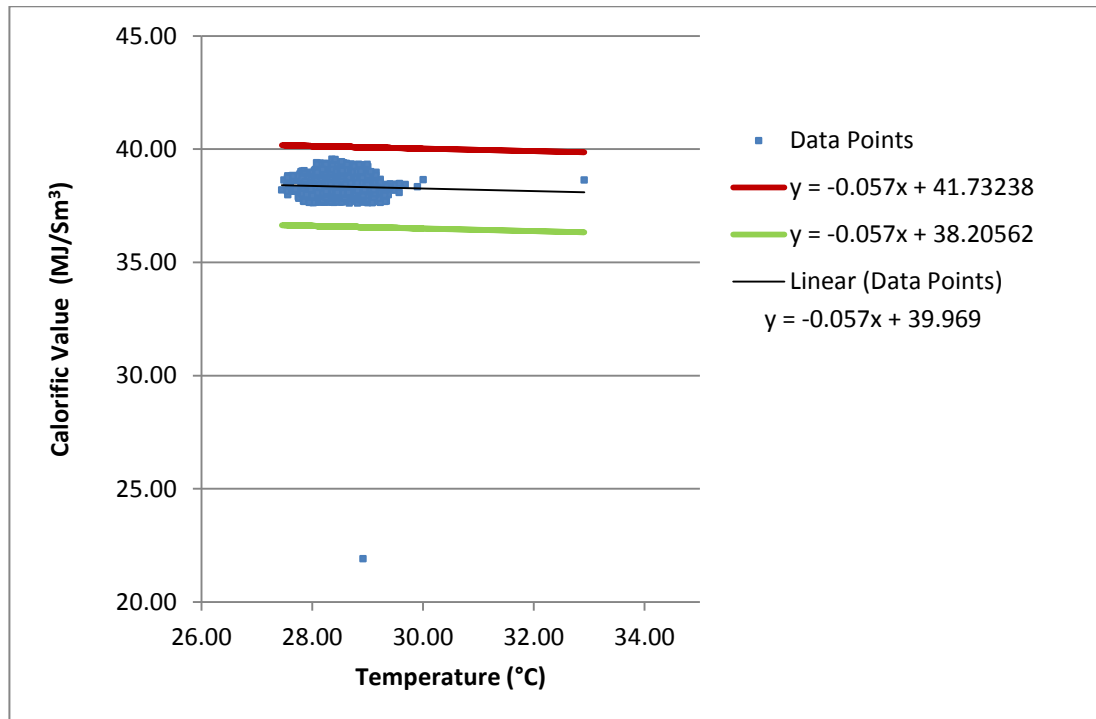


Figure C2: Graph of Pressure (MJ/Sm³) against Temperature (°C)

