

**NEURAL NETWORK PREDICTION OF FLASH POINT OF DIESEL FUEL  
FROM ITS CHEMICAL COMPOSITION AND PHYSICAL PROPERTIES**

**A MASTER'S THESIS**

**in**

**Chemical Engineering and Applied Chemistry**

**Atihm University**

**by**

**YOUNIS MUHSIN YOUNIS AL-ANI**

**DECEMBER 2017**

**NEURAL NETWORK PREDICTION OF FLASH POINT OF DIESEL FUEL  
FROM ITS CHEMICAL COMPOSITION AND PHYSICAL PROPERTIES**

**A THESIS SUBMITTED TO  
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**

**OF**

**ATILIM UNIVERSITY**

**BY**

**YOUNIS MUHSIN YOUNIS AL-ANI**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE  
DEGREE OF**

**MASTER OF SCIENCE**

**IN**

**THE DEPARTMENT OF CHEMICAL ENGINEERING AND APPLIED  
CHEMISTRY**

**DECEMBER 2017**

Approval of the Graduate School of Natural and Applied Sciences, Atılım University.

---

Prof. Dr. Ali Kara

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

---

Prof. Dr. Atilla Cihaner

Head of Department

This is to certify that we have read the thesis “Neural Network Prediction of Flash Point of Diesel Fuel From its Chemical Composition and Physical Properties” submitted by “Younis Muhsin Younis Al-ani” and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

---

Asst. Prof. Dr. Hakan Kayı

Supervisor

Examining Committee Members

Assoc. Prof. Dr. Murat Torun

Asst. Prof. Dr. Hakan Kayı

Asst. Prof. Dr. Enver Güler

---

Date: 27.12.2017

I declare and guarantee that all data, knowledge and information in this document has been obtained, processed and presented in accordance with academic rules and ethical conduct. Based on these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Younis Al-ani

Signature:

## ABSTRACT

### NEURAL NETWORK PREDICTION OF FLASH POINT OF DIESEL FUEL FROM ITS CHEMICAL COMPOSITION AND PHYSICAL PROPERTIES

Al-ani, Younis Muhsin Younis

M.S., Chemical Engineering and Applied Chemistry

Supervisor: Asst. Prof. Dr. Hakan Kayı

December 2017, 87 pages

The flash point is important in engineering calculations, so this study has two major purposes. The first purpose of the study is to predict the flash point from its chemical composition and physical properties by using artificial neural network to decrease time and cost spent on experimental analysis of flash point, and the second purpose is to find the simplest formula to predict the flash point.

Artificial Neural Networks is applied as a black-box type modeling for flash point prediction of diesel fuel. The experimental data used in this study is obtained from Erbil power station. Every truck holding diesel fuel needs to be monitored, especially for the flash point test.

In this study, the Levenberg-Marquardt training algorithm is utilized to train the neural network and to predict the flash point.

The network performance is evaluated through network test results, mean squared error analysis, regression corrections and error histograms. The findings obtained in this study indicated that the designed neural network performs quite well in the prediction of flash point of diesel fuel from its chemical composition and physical properties.

**Keywords:** Artificial neural networks, Flash point prediction, Diesel fuel, Chemical composition, Physical properties

## ÖZ

### DİZEL YAKITIN ALEVLENME NOKTASININ KİMYASAL BİLEŞİM VE FİZİKSEL ÖZELLİKLERDEN SİNİR AĞI İLE TAHMİNİ

Al-ani, Younis Muhsin Younis

Yüksek Lisans, Kimya Mühendisliği ve Uygulamalı Kimya

Tez Yöneticisi: Yrd. Doç. Dr. Hakan Kayı

Aralık 2017, 87 sayfa

Bir dizel yakıtın alevlenme noktası mühendislik hesaplamaları için önemlidir, bu nedenle bu çalışmanın iki ana amacı vardır. Çalışmanın ilk amacı, alevlenme noktasının deneysel analizine harcanan zaman ve maliyeti düşürmek için yapay sinir ağını kullanarak kimyasal bileşim ve fiziksel özelliklerinden alevlenme noktasını tahmin etmek ve ikinci amacı da alevlenme noktasını tahmin etmek için en basit formülü bulmaktır.

Yapay sinir ağları, dizel yakıtın alevlenme noktası tahmininde kara kutu tipi bir modelleme olarak uygulanır. Bu çalışmada kullanılan deneysel veriler Erbil enerji santralinden elde edilmiştir. Dizel yakıt depolayan her kamyon, özellikle de alevlenme noktası testi için izlenmelidir.

Bu çalışmada, Levenberg-Marquardt eğitim algoritması sinir ağını eğitmek ve alevlenme noktasını tahmin etmek için kullanılmıştır. Ağ performansı, ağ test sonuçları, ortalama kareli hata analizi, regresyon düzeltmeleri ve hata histogramları ile değerlendirilmiştir. Bu çalışmada elde edilen bulgular dizel yakıtın alevlenme noktasını kimyasal bileşiminden ve fiziksel özelliklerinden tahmin etmek üzere tasarlanmış sinir ağına oldukça iyi performans gösterdiğini ortaya koymuştur.

**Anahtar Kelimeler:** Yapay sinir ağları, Alevlenme noktası tahmini, Dizel yakıt, Kimyasal bileşim, Fiziksel özellikler

To my Parents, Brothers Dr. Asif, Hasam and Dr. Mustafa

To my dear wife and kids, for supporting me throughout my academic career. Without their moral support, encouragement for academic work, the completion of this effort would not have been possible.

## **ACKNOWLEDGMENTS**

I am extremely grateful to my supervisor Asst. Prof. Dr. Hakan Kayı for his support and guidance. He has helped me at various stages during my studies, by providing me with a well-defined thesis problem and willingly sharing his knowledge and ideas with me.

I would like to express my sincere gratitude to Prof. Dr. Erdoğan Alper for his sound advice in the evaluation of the results. His active participation at every facet of my research work has shaped my overall outlook.

I am grateful to the Mass Group Holding for permission to work in a laboratory at Erbil Power Station. I must thank Chairman of the Board Mr. Ahmad Ismail Saleh, Plant Manager Mr. Azad Abdukhalik Hasan and all staff work in the laboratory.

## TABLE OF CONTENTS

ABSTRACT.....	I
ÖZ.....	II
ACKNOWLEDGMENTS .....	IV
TABLE OF CONTENTS.....	V
LIST OF TABLES .....	IX
LIST OF FIGURES .....	X
LIST OF ABBREVIATIONS.....	XII
CHAPTER 1 .....	1
INTRODUCTION .....	1
1.1 Diesel Use and Challenges .....	1
1.2 About Hydrocarbons .....	3
1.2.1 Paraffins.....	3
1.2.2 Naphthenes .....	4
1.2.3 Olefins.....	4
1.2.4 Aromatics.....	5
1.3 Other Compounds .....	6
1.4 Diesel Fuel Chemistry.....	7
1.5 Diesel Fuel Parameters.....	9
1.5.1 Cetane Number .....	9
1.5.2 Aromatic Content.....	10

1.5.3 Density .....	10
1.5.4 Sulfur Content.....	11
1.5.5 Distillation .....	12
1.5.6 Viscosity .....	12
1.5.7 Cloud and Pour Points .....	13
1.5.8 Flash Point (Safety Factor).....	14
1.5.9 Smoke Point.....	15
1.6 Specification of Diesel Fuel .....	15
1.7 Neural Networks .....	16
1.7.1 Why Neural Networks? .....	17
1.7.2 How Does Human Brain Learn?.....	17
1.7.3 Artificial Neuron Model .....	19
1.7.4 Artificial Network Architecture.....	20
1.7.4.1 Feedforward Neural Network .....	20
1.7.4.2 The Learning Process .....	21
1.7.4.3 The Backpropagation Algorithm .....	21
1.8 Thesis Objectives and Problem Statement .....	23
CHAPTER 2 .....	24
BACKGROUND INFORMATION AND LITERATURE SURVEY .....	24
2.1 Introduction .....	24
2.2 Flash Point Prediction Model for Pure Compounds.....	25
2.3 Empirical Flash Point Prediction Based on Normal Boiling Point and Composition .....	26
2.4 Flash Point Prediction by Mathematical Regression Methods.....	27
2.5 Flash Point Prediction Using the ANN Methods .....	28

CHAPTER 3 .....	33
METHODOLOGY .....	33
3.1 Introduction .....	33
3.2 Experimental Work (Stage One) .....	35
3.2.1 Experiment on Sample (A) .....	35
3.2.1.1 Flash Point HFP 339 According to ASTM D93 .....	35
3.2.1.2 Fuel Analyzer (Cetane 2000) .....	38
3.2.2 Experiments on Sample (B) .....	41
3.2.2.1 Flash point HFP 339 using ASTM D 93 .....	41
3.2.2.2 Automatic Kinematic Viscosity Measurement System .....	41
3.2.2.3 Density Model (DMA <sup>TM</sup> ) .....	42
3.2.2.4 X-RAY Sulfur Meter .....	44
3.2.2.5 Automatic Distillation .....	45
3.3 ANN Modelling (Stage Two) .....	46
3.3.1 Introduction .....	46
3.3.2 The Basic Specification of the Model (Stage two) .....	47
3.3.2.1 Input Layer .....	47
3.3.2.2 Output Layer .....	48
3.3.2.3 Hidden Layer .....	48
3.3.3 Model Design for Sample (A) .....	49
3.3.4 Model Design for Sample (B) .....	52
CHAPTER 4 .....	54
RESULTS AND DISCUSSION .....	54
4.1 Sample (A) .....	54
4.1.1 Hidden Nodes .....	54

4.1.2 The Network Performance for Sample (A) .....	55
4.1.3 The Regression Plots for Sample (A) .....	56
4.1.4 Error Histogram of Sample (A) .....	57
4.1.5 Test of Network for Sample (A) .....	58
4.1.6 The Simplest Formula to Predict the Flash Point for Sample A.....	58
4.1.7 Summary of the Results for Sample (A).....	59
4.2 Sample (B).....	60
4.2.1 Hidden Nodes for Sample (B) .....	60
4.2.2 The Network Performance for Sample (B).....	61
4.2.3 The Regression Plots for Sample (B) .....	62
4.2.4 Error Histogram of Sample (B) .....	63
4.2.5 Test of Network for Sample (B) .....	64
4.2.6 The Simplest Formula to Predict the Flash Point for Sample (B) .....	65
4.2.7 Summary of the Results for Sample (B).....	65
CHAPTER 5 .....	67
CONCLUSIONS.....	67
REFERENCES .....	69
APPENDIX.....	73
A.    DATA RELATIVE TO CHAPTER 3 (SAMPLE A) .....	73
B.    DATA RELATIVE TO CHAPTER 3 (SAMPLE B).....	78
C.    POLYMATH REPORT SAMPLE (A) RELATIVE TO CHAPTER 4 .....	83
D.    POLYMATH REPORT SAMPLE (B) RELATIVE TO CHAPTER 4 .....	85