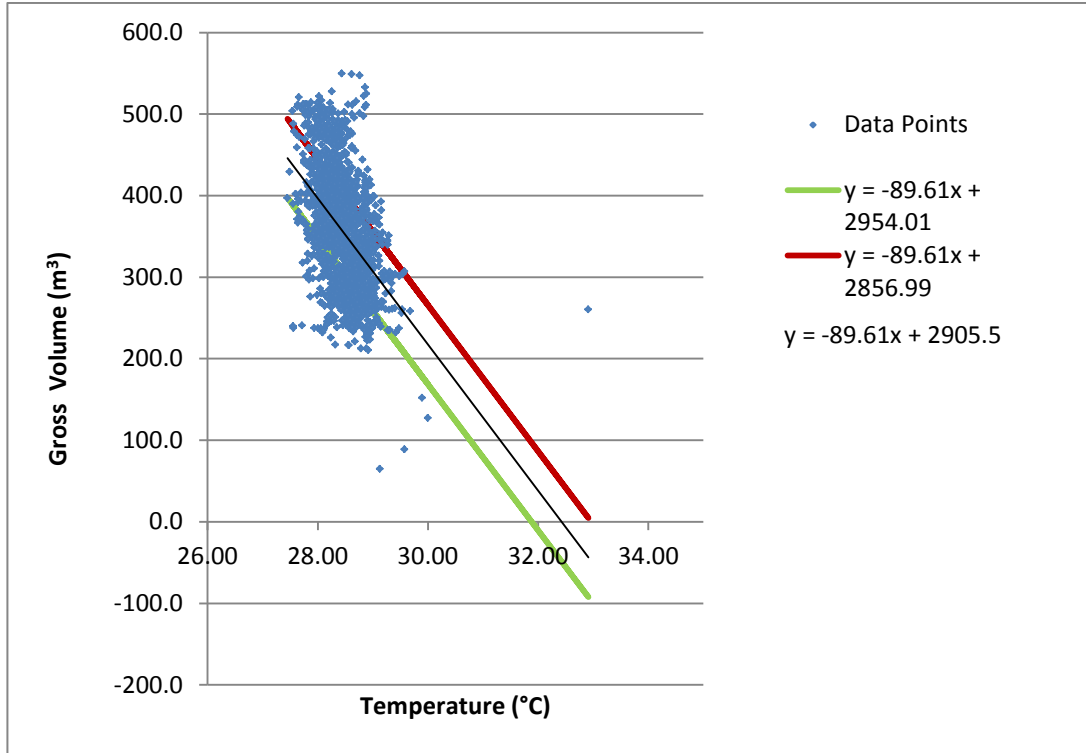


Figure C3: Gross Volume (m³) against Temperature (°C)

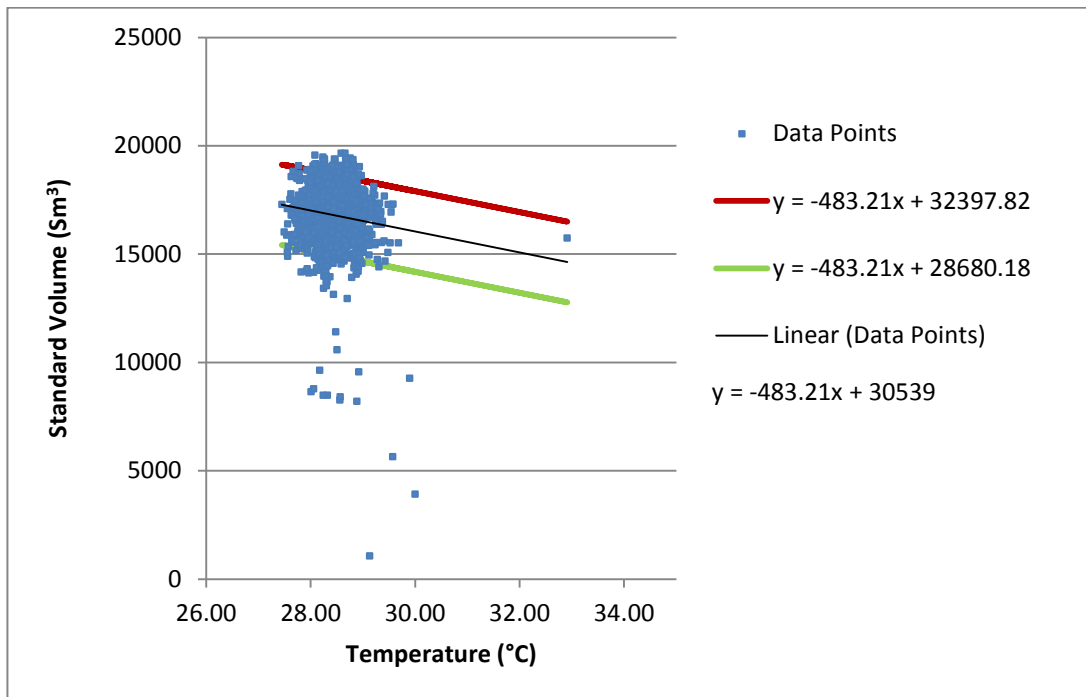


Figure C4: Standard Volume (Sm³) against Temperature (°C)

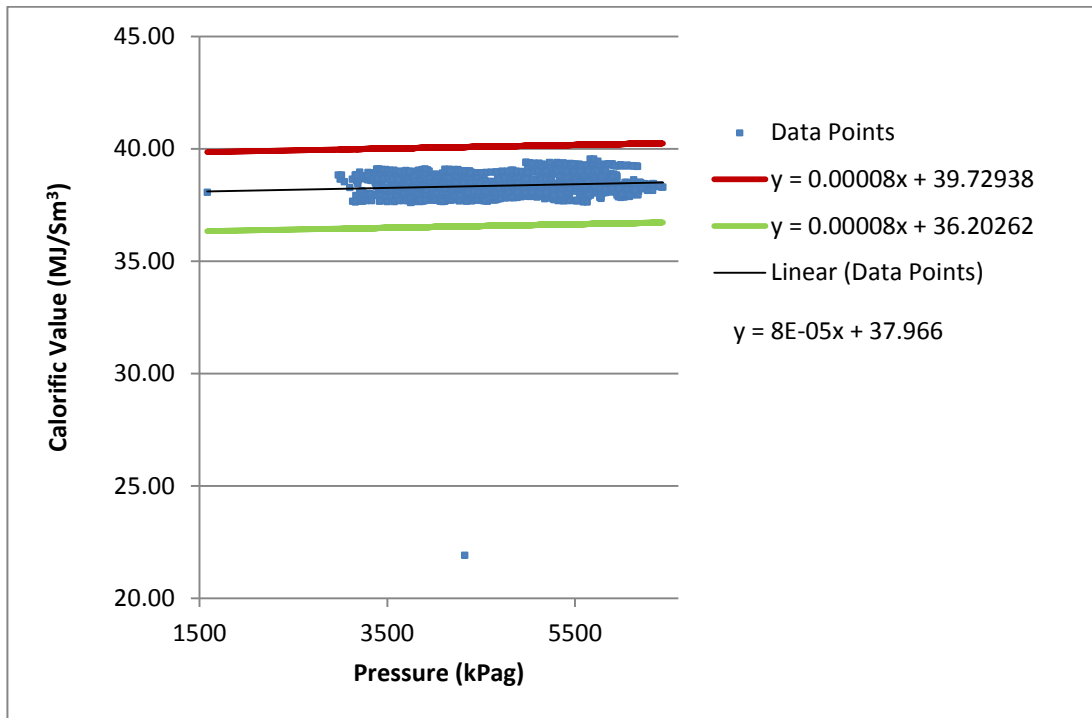


Figure C5: Calorific Value (MJ/Sm³) against Pressure (kPag)