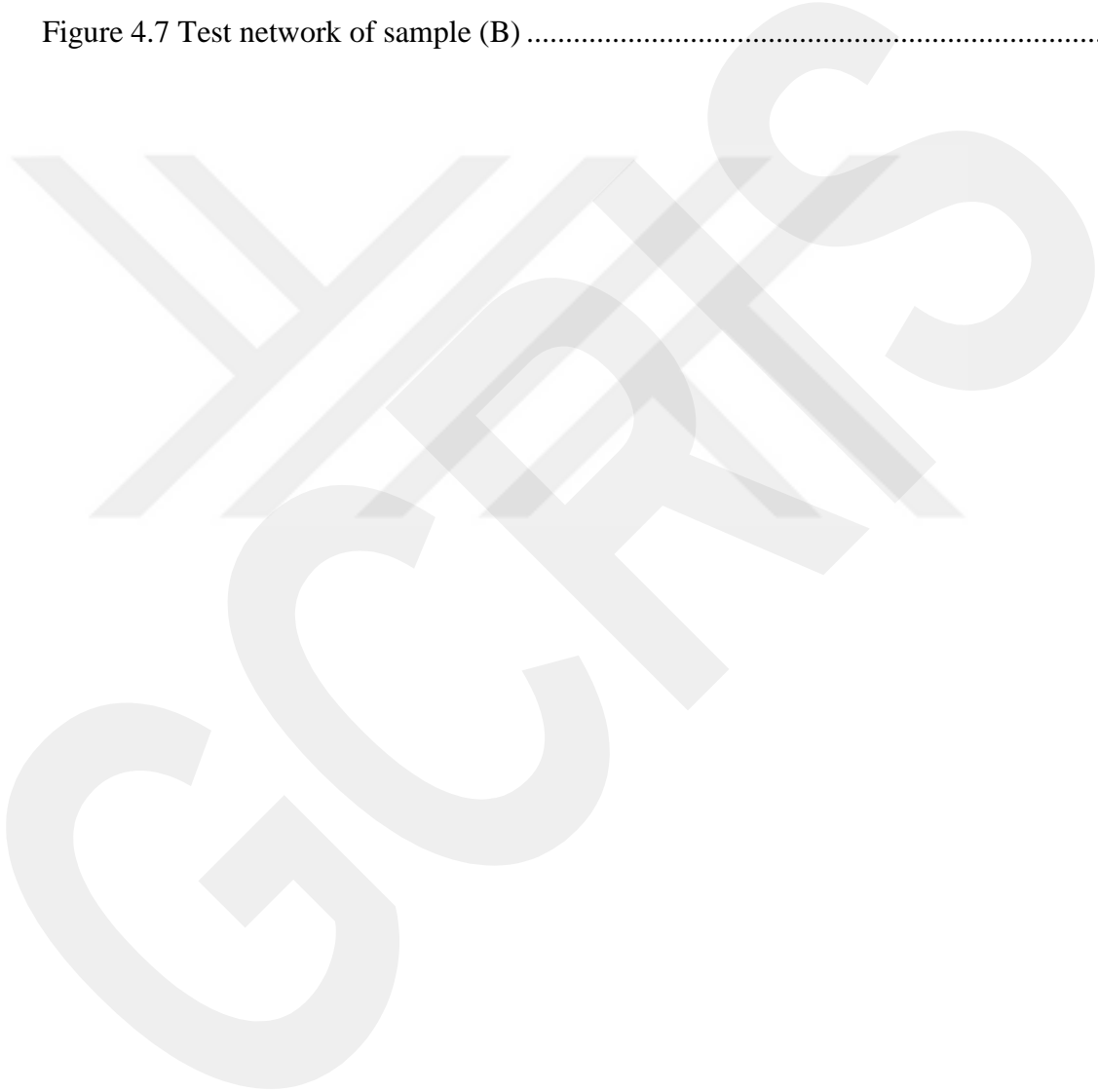
## LIST OF TABLES

Table 1.1 Specification of diesel fuel.....	16
Table 2.1 The standard flash point testing methods for various types of samples that follow the American Society of Testing and Materials (ASTM) standards .....	25
Table 2.2 Mean square error after generalization .....	31
Table 3.1 Input layer for sample (A).....	47
Table 3.2 Input layer for sample (B).....	47
Table 3.3 The best prediction for sample (A) .....	50
Table 3.4 The best prediction for sample (B) .....	52
Table 4.1 Hidden nodes regression results of sample (A) .....	55
Table 4.2 Constants of the equation for sample (A) .....	59
Table 4.3 Summary of the results for sample (A).....	59
Table 4.4 Hidden nodes regression results of sample (B).....	60
Table 4.5 Constants of the equation for sample (B) .....	65
Table 4.6 Summary of the results for sample (B).....	65

## LIST OF FIGURES

Figure 1.1 Typical carbon number distribution – diesel fuel.....	8
Figure 1.2 Typical distillation profile – diesel fuel.....	8
Figure 1.3 Apparatus used for both ASTM distillations.....	12
Figure 1.4 Cloud and pour points apparatus .....	14
Figure 1.5 Flash point apparatus .....	15
Figure 1.6 Interconnection of biological neural network .....	18
Figure 1.7 Synaptic junction.....	18
Figure 1.8 Basic artificial neuron .....	20
Figure 1.9 Feedforward neural network with one hidden layer.....	21
Figure 2.1 Architecture of the MLP network used to predict physical property parameters of diesel. ....	29
Figure 2.2 Flow chart for the training methodology of artificial neural networks. ....	30
Figure 3.1 Flow chart of the methodology used .....	34
Figure 3.2 Front view of the HFP339 .....	36
Figure 3.3 Unit with gas ignition .....	36
Figure 3.4 Flash point test result.....	38
Figure 3.5 Cetane 2000 apparatus.....	39
Figure 3.6 Fixed arm features .....	40
Figure 3.7 Standard sample run set-up .....	40
Figure 3.8 Sample preparation .....	40
Figure 3.9 Viscosity tester .....	42
Figure 3.10 Density instrument.....	43
Figure 3.11 X-ray sulfur instrument .....	45
Figure 3.12 Distillation unit tester .....	46
Figure 3.13 Flow chart of the second stage ANN model.....	49

Figure 4.1 Network performance of sample (A).....	55
Figure 4.2 Regression plots for sample (A).....	56
Figure 4.3 Error histogram of sample (A).....	57
Figure 4.4 Network performance of sample (B).....	61
Figure 4.5 Regression plots of sample (B).....	62
Figure 4.6 Error histogram of sample (B).....	63
Figure 4.7 Test network of sample (B).....	64



## LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
ARO	Aromatic
ARO-JET	Aromatic jet fuel
ASTM	American Society of Testing and Materials
CN	Cetane Number
FAME	Acid Methyl Ester (biodiesel)
FP	Flash Point
L-M	Levenberg -Marquardt BackPropagation
MSE	Mean Square Error
MLP	Multiple linear regression
NN	Neural Network
NARX	Nonlinear Autoregressive with External input
PNA	Polynuclear Aromatic
R	Linear Correlation Coefficient
R <sup>2</sup>	Coefficient of Determination
RMSE	Root Mean Square Error

# CHAPTER 1

## INTRODUCTION

### 1.1 Diesel Use and Challenges

At the present time, diesel engine is quite well known in numerous applications. The reputation of the diesel engines is growing as it is frequently used in commercial and personal vehicles. The issues with using diesel engines concentrate on two primary concerns: the overall quantity of diesel fuel oil utilized the exhaust secretions of the combustion process and the limited amount of crude oil reserves. The earth's oil reserves are predicted to last for only four decades. If demand for oil outweighs the supply, the price of oil is anticipated to increase. Major problems could be caused due to this, not only in the operation of engines but also the stability in the world and in the most damaging case, cause a widespread pandemonium. Thus, substitute fuels need to be found so as to enable the use of engines in the coming future [1].

During the late eighteenth century, the development of the internal combustion engine was vanished. Slow but steady advancements had been made over this century. Rudolf Diesel received a patent in 1892 to compress the ignition engine. However, its initial design, which uses coal dust and fuel, did not work. In the same century, in 1859, crude oil was found out in Pennsylvania. Lamp oil (kerosene) was the first product used from crude oil. Rudolf discovered that the change in fuel (along with some changes of mechanical design) led to design of a successful prototype engine in 1895. Thus a working prototype engine was born in 1895, with some mechanical modifications along with the change in fuel [1].

Consumption rate of diesel fuel in diesel engines is extremely high. For instance, the road usage of diesel engines has an eighty one percent portion of the total energy usage in the transportation sector. As suitable alternative fuels come into existence, this leads to an issue with the volume of production of these fuels. The reduction of fuel consumption is essential in this present situation. The pollution in the entire world is in fact greatly affected by exhaust gas emissions of diesel engines [2] .

In the current time, individuals require a cost effective fuel to meet their demands as the prices of fuel are rising due to increase in demand and decrease in fuel supply. The diesel engine industry has grown to be greatly cost effective and efficient due to Rudolph Diesel efforts. Gasoline is relatively cheaper than diesel fuel however diesel has a higher energy density. This means that larger amounts of energy can be obtained from diesel as compared with the same volume of gasoline. As a result, diesel engines in automobiles provide greater mileage. The merits of diesel engines can make them an appropriate choice for heavy duty vehicles and equipment. Diesel is oiler and heavier than gasoline, and has a boiling point higher than water. Diesel engines attract more attention than other engines due to their efficiency and cost effectiveness.

Diesel fuel applies an important function in reinforcing of the world economy and the standard of living from the electric power production, to improve the efficiency of farms and consuming goods carried around the globe. The main uses for diesel fuel [1]:

- Transportation
- Shipping
- Electric power generation
- Military transportation
- Road transportation
- Off road uses (e.g. mining and construction)
- Farming

Diesel engine is also called the diesel cycle after Rudolf Diesel, who invented it in the 19<sup>th</sup> century. While old diesel engines may not be applicable to our definition of an alternative source of energy, they are still a key addition to a remote power or network backup system. According to the 2011 Annual Energy Review by the Energy Information Administration (EIA), the consumption of the total energy in the US involves 36 % coming from

petroleum, 26 % coming from natural gas, 20 % coming from coal , 9 % coming from renewable energy, and 8 % coming from nuclear-electric power. Also, the energy consumption in the US was categorized by sectors as transportation sector consumed about 28 %; industrial sector consumed about 21 %; residential and commercial sectors consumed about 11 % and electric power sector consumed about 40 % [3].

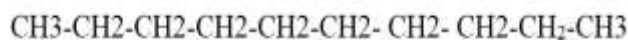
The usage of petroleum based liquid fuel decreased in 2012 by 2.1 % and later in 2013 rose by 1.7 %. Clearly, transportation sector has the maximum consumption of liquid fuel between all the sectors. EIA predicts petroleum based liquid fuel fall of by 0.4 % in 2014 [4].

## 1.2 About Hydrocarbons

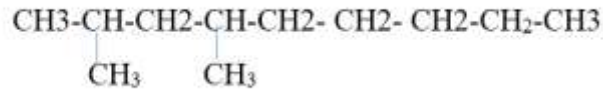
Hydrocarbons can be classified as organic compounds completely made up of hydrogen and carbon atoms and there are four main categories of hydrocarbons ; paraffin, olefin naphthene, and aromatic. Each class belongs to a family of hydrocarbon molecules which share the same structure of mutual, but vary in number of carbon atoms or geometry. The classes also vary in the ratio of carbon to hydrogen atoms and the fashion in which the carbon atoms are attached to the other carbon atoms in the molecule [1] .

### 1.2.1 Paraffins

$C_nH_{2n+2}$  is a general formula for paraffin, in which (n) is the numbers of carbon atoms within that molecule. Paraffins contain two categories; normal paraffins and isoparaffins. Paraffin chains are formed by joined carbon atoms such as a molecule. Isoparaffins have a likewise carbon structure, but they also contain one or more carbons branching off from the structure. For instance 2, 4 dimethyloctane and decane have the same chemical formula  $C_{10}H_{22}$ , but their physical properties and compositions are different. The compounds are similar to this, with a similar chemical structure but a dissimilar grouping of atoms are well- known as isomers [1] .



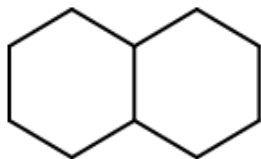
**n-Decane  $C_{10}H_{22}$**



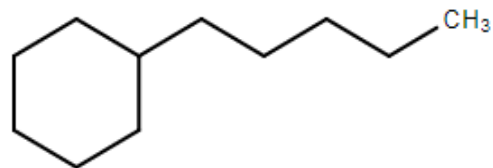
2,4 – Dimethyloctane

### 1.2.2 Naphthenes

The general formula of the naphtha is  $C_nH_{2n}$  and has one or more rings, Naphthenes contain few of their carbon atoms assorted as a ring. Naphthalene in diesel fuel contain rings of carbons 5 or 6. In some cases, they were joined by two rings or more together, with little carbons shared by next to neighbouring rings. Naphthene is the phrase used in the oil industry to explain saturated cyclic or ring hydrocarbons, these compounds are also referred to as cycloalkane and cycloparafins [1].



Decalin  $C_{10}H_{18}$



Butylcyclohexane  $C_{10}H_{20}$

### 1.2.3 Olefins

General formula  $C_nH_{2n}$  refers to olefins with a single double bond, they are like paraffins, however; they have fewer hydrogen atoms and have a minimum of one double bond between carbon atoms. Olefins rarely happen in crude oil; similar to paraffins, olefins with four or more carbons can happen as structural isomers [1].