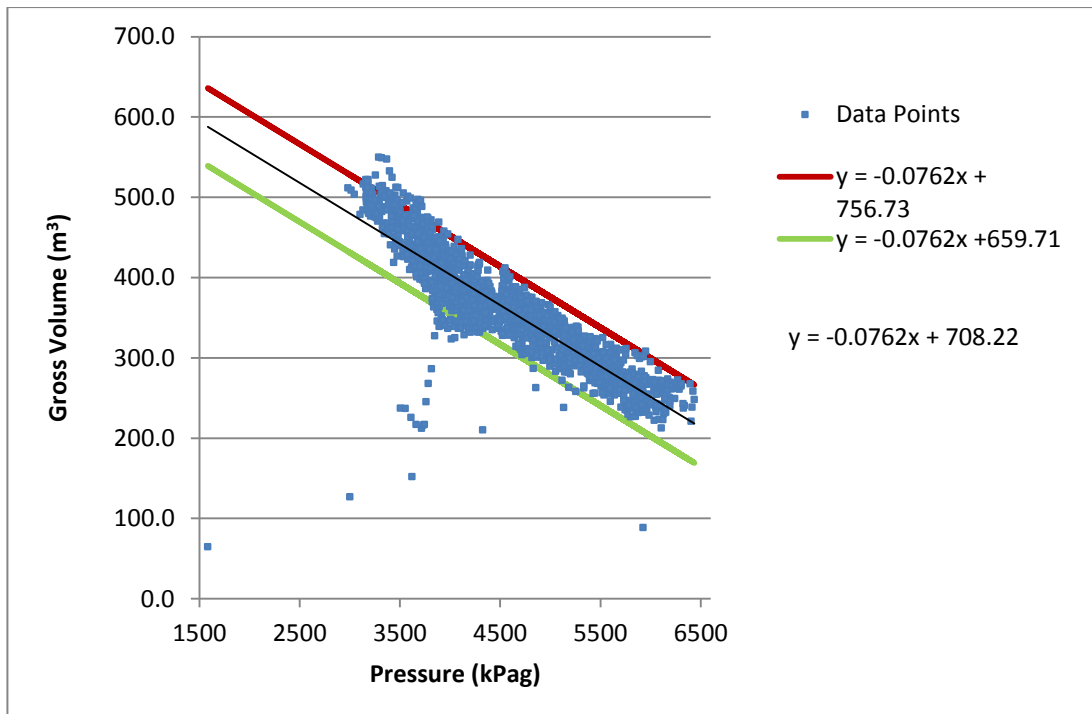


Figure C6: Gross Volume (m³) against Pressure (kPag)

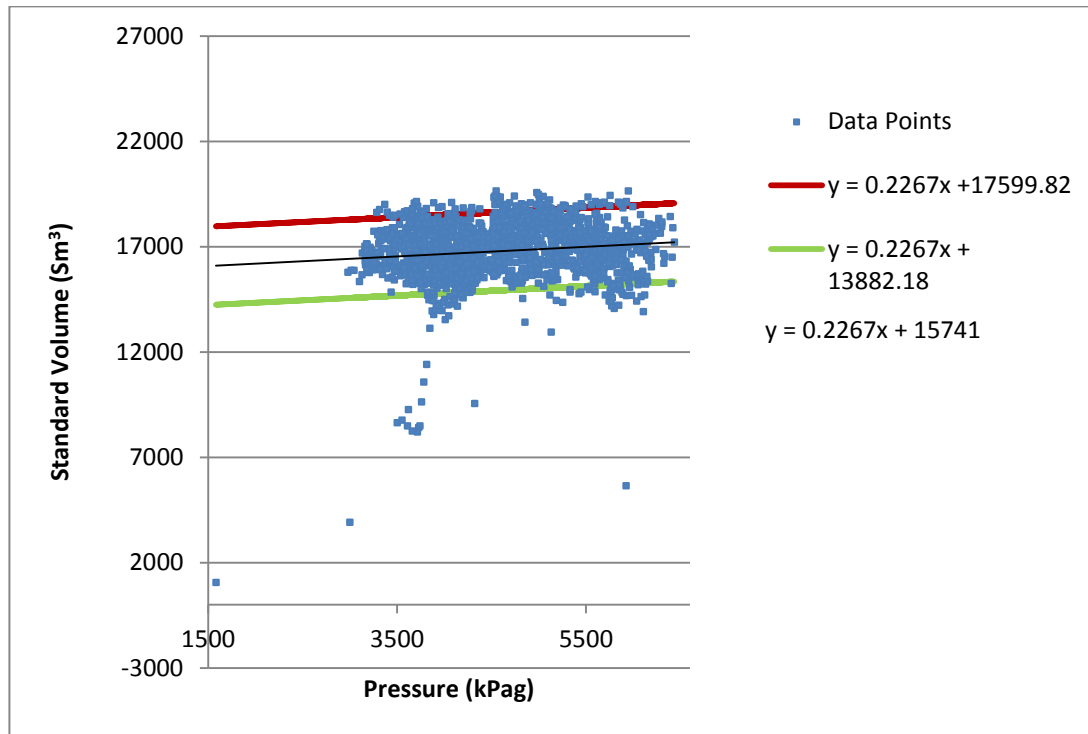


Figure C7: Standard Volume (Sm^3) against Pressure (kPag)