1- Decene  $\text{C}_{10}\text{H}_{20}$

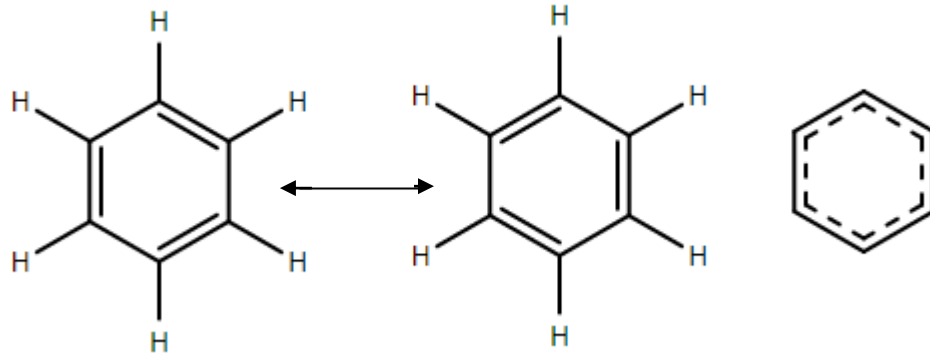
#### 1.2.4 Aromatics

Aromatic hydrocarbon rings are formed of six carbon atoms. Some carbon atoms are arranged in the ring, which is similar to naphthenes and the most basic aromatic compound is benzene and structure was initially imagined as two equal structures with double bonds.

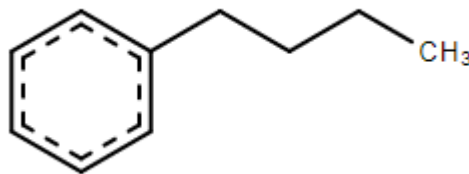
One ring aromatic compounds has general formula  $\text{C}_n\text{H}_{2n-6}$ . Shorthand representation is used to present a hexagon with a circle inside the imaging of aromatic bonds. Polycyclic aromatics are compounds with two or more of aromatic rings. These aromatic rings join together and share some neighboring rings carbonates. Naphthenes and paraffins are classified as saturated hydrocarbons due to the fact that no more hydrogen can be added except if the carbon backbone is broken. Olefins and aromatics classified as unsaturated hydrocarbons, and composed of carbon-carbon bonds or aromatic bonds [1].

While olefins saturate with hydrogen, they are converted into paraffins; when partially saturated, they lead to cyclic olefins, when aromatics are saturated with hydrogen, they lead to naphthenes [1].

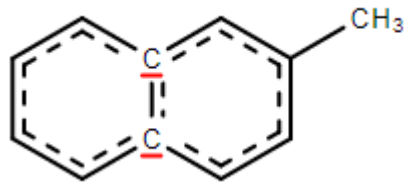
Certain molecules comprise of structural properties of two or more hydrocarbons. For instance, multiple methods can be classified to a molecule containing ring for every an aromatic, a naphthenic and a paraffinic. Thus, chemists determined the hierarchy of hydrocarbon structural characteristics, with aromatics at the top, then olefins and naphthenes, and paraffins at the bottom. A composition containing properties of two classes or more is put in a higher class in the hierarchy. As a result, in this instance, the aromatic molecule is classified [1].



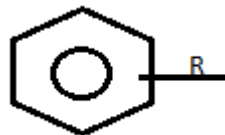
Two equivalent structure of benzene  $C_6H_6$



Butylbenzene  $C_{10}H_{14}$



2-Methylnaphtalene  $C_{11}H_{10}$  Polynuclear aromatic (PNA) or polycyclic

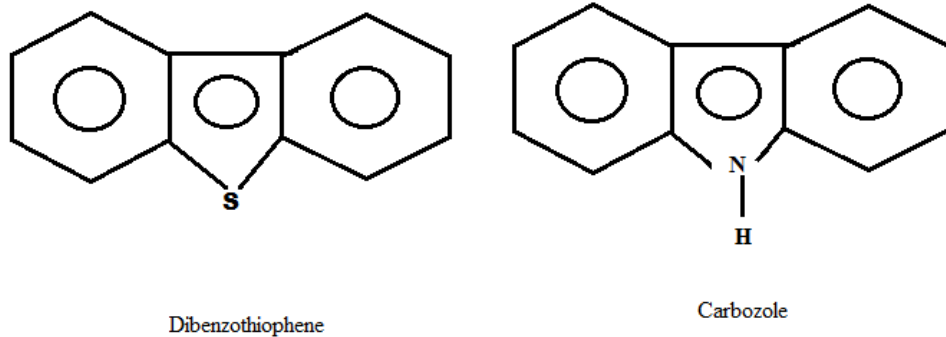


(R) represents an alkyl group that can be linked with any carbons with the benzene ring

### 1.3 Other Compounds

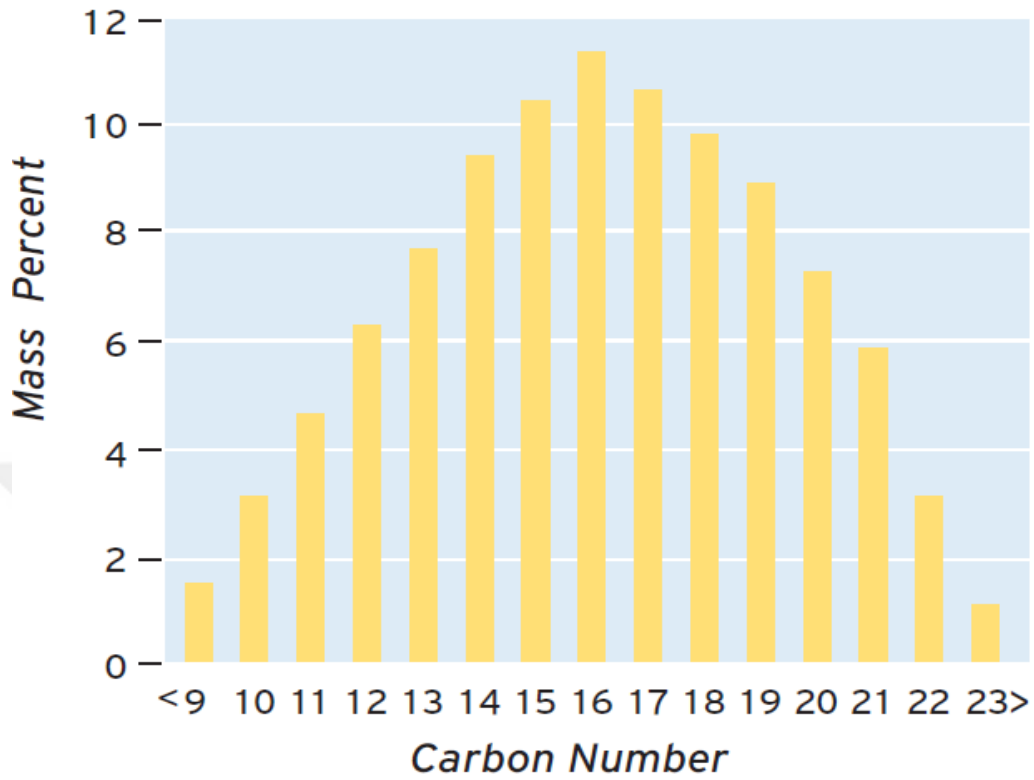
Although hydrogen and carbon are common elements in diesel fuel, other elements will be found in this fuel such as, the low amounts of sulfur, oxygen, and nitrogen. These

elements are known as heteroatoms (found in the carbazole and dibenzothiophene molecules, and these examples are available in diesel fuel). Even though these compounds are only in minute amounts, they have a major role in reducing specific diesel properties. [1].

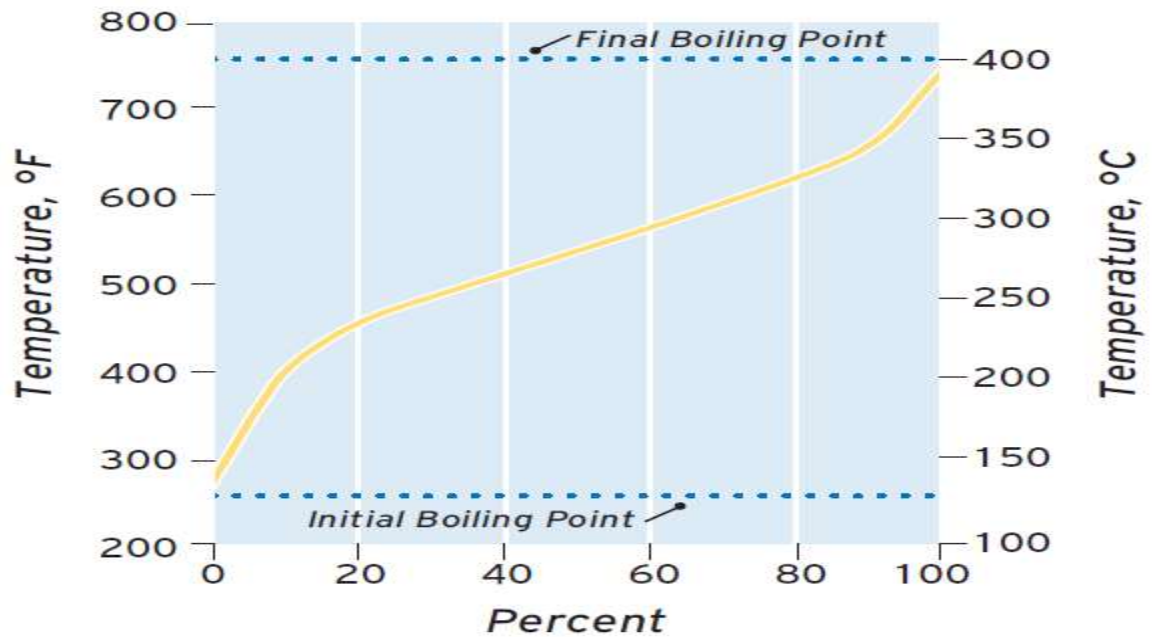


#### 1.4 Diesel Fuel Chemistry

*“Diesel fuel is a very complex mixture of thousands of different compositions”* [1]. Most of them are with carbon numbers between eleven and twenty-two. A large part of these compositions are constituents of the paraffinic, naphthenic, or aromatic category of hydrocarbons, where every category has distinct chemical and physical properties. Figure 1.1 shows a standard carbon number allocation for diesel fuel and Figure 1.2 shows a usual distillation profile [1].



**Figure 1.1** Typical carbon number distribution – diesel fuel [1]



**Figure 1.2** Typical distillation profile – diesel fuel [1]

## **1.5 Diesel Fuel Parameters**

The components of diesel fuel are hydrocarbon chains, specifically unsaturated (olefin), saturated hydrocarbons (paraffin), and aromatic chains that contain from 10-19 carbon atoms, with a boiling point range from approximately 400 to 670 °F [5].

The following are the main parameters that are subjected to diesel fuel specifications in terms of engine performance or emissions [6].

1. Cetane number
2. Aromatic content
3. Density
4. Sulfur content
5. Distillation
6. Viscosity
7. Cloud and pour points
8. Flash point
9. Smoke point

### **1.5.1 Cetane Number**

In the 1930s, this diesel testing procedure was developed by the Cooperative Fuel Research Committee (CFR) and it was joint after that by ASTM D613. The fuel is run in a cylinder engine with a constantly variable compression ratio for this method. This is done within a fixed set of conditions. This method is still based originally on the original engine design though it has been updated over time [1].

This is the result of the engine test for diesel fuel ignition delay, therefore, two types of reference fuel are selected for this test. The first one is normal cetane C16 and an isomer of cetane known as heptamethylnane. 100 is the arbitrary number given to the normal cetane, whereas the isomer with another reference fuel has specific cetane number of fifteen. The fuel that will be tested runs in a standard experimental engine. The number of cetane was found out by contrasting the ignition period of the test diesel to a mixture of two reference fuels, and the number of cetane is derived utilizing the equation below [6], [7].

$$\text{Cetane Number} = \text{normal cetane \%} + 0.15 \times \text{heptamethylnonane \%}$$

### 1.5.2 Aromatic Content

Single ring aromatics, multiring or polyaromatic hydrocarbons can have their measured aromatic amount of diesel fuel. Some studies show that decreasing the aromatics causes each of the regulated emissions to be reduced, while other studies show that the low emissions of non-burning hydrocarbons, and particles can only occur by reducing polycyclic aromatics or multi ring aromatics [6] .

The reduction of total aromatics do not effect on the emissions of particulate matter (PM) and the hydrocarbons shown by the majority of studies. On the other hand, nitrogen oxides  $\text{NO}_x$  emissions are reduced when the total aromatics are decreased from 30 % to 10%. Studies carried out on the effect of PNA hydrocarbons depict that reducing tri and di aromatics decreases emissions of nitrogen oxides, hydrocarbons (HC) and particulate matter [8] .

### 1.5.3 Density

Mass per unit volume is the way to calculate density, in this way, low-density diesel fuel needs a greater injection time to volume, a similar mass of diesel fuel in the cylinder volume [6].