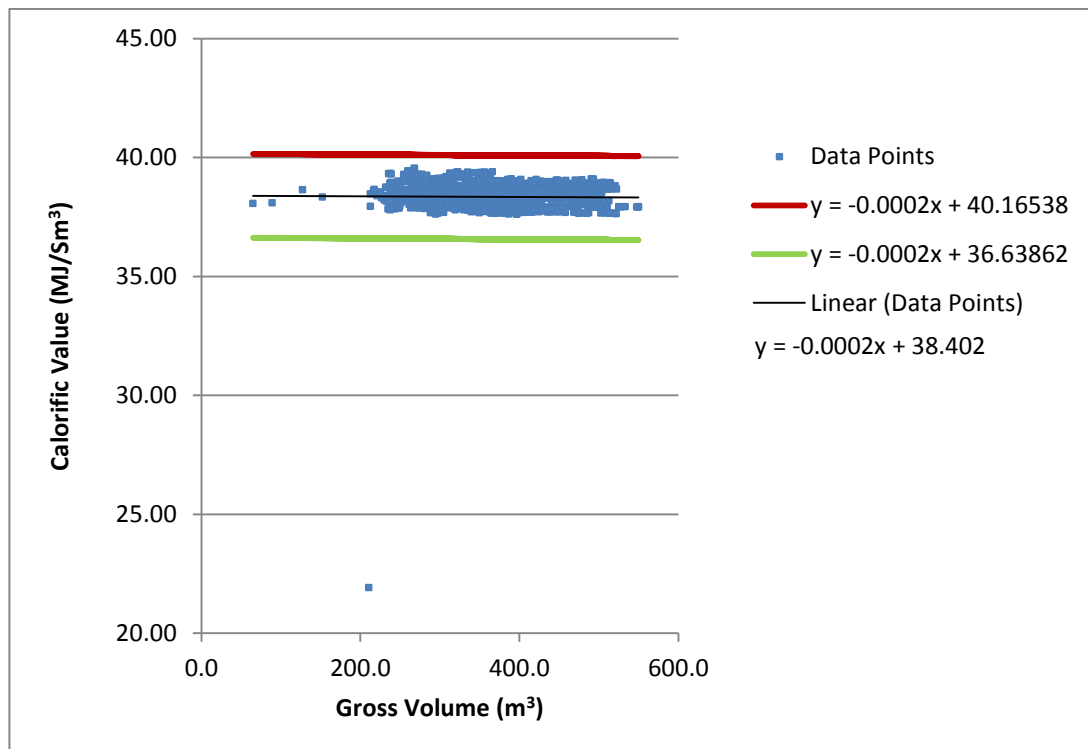


Figure C8: Calorific Value (MJ/Sm^3) against Gross Volume (m^3)