Specific gravity (SG) of the liquid system known as liquid density to water, and thus, it is a quantity that is dimensionless. But, the temperature at which the specific gravity is stated is necessary to be specified. Relative density versus absolute density is another name for specific gravity [9]

$$API = \frac{141.5}{SG} - 131.5$$

Density ASTM D405, density, relative density and API gravity of liquid petroleum products and crude petroleum are determined by hydrometer measure. The fuel is poured to a cylinder container and a hydrometer is prudently lowered into the cylinder and leave

to be steady. Then let the sample temperature calibration, the number record on the hydrometer scale at the sample surface and the sample temperature are taken in the record. The API gravity always should be mentioned that degree, the hydrometer test can be found at 60<sup>0</sup>F using the following equation [10].

$$\text{Specific gravity} = 141.5 / (131.5 + \text{API})$$

The density increases when the number of carbon compounds that are of the same class. The order of increasing density is aromatic, paraffin, and naphtha for the compounds that have the same number of carbon. For these compounds, increasing heating combustion value by category is aromatic, naphthene and paraffin on the basis of their molecular weight. On the other hand, a reverse order takes place for on the basis of volume, the aromatic has the highest point and paraffin is being at the lowest point. The same order is true with fuel. Lighter fuels, such as gasoline, it have a greater heating value on a weight basis, while fuel is heavier, for example diesel has the highest heating value based on its volume [1].

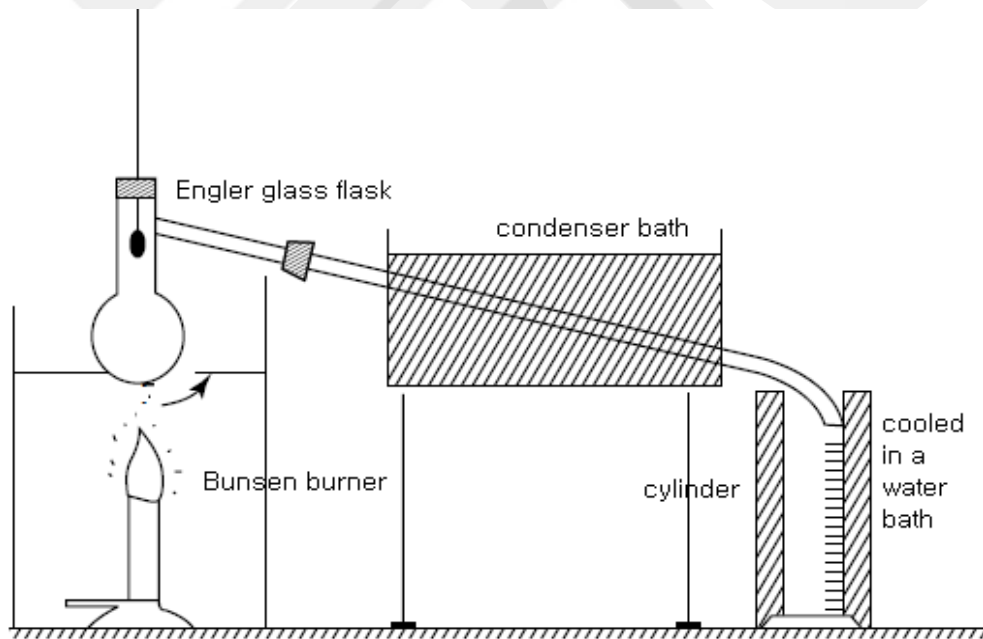
#### **1.5.4 Sulfur Content**

The sulphur content of diesel should be reduced if the pollution ratio caused by fuel is high. Catalytic desulfurization process is done in refineries to remove sulphur from diesel oil. Thus, the content of sulphur should be not more than 0.25% of the oil weight [7]. In fact, a common proportion of a sulphur is about 0.05% of the diesel fuel weight. Most of the time, this amount of sulfur can only be achieved by property hydrotreating or hydrocracking to remove the sulfur from the middle distillation which should be used in the fuel product mixture [6]. It is recommended that the sulfur content of diesel fuels should not overtake 0.05 wt % and the flash point should be not less than 55 °C [11]. Sulfur dioxide is composed of burning sulfur in diesel fuel, a part of which is oxidized to a sulfate which forms a bond with water to compose part of particles. The make contribution of sulfates to particulates is safely small since only a small percentage of total sulfur in fuel is oxidized to sulfates. But, if it uses a catalyst oxidation to decrease the emission of hydrocarbons, carbon monoxide (CO), and particle then a large amount of the sulfur dioxide (SO<sub>2</sub>) is changed to sulfates thus causing an important make a contribution to particles in the emission process [6].

### 1.5.5 Distillation

The most important separation method is distillation in a petrol refinery, and petroleum products are known through the range of boiling temperature. In addition, the distillation process is used to separate crude oil [1].

Two types of ASTM used in the refinery at the finished product. They are known as the ASTM D86 and ASTM D156. The D86 is utilized for the distillation of diesel fuel and naphtha and for kerosene and the D159 is utilized for the distillation of atmospheric gasoil. The main distinction between these two tests is the sample's volume used up. With regards to the D86, the sample is poured 100 ml of fuel into the Engler glass flask while the D156 should be 200 ml. In Figure 1.3 shows the instrument uses ASTM D 86 [6].



**Figure 1.3** Apparatus used for both ASTM distillations [6]

### 1.5.6 Viscosity

*“The measure of its resistance to a flow of any petroleum product is the viscosity”* [7]. This measurement is essential for many aspects of the design process and is actually an important quality for a lot of finished products [7]. Viscosity is mainly related to molecular

weight, and not to hydrocarbons category. Naphthenes usually have slightly larger viscosity than paraffins or aromatics for a given carbon number [1].

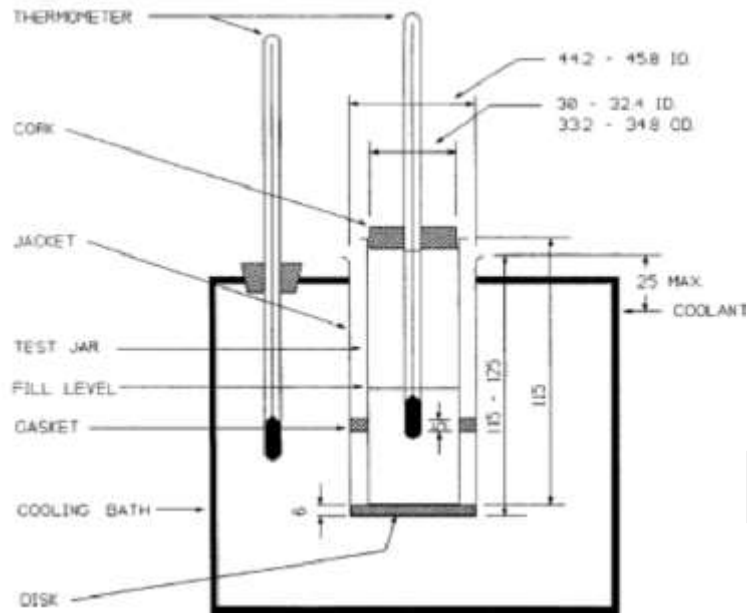
There are two primary viscosity parameters. These are:

1. Kinematic Viscosity
2. Dynamic viscosity

The injection spray will be very smooth and will be unable to penetrate completely away in the cylinder if the fuel viscosity is very low, which also causes loss of efficiency and power, and also causes corrosion [6].

### **1.5.7 Cloud and Pour Points**

Cloud and pour points are expressions which reference the relative coagulation of wax in diesel fuel; where the performance of diesel fuel in cold weather is an important consideration for users. The temperature at which the wax starts in the sediments of the fuel is considered as cloud point, and when the temperature lower than the cloud point, the longer paraffinic molecules sediment as a wax. This results in unheated fuel lines and filters due to this wax. Furthermore, most paraffin fuels have the highest cloud point. In some areas of this globe, pour point is an indicator of the lowest temperature at which fuel that will be pumped. The cloud point is usually between 4-5 °C which is higher than the pour point [6].



**Figure 1.4** Cloud and pour points apparatus [6]

### 1.5.8 Flash Point (Safety Factor)

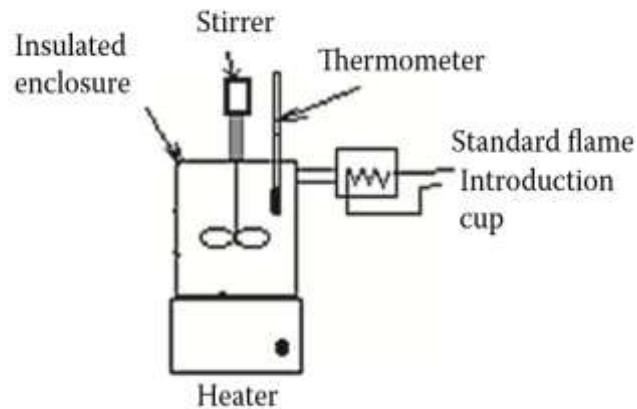
The temperature must be heated to create a flammable or ignitable mixture of vapour and air on the surface of the fuel liquid called the flash point. This feature is only necessary for safe fuel storage and handling. Therefore, fuel performance characteristics and emissions properties are not affected by it. Possible fire or explosion may happen if very low flash point is available [6].

### Flash Point Test Method (D93)

The flash point of finished and intermediate petroleum products can be determined using two methods. The first is Pensky Marten Closed Cup method (ASTM D93), and the second is Tag Closed Cup method (ASTM D56). ASTM D93 method is used for all fuel product distillates and distillation residues containing flash points above 148°F, while the ASTM D56 method is used only for products ranging from 148°F to 68 °F. On the basis of this principle, the ASTM D56 is utilized specifically for the kerosene. The Pensky Marten D93 will be described here only.

The test specimen is filled in a copper test cup up to the inside mark line. A cover on specimen dimensions is then fitted, as shown in Figure 1.5. The specimen is stirred and

heated at certain rates and then a small flame, the electric or gas ignition source, is directly supplied into the cup at equal intervals of each temperature ( $^{\circ}\text{C}$ ). The temperature of the sample is registered when the flash point is happened [6].



**Figure 1.5** Flash point apparatus [6]

### 1.5.9 Smoke Point

*“The smoke point of the test sample is quoted as the maximum height of flame that can be achieved without smoking”* [6]. The height of the flame (in millimetres) created by the oil in the wick of a stove or a lamp without producing any smoke is the smoke point. The burning quality is better if the smoke point is greater [7].

### 1.6 Specification of Diesel Fuel

All of the physical properties of automotive diesel fuel are described by ASTM. As shown in the Table 1.1 below, ASTM needs to involve the mixing of up to 7% from biodiesel (FAME) with classical diesel [12].

**Table 1.1** Specification of diesel fuel

Property	Unit	Lower limit	Upper limit
Cetane number	-	50.0	-
Cetane index	-	48	-
Density at 15 °C	kg/m <sup>3</sup>	820	870
Viscosity at 40 °C	mm <sup>2</sup> /S	1.60	5.50
Sulphur content	mg/kg	-	10.0
Flash point	°C	55	85
Carbon residue	%m/m	-	0.30
Ash content	%m/m	-	0.01
Water content	mg/kg	-	200
Total contamination	mg/kg	-	24
Fatty Acid Methyl Ether (FAME) (biodiesel) content	% v/v	-	7

### 1.7 Neural Networks

A neural network is an enormously distributed parallel processor comprised of simple processing units that have a natural tendency to save experimental knowledge and enable them to use it. It is similar to the brain in two ways [13].

1. Information is collected by the network from the environment at present during the learning process.
2. Synaptic weights, which are interneuron connection strengths, used to save accumulated information.

An artificial neural network (ANN) is normally talked about when a person mentions a neural network. ANN is a system that uses the biologically nervous system such as the brain as a basis to process the information. The structure of biological nervous systems is the key idea of this example, in which a huge number of interrelated processing elements come together to solve a problem. Neuron is the name given to such constituents. ANNs learn in the same way as human being. The process that is utilized to carry out the learning process is known as the learning algorithm. The synaptic weights of the network are adjusted by the learning algorithm to form a necessary design purpose [13].

ANN is regarded as a nonlinear statistical data modelling device. It processes information utilizing a connectionist advance to the computation and it is comprised of an interrelated

group of artificial neurons. In the oil and gas industry, the fame of the ANN is rising quickly. One of the most widely used forms of ANN is the multilayer feed forward network and it is able to model the unknown input/output relation of a multitude of complex systems.

### **1.7.1 Why Neural Networks?**

It has been in recent years, neural networks receive a large deal of attention. The primary cause for this is that, compared to the conventional approach, they facilitate a different approach to solve problems. The conventional approach is often implemented when a user knows the algorithm and teaches the computer to follow a group of instructions to solve the issue. The user needs to completely understand the problem and elucidates the steps that the computer has to follow to solve the problem or else the computer is unable to resolve the problem in this conventional approach. The conventional approach is thus limited due to these requirements. Therefore, the constraint of this approach is that the user has to already understand the problem and knows how to solve it [14].

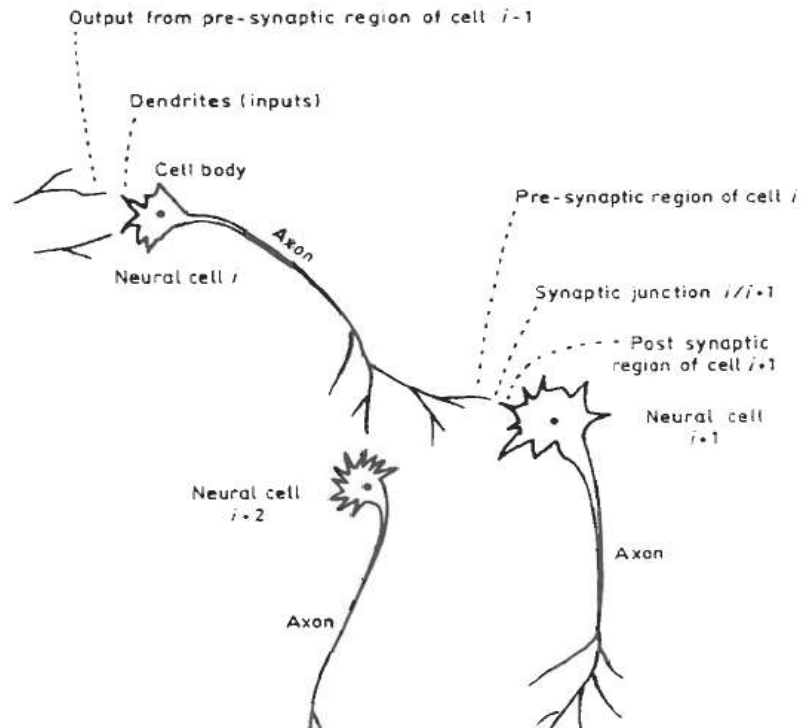
The neural networks are dissimilar of conventional approach in some cases. They are used when a person does not exactly know how to solve a problem. A neural network cannot be programmed to execute a certain task or to follow a group of instructions. Neural networks learn by example and data have an imperative role in them. There should be a careful selection of the first examples to ensure the network functions correctly. Another common issue is that there is not any way for the neural network to recognize a faulty system [14].

### **1.7.2 How Does Human Brain Learn?**

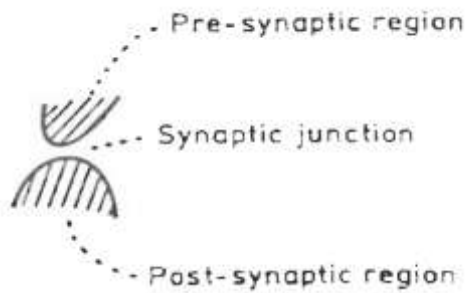
Since early times, people have been aware of the fact that the brain is a centre of learning. It has in fact been possible for neuroscientists during the last decade to move inside the brain and understand the learning process. The basic processing elements in the

brain are neurons who receive their feedback from other sources, merging the execution of a nonlinear process, at last distributing the result to other neurons present [14].

Nerve cells make up the biological neural network given in Figure 1.6, which are thus interconnected as in Figure 1.7.



**Figure 1.6** Interconnection of biological neural network [15]



**Figure 1.7** Synaptic junction [15]

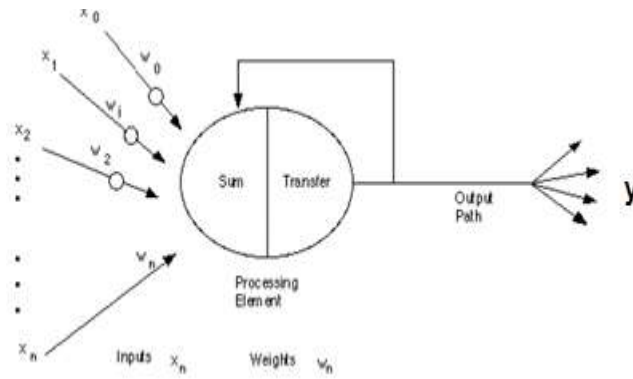
The majority of the neural computations occur in the cell body neuron are composed of the nucleus of neurons as well. Neural activity goes from one neuron to the other as electrical triggers. The triggers go from one cell to another under the neuron's axon, through an electrochemical procedure of voltage gated ion exchange along the axon and of propagation of neurotransmitter molecules across the membrane on top of the synaptic gap. Axon can also be seen as a wire connection. But, the system of signal flow is not through electrical conduction. This process of transportation transfers along the cell of the neuron, under the axon and then across synaptic junctions at the maximum axon across a very thin area to the dendrites.

As it is indicated in Figures 1.6 and 1.7 that since a neuron might have numerous synapses, a neuron is able to connect to a lot of other neurons. Likewise, a single neuron can receive messages from multiple other neurons since there are a lot of other dendrites per each neuron. In such a way, the biological neural network connects.

### **1.7.3 Artificial Neuron Model**

Biological neurons have some important functionalities. First a neuron receives inputs from other sources, then processes these inputs and finally outputs the result. Even though the biological neurons are significantly more complex than this structure, this primary structure is designed on the basis of neurobiological neurons and replicates their primary function. The structure of artificial neurons is shown in Figure 1.8. Inputs are represented by the symbol of  $x_n$  and are then multiplied with the corresponding weight  $w_n$ , represents the weight at the inputs,  $z$  is the output of the node and  $f_N$  is a nonlinear which will be mentioned later in chapter 4 Then a conversion function to the sum of products is applied that generates a result as an output [14], used to get neuron's output  $y$  as shown in Figure 1.8.  $y$  is a nonlinear operator that will be clarified later.

$$z = w_i * x_i$$



**Figure 1.8** Basic artificial neuron [14]

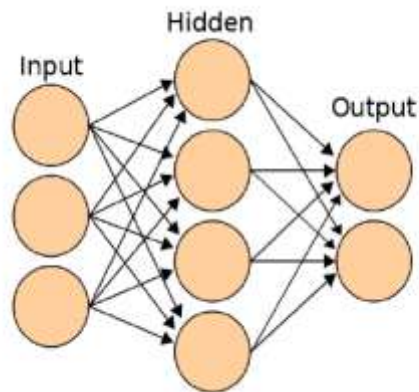
#### 1.7.4 Artificial Network Architecture

Similarly to a human brain, (artificial) neurons are the main processing elements in the artificial neural network. An ANN comprises of interconnected neurons arranged in the following layers.

- The input layer: The raw information for the network is represented through this layer.
- Hidden layer: One or more hidden layers to send the inputs into something that is able to be used by the output layers
- Output layer: The generated result is sent as output to the outside world [14].

##### 1.7.4.1 Feedforward Neural Network

The feedforward neural network starts with the input layer in the same way as the numerous other kinds of neural networks. The input layer can be connected to a hidden layer or directly to the output layer. If the input layer is connected to the hidden layer, it is easy to connect to another hidden layer or immediately after the output layer. As long as there is a minimum of a one hidden layer or an output layer provided, there could be any amount of hidden layers. Generality of neural networks have a one hidden layer mostly, it rarely happens that the neural network having two or more hidden layers. Figure 1.9 shows the feedforward neural network with one hidden layer.



**Figure 1.9** Feedforward neural network with one hidden layer [14]

#### 1.7.4.2 The Learning Process

The values of the connection weights between the neurons and the neuron thresholds are utilized to represent the knowledge of a neural network. The learning process in this is the determinant for these values [14].

The learning processes are categorized into two fields [16]:

- Supervised Training - is a method that gives a neural network its expected outputs. These desired outputs can then be produced by training the neural network.
- Unsupervised Training - is a method that the neural network is not provided with expected outputs.

In this study, the supervised learning process is used thus the unsupervised learning process will not be further discussed.

#### 1.7.4.3 The Backpropagation Algorithm

One of the most common supervised learning algorithms to train the network is the backpropagation algorithm. First, the weight matrix is initiated by some minute random values. The output is then found out by applying the input data to the network. Feedforward is the name given to this phase. In backpropagation, which is the next step, the error of each and every neuron is calculated in the backward direction. This means

these so called errors are first computed from the output layer and after that the hidden layer just before the output layer and in the same manner it is continued. Backpropagation thus gets its name from this process. The weight of the neurons is then altered by using these errors until the errors get smaller and the procedure is repeated until the error becomes negligible. In the case of supervised learning, a learning algorithm is used to attune the weights of the units in such a manner that error between the output of the model and supervised outputs is decreased. The neural network thus calculates the error derivative of the weights to fulfil this process. The weight of these units is increased or reduced marginally by the algorithm and is then recomputed to see how the error undergoes a change.

For supervised training, there are two values that needed to be considered in concluding the rate of error. Initially, one must compute the error for each element of the training set as it is progressed. Secondly, the aggregate of the errors for all of the elements of the training set along each sample needs to be computed.

The steps for backpropagation are as follows [13] [14]:

1. Start the weight's matrix with random values
2. Predict the output for the input example (feed forward)
3. Calculate errors at the output layer for output neurons
4. Calculate errors of the weight between the output and the hidden layer (backpropagation)
5. Calculate errors of the weights between input layer and hidden layer (backpropagation)
6. Revise the weights based on the calculated errors
7. Replicate steps 2 to 6 until the minimum error is reached (based on defined conditions).

## 1.8 Thesis Objectives and Problem Statement

The goal of this study is to predict the flash point of diesel from its composition to decrease the time and cost spent on experimental analysis of these properties. The experimental determination of the flash point of diesel is expensive, tiring and a time consuming procedure. The location of this study is in Erbil Gas Power Station (EGPS) 1500 MW situated in North Iraq, about 22 km from Erbil city centre. The kind of fuel used was natural gas as the first option and diesel as the second option. The diesel fuel is kept stored in 8 tanks. Each tank can hold up to 4000 m<sup>3</sup>. Every truck holding diesel needs to be monitored, especially the flash point. The artificial neural network (ANN) is considered as a means for predicting the flash point of fuel from the physical properties and the composition of diesel. And additionally, it is aimed to find a simple equation according to the fuel sample tests performed in the lab.

The initial research problem is how to predict the flash point by using a fuel analyser that facilitates our fuel composition. For instance, amount of aromatic rings, naphtha, density and smoke point, upper and lower heating value combinations are considered to find the best flash point prediction model utilizing the Levenberg-Marquardt back propagation algorithm.

The second research issue is the way to predict the flash point by using physical properties, such as, viscosity, density, sulphur content and distillation residue, to find the flash point using Levenberg-Marquardt back propagation algorithm and to find a formula that predicts flash point using multi-linear regression approach.

## CHAPTER 2

### BACKGROUND INFORMATION AND LITERATURE SURVEY

#### 2.1 Introduction

The flash point is the smallest temperature the vapour of fuel ignites or flashes when it is connected with a flame or spark [17]. It is very important to have suitable experimental flash point data and estimation methods for evaluation of the handling of combustible liquids for safety guidelines. The closed-cup method or the open-cup method can be utilized to acquire experimental flash point data. Table 2.1 lists the standard flash point testing methods for multiple types of samples that follow the American Society of Testing and Materials (ASTM) standards. In this study, Pensky Martens closed-cup tester D93 is utilised because an open cup tester allows the low boiling point component in the sample mixture to escape into the surrounding atmosphere before the flame is applied [17]. The sample in a closed cup tester is detached from the surrounding atmosphere, and it stops mass transfer between the surrounding atmosphere and the sample. Reliable and consistent results can be obtained using a closed-cup tester [18], but the open-cup method is recommended for heavy compounds with relatively high flash points, such as lubricating oil [19].

Generally, acquiring a flash point is experimentally preferable. It is still quite difficult to analyse the flash point of some radioactive and toxic compounds and it is in fact time consuming. The development of a faster estimation technique is needed to find out the flash point of an ignitable liquid.