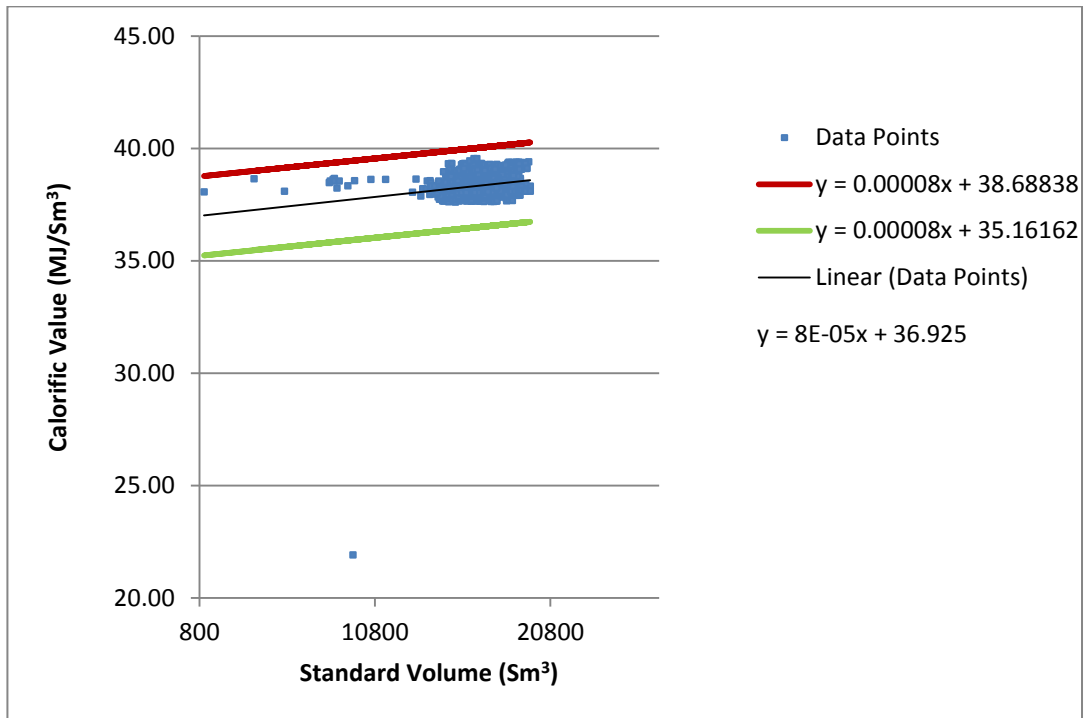


Figure C9: Calorific Value (MJ/Sm³) against Standard Volume (Sm³)

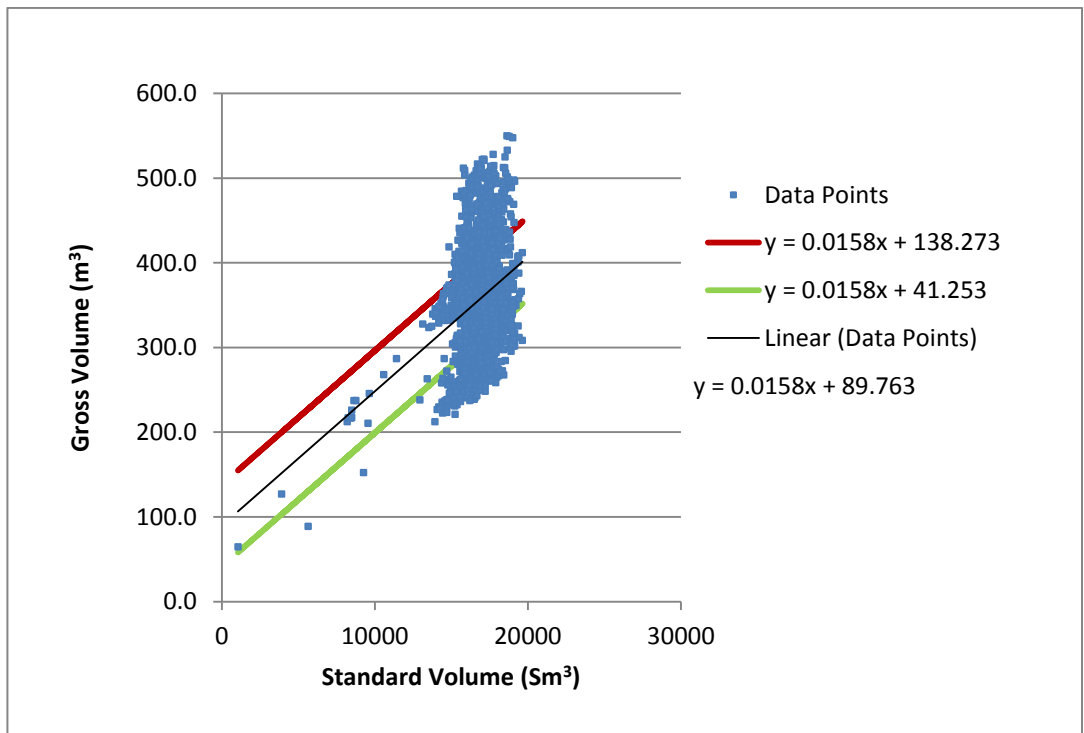


Figure C10: Gross Volume (m³) against Standard Volume (Sm³)