**Table 2.1** The standard flash point testing methods for various types of samples that follow the American Society of Testing and Materials (ASTM) standards

Test method	ASTM	Flash point temperature	Scope
Cleveland open-cup tester	D92	between 79 °C and 400 °C	petroleum products (except fuel oils)
Tag open-cup apparatus	D1310	between 18.5 °C and 165.5 °C	liquids
Tag closed-cup tester	D56	flash points below 93 °C	liquids with viscosity at least 5.5 mm <sup>2</sup> /s
Pensky–Martens closed-cup tester	D93	40 °C and 360 °C diesel and biodiesels between 60 °C and 190 °C	Petroleum products
Modified continuously closed-cup tester	D7094	between 35 °C and 225 °C	Fuel, lubricants, solvents, and other liquids
Continuously Closed-cup tester	D3828	between 30 °C and 300 °C	Petroleum products and liquid fuel biodiesel
Small Scale Closed-Cup Tester	D3278	flash points between 0 °C and 110 °C	paints, enamel, varnishes, lacquers, and related products with viscosity is less than 151 St at 25.5 °C

## 2.2 Flash Point Prediction Model for Pure Compounds

In 1992, by K. Satyanarayana and Rao, the correlation and nonlinear exponential model were improved and they approximated the flash points of organic compounds and petroleum fractions as a function of boiling temperature. Over a big range of normal boiling temperatures, the model was tested on an estimated 1220 compounds from various chemical families, and a unique correlation was associated with each chemical family with less than 1% of the average absolute error [20].

Carroll et al., measured the flash points of 124 linear and branched acyclic alkenes and predicted them [21].

Laurent Catoire mentioned that the key determining factors of the flash point are the number of carbon atoms, a standard enthalpy of vaporization at 298.15 K and normal

boiling point. A case study approach used a general flash point prediction equation for organic compounds from various chemical families [22].

Models for the flash point of pure components were designed by Edna M. Valenzuela , and a model was tested on 611 chemical species from various families, in which both chemical and physical properties are comprised in the model [23].

Farhad Gharagheizi et al, introduced a study on the amount of carbon atoms and the normal boiling temperature. A sum of 77 chemical families was utilized with 1.471 pure compounds to ascertain a general relationship correlation. The experimental association, including the number of halogens, sulfur and oxygen atoms, was taken into account by them [24].

Many researchers have investigated the structural group contribution or molecular structure. Juan A. Lazzus studied seven thermal properties, melting point temperature, upper flammability temperature and boiling point temperature [25]. These properties were approximated utilize a hybrid method that involves an artificial neural network (ANN) with an optimal particle swarm [25]. Farhad Gharagheizi et al., studied and found out the greatest point of flash of chemical compounds by utilizing a chemical based structure model [26]. Mohamed Hussein Keshavarz worked on a basic way to anticipate the flash point of various classes of unsaturated hydrocarbons which include alkins [27]. The aromatics are estimated based on the amount of hydrogen and carbon atoms. It was found out that more trustworthy experimental flash point data is needed to form and train flash point prediction methods based on ANN to enable further improvements, specifically the ones that relate molecular structures with flash points [27].

### **2.3 Empirical Flash Point Prediction Based on Normal Boing Point and Composition**

The flash point of a fuel mixture is possible to be empirically correlated to the boiling point of the fuel mixture and the chemical composition range of the flammable component available in the liquid mixture [17].

## 2.4 Flash Point Prediction by Mathematical Regression Methods

It is relatively easy and simple to form empirical flash points from experimental data. The models are mostly formulated from artificial neural network (ANN) techniques and or by using mathematical regression techniques.

Many researchers have predicted flash point by mathematical regression summarized and compared below:

The Catoire model, Catoire and L.Paulmier estimated the flash point of pure compounds as well as of multicomponent mixtures, nevertheless, more experimental validation is required; therefore the flash point temperature is concluded through the following equation, including the amount of carbon atoms present in the fuel vapour mixture that is above the liquid phase ( $n$ ) and the normal boiling point of the fuel mixture ( $T_{eb}$ , K) [28], [29] .

$$T_{FP} = 1.477 \times T_{eb}^{0.79686} \times \Delta H_{vap}^{0.16845} \times n^{-0.5948}$$

The Wickey Model, Wickey: and Chittenden developed a formula to estimate the flash points of petroleum blends which are restricted to hydrocarbon mixtures that behave similar to ideal solutions. This model is not advised for non-ideal mixtures.

$$T_{FP} = \frac{2414}{618 + \log(I_{mix})} + 42.59$$

The formula is originated by being based on the index of the mixture,  $I_{mix}$ . Averaging the value of the flash point indices of the petroleum component with their volume fraction ( $v_i$ ) is used to acquire  $I_{mix}$ . The mathematical expression for this is  $I_{mix} = \sum I_i v_i$ , in which the flash point index,  $I_i$  is concluded centred on the flash point of the pure component  $I$ ,  $T_{FP}$  (expressed in the units of Kelvin) as shows by formula [17].

$$I_i = -6.1188 + 2414 / (T_{FP} + 503.71)$$

The Garland Model: Garland and Malcolm, established a new model which is used for mixtures of organic acid with anhydrides and water. The model was experimentally created by using statistical mixture modelling procedures. The weight percent (wt %) of

the respective component in the liquid mixture was correlated by this model to the mixture flash point. The  $R^2$  was stated to be a value of 0.946 for this model [17].

$$T_{FP} = 267.53 - 1.5927 (\text{acetic acid wt\%}) + 1.3897 (\text{propionic acid wt\%}) - 1.0934 (\text{butyric acid wt\%}) + 0.0027 (\text{acetic acid wt\%}) * (\text{butyric acid wt\%})$$

The Hristova model: In 2011 Mariana Hristova et al, created a simple data correlation for binary aqueous alcohol mixture, 3<sup>rd</sup> order polynomial correlations for  $T_{FP}$  is used to fit particular sets of experimental data.

$$T_{FP} = b_0 + b_1x + b_2x^2 + b_3x^3$$

The adjusted parameters ( $b_0, b_1, b_2, b_3$ ) of model were fitted using experimental data [30].

In 2010, Kim et al. estimated the results of the FP of liquid mixtures utilizing partial least squares method [31].

Mejia et al, in 2013, correlated the flash points of the diesel-biodiesel blends to the volume fraction of the biodiesel in the blend empirically. They concluded that these prediction models are only valid for particular blends that are used to develop correlations. However, the model is not suitable for the prediction outside the specified range of the flash point. Thus, it was correlated to the mix composition by experimental equations. The following correlations were suggested in that work.

$$T_{FP} = 343.03 - 48.41vp + 120.9vp^2 \text{ (diesel-palm oil biodiesel blend)}$$

$$T_{FP} = 350.28 + 0.0046e^{10.72vc} \text{ (diesel-castor oil biodiesel blend)}$$

$$T_{FP} = 430.00 - 294.14vc + 167.86vc^2 \text{ (castor oil biodiesel blend, palm oil biodiesel)}$$

Where,  $vc$  is the volume fraction of castor oil biodiesel and  $vp$  is the volume fraction of biodiesel in the blend [32].

## 2.5 Flash Point Prediction Using the ANN Methods

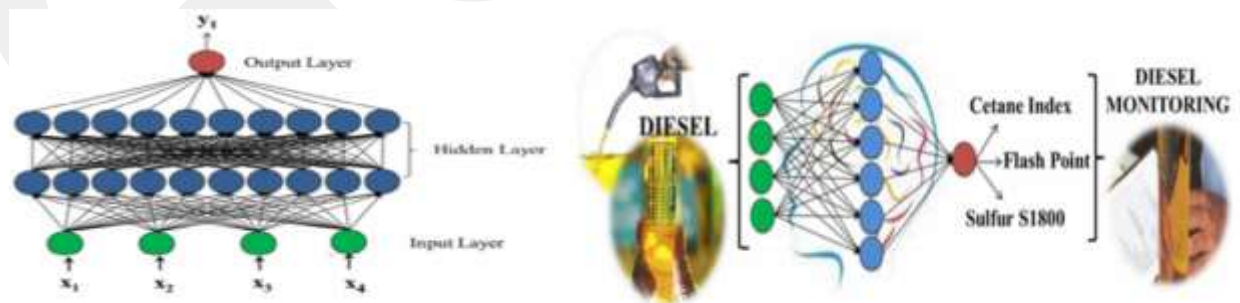
ANN methods are strong modelling tools to establish diesel properties and they have the ability to recognize very complex relationships or obtain knowledge using data input and output only [33] [34].

In 2002, Korres and co-authors determined the lubricity of the fuel by a neural network with radial basis function having the specific gravity, conductivity, sulfur content, viscosity, and 90% distillation point as input database [33].

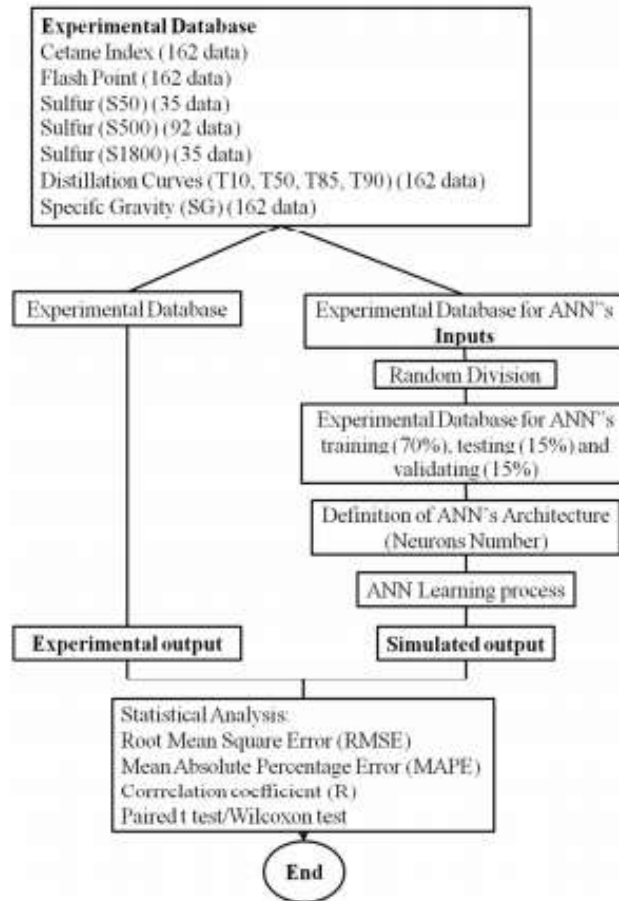
In 2002 Basu et al., predicted the cetane amount of diesel samples with a backpropagation algorithm by using spectroscopic device data as input and by using an ANN [35].

In 2017 Fernanda M. de Oliveira et al. predicted the flash point, sulfur content and cetane number. They used ANNs of diesel blends as 7% biodiesel using the parameter such as distillation curves (ASTM D405), specific gravity at 20 °C (ASTM D405), sulphur content (ASTM D4924), flash point (ASTM D93) and cetane index (ASTM D4737) [36]. A case study approach was used for 162 samples, which were randomly divided into three groups; 15% for testing, 15% for validation and 70% for the samples of training. The aim was to enhance the generalized capability of the models. Therefore a cross validation process was carried out during training of the network such that the amount of iterations in this step was conditioned by square error in the test samples [36].

Feasibility was shown by ANN in the suggested method to accurately predict cetane index, sulphur content and flash point, with an aggregate absolute percentage error of 4.6%, 0.4% and 3.3 % correspondingly. These models were proved to be quite successful as they used the input matrix, data of physicochemical properties such as distillation curves, sulfur content specific gravity at 20 °C, cetane index and flash point that were already a component of the diesel specification routine. As a result, few of the physicochemical was used. Figures (2.1, 2.2) are the description of methodology of ANN [36].



**Figure 2.1** Architecture of the MLP network used to predict physical property parameters of diesel [36].



**Figure 2.2** Flow chart for the training methodology of artificial neural networks [36].

In 2010, Jatinder Kumar and Ajay Bansal estimated the flash point, density, viscosity and fire point of FAME and diesel mixtures [37]. A study, which was made to estimate these properties without experiment can have significant advantages. An intriguing finding is that the ANN is the best choice for this system. Thus, the characteristics have a very small error in the case of artificial neural network and that gives a better approximation of these characteristics compared with appropriate statistical technique of the fitting the curve (the principle least squares). The Table 2.2 that follows demonstrates the mean square error after generalization [37].

**Table 2.2** Mean square error after generalization [37]

Property	Mean Square Error	
	(least square )	ANN
Flash point	5.74	0.16
Fire point	4.11	0.74
Viscosity	0.06	0.02
Density	$5.30 \times 10^{-6}$	$5.54 \times 10^{-6}$

Agarwal et al. [38] .Compared to linear regression techniques and ANN to predict the properties of FAME. The finding of this study showed that ANNs were able to predict the properties of FAME much better than a linear regression. The fatty acid composition of vegetable oil was the input data. The most correct predictions of biodiesel flash point were acquired using a single hidden layer ANNs. The ANN model was made using MATLAB software for anticipating properties of biodiesel such as pour point, cloud point, viscosity, density, iodine value, saponification value, fire point and flash point as a function of fatty acid content. The fuel properties of biodiesel were found out according to ASTM methods using basic apparatus for the measurement. The Pensky Marten apparatus was used to measure the flash point and the fire point. ASTM methods were used to measure cloud point and pour point by using a pour point measuring apparatus. According to the ASTM methods prescribed, saponification and iodine values were measured through titration. As a whole, this study reinforces that for the flash point and fire point. Linear regression gives weak results for the oils utilized. On the contrary, properties were predicted better by ANN as compared to linear regression for a given set of data [38].

In 2007, Guzu Liu et al. investigated the relationship of the properties of the chemical composition of the eight fuel related to chemical composition and many characteristics specifications including freezing point, density, aniline point, flash point, net heat of combustion. The chemical composition of the jet fuel was concluded via composition property relations. Moreover, it was ordered eight categories of hydrocarbons; these include isoparaffins, tetralins, normal paraffins, dicyclopraffins, monocyclopraffins, alkylbenzens, hydroaromatics, and naphthalenes [39]. The ANN new approach on chemical composition property relations is a reliable and trustworthy tool for the

formation of next generation jet fuel and the fast quality control technique with the highest promise [39].

In 2015, Alex Oliveira and co-workers, reached the decision of the adequate performance of the model, such that they were able to anticipate the quality parameters of biodiesel utilized by ANN [40].

A paper was published by Gatinder Kumar in 2007, in which a comparison of statistical and ANN data was used to predict the physical properties of a mixture of distinct diesel and biodiesel. The author stated that the results showed that ANN is the best option for this particular system. Viscosity, fire point, flash point and density were the tests carried out in this situation [41].

In 2002, Yang and co-workers used a general regression neural network to predict the cetane number and specific gravity of diesel on the basis of its chemical composition [42].

## CHAPTER 3

### METHODOLOGY

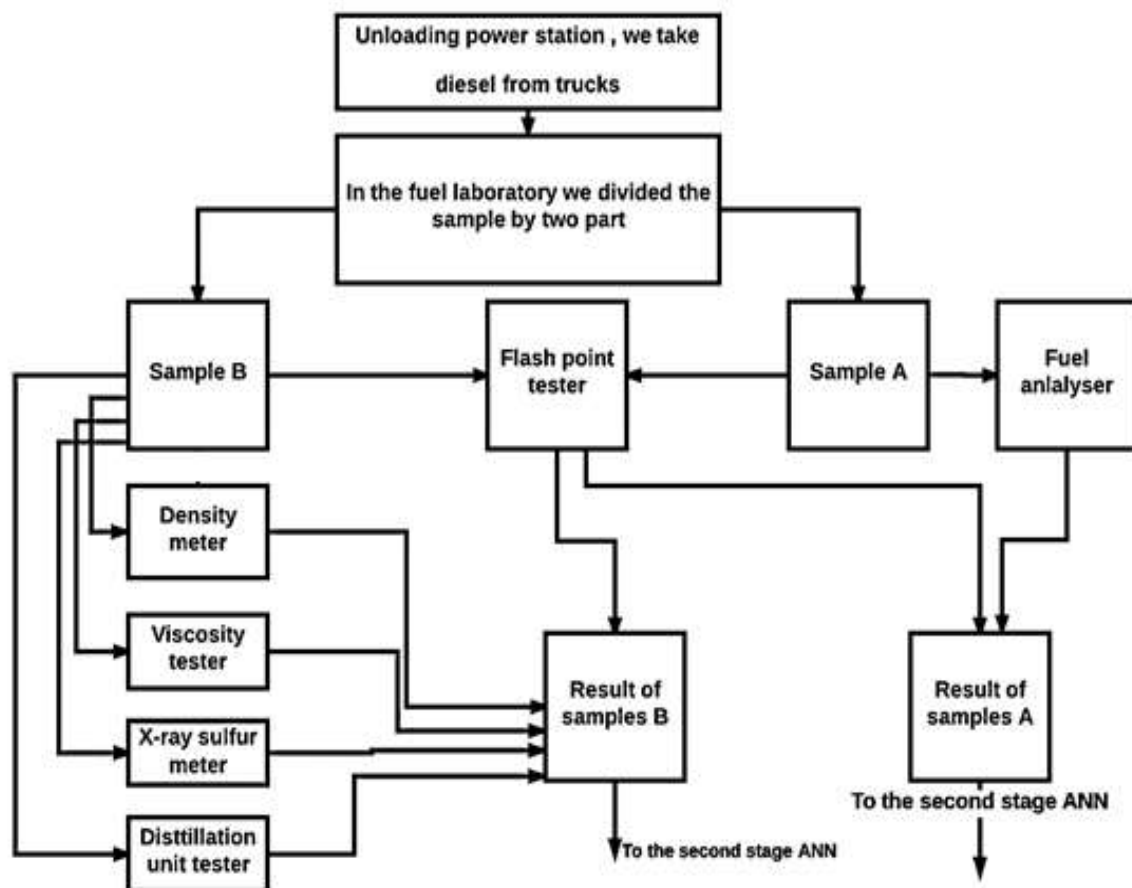
#### 3.1 Introduction

This chapter was divided into two stages. The first one involves working in a laboratory to measure the composition and physical properties of diesel fuel, and the second stage introduced the ANN that was used for prediction of flash point. The flow chart in Figure 3.1 reveals the stages of the methodology used. The figure explains the procedure of sampling method, where the sample is taken from the truck that brings fuel to the power station and then divided into 2 parts, namely (A) and (B). The case study approach is used by many fuel devices to get more information of the results.

Many researchers have utilized the chemical composition of fuel to measure the flash point. For this reason, the first sample (A) is used in a fuel analysis device that gives an in-depth a report about the chemical composition of diesel in a short period of time compared to the flash point

Furthermore, physical properties of fuel are employed by a number of researchers to predict the flash point. For this reason, in the second sample (B) the following tests, sulfur content, density, viscosity and distillation are used.

After collecting the samples, ANN model is used to predict flash point through the network training using Matlab 2017b software. Figure 3.13 shows the model method ANN.



**Figure 3.1** Flow chart of the methodology used

## **3.2 Experimental Work (Stage One)**

### **3.2.1 Experiment on Sample (A)**

To determine the flash point and other analysis, diesel fuel samples were collected from trucks that come to the Erbil Gas Power Station (EGPS). 137 of diesel samples (approximately 100 ml per sample) were stored in sealed bottles at room temperature until using them in flash point tests and fuel analysis

#### **3.2.1.1 Flash Point HFP 339 According to ASTM D93**

Automatic flash point testers were used a large data set of measured flash points experimentally in the Series HFP 339 apparatus. The determination of the flash point demonstrates an empirical method. As a result it is vital, with respect to the unit and the procedure, to follow the standard specifications as accurately as possible. The flash point of a material is employed in transporting and in safety requirements for the definition of combustible and flammable substances. The flash point is one of a number of material properties which is used for the determination of the overall flammability hazard of a fuel. The flash point determination of the sample is carried out by charging it into a test crucible up to a specific mark. The sample is heated under a steady heating rate under constant stirring. An electrical igniter or a small gas flame is put into the crucible at steady intervals, and the stirring is interrupted at the same instant. The flash point is regarded as the least temperature of the liquid at which the gas mixture above the sample is ignited by an ember. The measurement time turns out to be roughly 30 to 60 minutes, which depends on the flash point of the material. The demands on the laboratory assistant to achieve the measurement are high during a manual measurement, especially in handling of the sample (stoppage of the loss of volatile materials, accurate observation of the flash point) along with the final cleaning of the unit.



**Figure 3.2** Front view of the HFP339



**Figure 3.3** Unit with gas ignition

### Steps of the experimental work with the HFP 339 apparatus

- a) Pour a diesel sample 20 ml into the standard oil cup to the marked line, Figure 3.2 show the front view apparatus.
- b) Place the cup into the apparatus and cover the lid. By connecting the electrical ignition in the electric igniter, the black and red connections are used
- c) FP sensor. The flash point is obtained thermally via a FP sensor (3-prop plug on the right)
- d) The thermometer is connected. The HFP339 is installed with a resistance thermometer and the temperature of the charged sample is measured through this.
- e) The stirrer cable is connected. Connect the flexible cable to the stirrer coupling on the housing on the crucible. Right after starting the test, the stirrer engine starts to turn the flexible cable around. Ascertain that the square connections are completely engaged into the opposite member, Figure 3.3 shows gas ignition unit and explain FP sensor, thermometer, and stirrer cable.
- f) Choose the procedure ASTM D93 and enter expected flash point.
- g) Click TEST the apparatus starts heating at the rate  $5.8\text{ }^{\circ}\text{C}/\text{min}$  and 105 rpm stirring rate.
- h) When it gets to a temperature of  $23\text{ }^{\circ}\text{C}$  before the expected flash point, the apparatus begins to ignite automatically. The apparatus will ignite once automatically for every  $1\text{ }^{\circ}\text{C}$  the increase in the sample.
- i) The apparatus will alarm automatically when the flash point appears.
- j) Click PRINT to print the result of the test. The format of the report after the end of the experiment is demonstrated in Figure 3.4.

HERZOG	HFP339 FLASH POINT
RESULT	
Sample number:	Test 004 Diesel
Measur.program (3):	ASTM D93-A
Time:	09:45
Expected flash point:	60 C
Measured flash point:	61 C
Airpressure:	985 mbar
Corrected flash point:	61.5 C

**Figure 3.4** Flash point test result

### 3.1.1.2 Fuel Analyzer (Cetane 2000)

The sample approximately uses 10 ml diesel fuel to be used in the second step examination. Cetane 2000 is the initial analytical tool which involves analyses for, ARO, PNA, ARO-JET, naphthalene (total aromatic ASTM D5186), smoke point (ASTM D1322), net heat of combustion and cross heating combustion according to (ASTM 4868), density (ASTM D405) and cetane number (ASTM 613) , Figure 3.5 shows the Cetane 2000 apparatus.

Steps of the experimental work

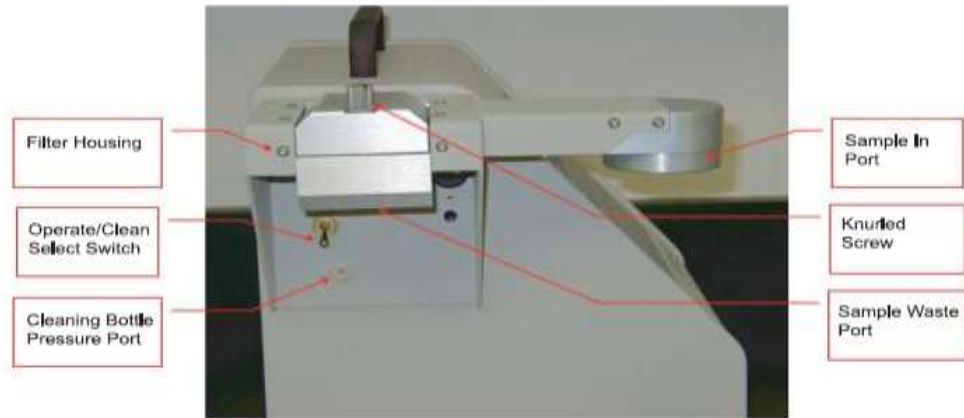
1. A minimum 75% of the sample bottle should be filled with the diesel sample to obtain the best results. A suitable dip tube should be employed for the selected sample bottle. Then, the sample bottle is screwed tightly on the sample input port. However, it is vital to avoid over tightening the sample bottle to protect it from breaking. (Connect the amber waste bottle to the sample waste port and do not over tighten the bottle as it can cause bottle breaking), Figure 3.6 shows fixed arm features , Figure 3.7 show the standard sample run setup .
2. The PUMP function enables the user to determine the length of the time that the fuel pump works before starting the analysis process. At least 3 ml of fuel should

be pumped by the fuel analyser through the system within a chosen pump time. Figure 3.8 shows the sample preparation .

3. The PRINT set up function lays down the format of printed analysis results. A user is able to choose either TABLE or REPORT format. The table format lists the sample identification code, the time and the date on which the analysis was carried out, and all the analysis results. The report format prints the same data but allows room for a description of the sample, results from carrying out a calibration standard, and a location for the operator's signature.



**Figure 3.5** Cetane 2000 apparatus



**Figure 3.6** Fixed arm features



**Figure 3.7** Standard sample run set-up



**Figure 3.8** Sample preparation

The Appendix A illustrates the results of each sample (A) so we can start working on the second stage.

### **3.2.2 Experiments on Sample (B)**

To determine the flash point and other analysis, diesel fuel samples were collected from trucks that come to the Erbil Gas Power Station (EGPS), and 192 diesel samples (approximately 200 ml per sample) were stored in sealed bottles at room temperature until using them in flash point test and other tester.

#### **3.2.2.1 Flash point HFP 339 using ASTM D 93**

The same methodology used in section 3.2.2.1 is also used for sample (B).

#### **3.2.2.2 Automatic Kinematic Viscosity Measurement System**

This system is used to determine the automatic viscosity and kinetic. The report is fully automated when the sample is set by a series of operations; these operations are composing of timing, cleaning and drying the viscosity tube. Figure 3.9 demonstrates this instrument.

#### **Steps of the experimental work for automatic kinematic viscosity measurement**

- a) According the ASTM D445 the automatic kinematic viscosity will be operated.
- b) Turning the instrument on will give the details of both the viscometer as well as the pre-set and actual bath temperature.
- c) Then charge sample lift and turn on the solvent dispensing cap, which will be indicated on the screen.
- d) Approximately 15 ml of diesel fuel is poured into the viscometer tube.
- e) There is an inbuilt over flow tube in the viscometer along with the sample reservoir tube.
- f) Press the start button on the control panel. The machine then automatically goes through a series of pre-test checks such as soaking and adjustment of optical sensors.

- g) The suction of the sample is begun automatically after this and the sample begins to flow through the tube into the testing bulbs.
- h) Immediately when the sample of the fuel goes through the 1<sup>st</sup> optical sensor located at the beginning of the lower bulb, the timer starts and it then stops when it goes through the 2<sup>nd</sup> optical sensor.
- i) Identical measurements of time are carried out when the sample goes through the 2<sup>nd</sup> bulb as well, and as soon as the sample passes the top bulb of the viscometer, the timer stops and then calculates the test result.
- j) When the testing is finished, the remaining sample is sucked out of the viscometer tube, and the cleaning ritual is started. Then the solvents used for cleaning are dispensed through the special solvent dispensing cap as per the required sequence. Then the viscometers are dried via the vacuum and the machine is ready for the coming test.



**Figure 3.9** Viscosity tester

### 3.2.2.3 Density Model (DMA™)

DMA™ 35 Portable Density Meter is used to measure density and concentration sample and it is compatible with ASTM D405. The sample was put into the measuring

cell of DMA™ 35 by only pushing the pump lever. DMA™ 35 measures the density of sample, or decides concentration using the initially installed density concentration tables. It is very useful to measure specific samples frequently. Approximately 2 ml, which is a small sample of the volume, is absorbed into the device and after a small amount of seconds, the result is visible and we can see the result and the temperature at the same time. Figure 3.10 demonstrates this portable density meter DMA™ 35.



**Figure 3.10** Density instrument

### **The steps of the experimental work for density measurement**

- a. Press down the pump lever as far as it can go
- b. Submerge the filling tube in the sample.
- c. Slowly release the lever of the pump.
- d. Use the table of temperature – density to usual methods, convert the density from the test temperature to the temperature at 16 °C.

#### **3.2.2.4 X-RAY Sulfur Meter**

X-ray sulfur meter is used to determine total sulfur content in petroleum products, such as diesel, using X-ray fluorescence. The device is accurate, non-destructive and compatible with ASTM D4294.

#### **Specifications of the device X-ray Sulfur content**

1. Precise X-ray fluorescence analyser using a compacted X-ray tube
2. Quick analysis of sulfur content in the oils by the non-destructive test
3. Great span of measurement from 20 ppm to 60 wt%
4. Conforms to the ASTM D4294
5. Complete unit with a display, operation keys and a built in operator
6. Small, portable and light weight (12kg)
7. Very uncomplicated operation panel
8. The sample's cell is made up of PTFE, which can be used multiple times.

#### **Steps of the experimental work for sulfur content measurement**

- a. Place a cell with a sample that was made on a measuring position, to prepare the cell using Teflon sample and Mylar film. Then put the diesel around 2-3 ml inside the Mylar film (Mylar film is used only once), Figure 3.11 shows method of sample preparation.
- b. Shut the cover of the measuring compartment.
- c. The start key is pressed to begin the measurement.

- d. A RS-232C port can be used to display, print out and get the output measurement results.
- e. A multi-point calibration or a one or two point simplified calibration is available by using a standard reagent.



**Figure 3.11** X-ray sulfur instrument

### 3.2.2.5 Automatic Distillation

It is applicable for ASTM D 86 and for finding the end point temperature and the amount of residue weight. This work is demonstrated in Figure 3.12.

#### Steps of the experimental work for automatic distillation

1. The cylinder should be filled up to 100 ml
2. The 100 ml of sample is transferred into a 125 ml flask
3. Place some pumice stone in the sample for a much better homogeneity during the boiling process.
4. Plug the stopper on the flask

5. Apply force downward on the top door such as to lower the plate, bring in the flask tube in the condenser stopper and lock the metal arm.
6. The cylinder is then put in the chamber
7. The instrument is operated and the temperature is monitored. Then record the end temperature and also calculate the residue temperature.



**Figure 3.12** Distillation unit tester

The results for each sample (B) are given in Appendix B. So we can start working on the second stage.

### **3.3 ANN Modelling (Stage Two)**

#### **3.3.1 Introduction**

An analysis of the data makes it possible to interpret the relation between the variables and recognize the input variables of the model. Thus a proper model can be formed by using this knowledge and predict the outcomes that this model will give us. The model is designed on the basis of analysed data in this section. At the beginning the primary specification of the model is studied. Various methods made for trial. These methods are dissimilar in their details such as hidden nodes, inputs and outputs of the model [14].

In developing an artificial Neural Network for any application, The following steps must be followed; collect data, pre-process the data, determine the structure of the network (number of hidden layers and nodes), create the network, initialize weights and biases, train network, validate network and test network [14]

### 3.3.2 The Basic Specification of the Model (Stage two)

The model to be used a multilayer perceptron MLP.

#### 3.3.2.1 Input Layer

It is selected from the nine possible inputs (sample A) of the model, Table 3.1 below shows the input layer for sample A

**Table 3.1** Input layer for sample (A)

No.	Input sample A
1.	Cetane number
2.	Density
3.	Cross heating combustion
4.	Net heating combustion
5.	PNA
6.	ARO
7.	Naphthalene
8.	Smoke point
9.	ARO-JET

It is selected from the five possible inputs (sample B) of the model, Table 3.2 below shows the input layer for sample B.

**Table 3.2** Input layer for sample (B)

No.	Input sample B
1.	Sulfur content
2.	Density
3.	Viscosity
4.	Distillation end point
5.	Distillation residue

### **3.3.2.2 Output Layer**

The flash point prediction is the aim of this study.

### **3.3.2.3 Hidden Layer**

There are two questions that need to be answered when the layer is being designed. Initially, how many hidden layers are required in one neural network?

Secondly, how many neurons are required in every layer?

For the answer of the first question, it is vital to know that most neural networks have only one hidden layer, and the use of 2 or more hidden layers is very rare. In actual fact, numerous practical problems can be solved by using one hidden layer only.

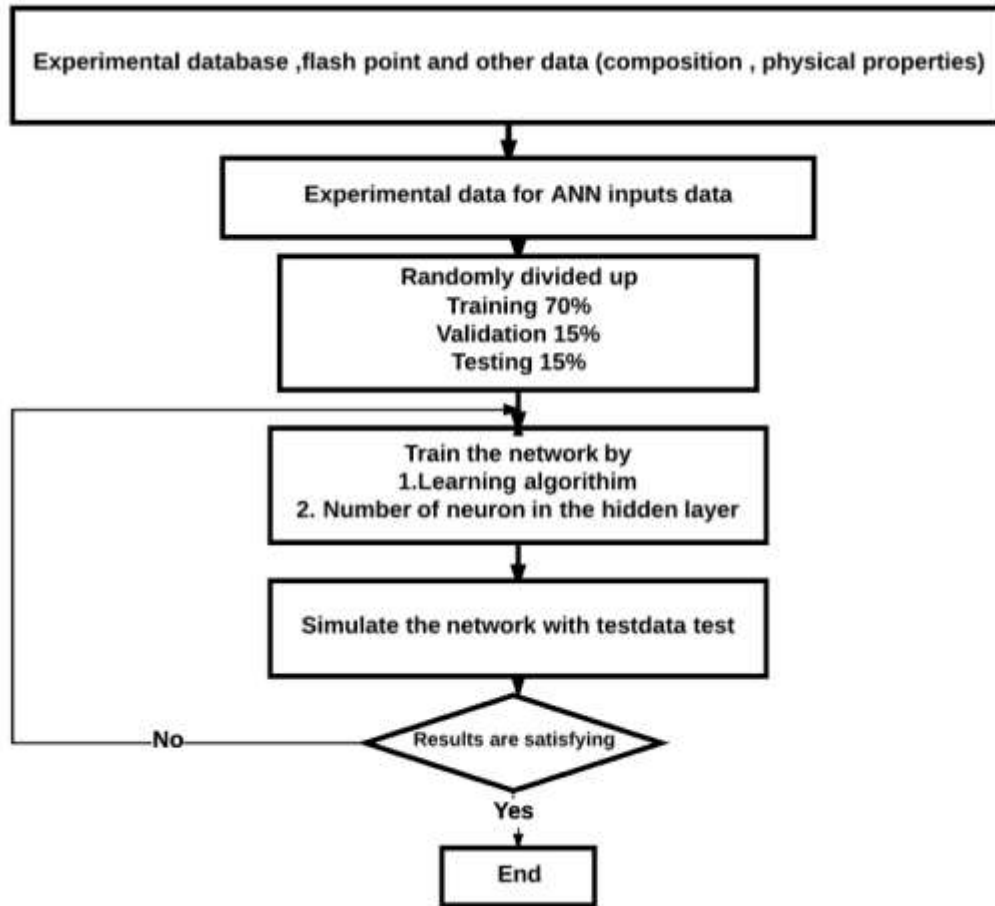
On the basis of what we described, for the number of hidden layers we have decided to choose a single layer.

Choosing the hidden layer number is quite an easy decision due to the fact that many problems can be solved only using a single layer. Finding the number of neurons in this layer is the chief question to be addressed while the neural network is being designed.

There are numerous thumb rules to find the amount of nodes in hidden layers, and the most common rule is that the amount of nodes in the hidden layer is the average of the output node and input nodes.

To find the correct number of neurons in the hidden layer, these rules are used as a starting point and trial and error approach is utilized to find the number of nodes.

Description of this methodology is given by the flow chart in Figure 3.13 (sample A and sample B have the same methodology).



**Figure 3.13** Flow chart of the second stage ANN model

### 3.3.3 Model Design for Sample (A)

The following steps were taken in training the network using Matlab R 2017b software to model our network. Matlab has the flexibility to select inputs, outputs and hidden neurons. There is the possibility to easily change the hidden neurons and test the result. On the other hand there are some different methods to divide the dataset such as block or random. The percentage of training, validation and test subsets can be changed easily. There are different tools like performance plot, regression plot and histogram to analyse network performance.

Prediction is a type of dynamic filtering, where the previous values of one or a number of time series are utilized to predict future values. In this study, prediction producer can be summarized as follows:

1. In this model, the thumb rules were applied to determine the number of nodes in the hidden layer. The most common rule for the number of nodes in the hidden layer is the average input and output.
2. The best method of prediction is chosen. The solution of NARX was applied similar to the solution that was addressed by Asadolahi Reza (2014) [14].

**Table 3.3** The best prediction for sample (A)

Method	Nonlinear input – output		NN		NARX	
	MSE	R	MSE	R	MSE	R
Training	2.4	0.782	1.8	0.912	<b>2.9*10<sup>-3</sup></b>	<b>0.999</b>
validation	18.2	6.5	6.2	0.867	<b>2.73*10<sup>-1</sup></b>	<b>0.985</b>
testing	10.8	1.4	6.6	0.427	<b>6.3*10<sup>-1</sup></b>	<b>0.983</b>

In the first kind of time series problem, the future values of a time series  $y(t)$  should be predicted from past values of that time series and past values of a second time series  $x(t)$ . This form of prediction is called NARX, the formula below.

$$y(t) = f(y(t-1), \dots, y(t-d), x(t-1), \dots, (t-d))$$

3. Getting data from workspace input time series  $x(t)$ , and target (FP) time serie , and defining the desired output  $y(t)$ .
4. Selecting the time series format matrix that means Input “compositions” are a 137x9 matrix, and inputs represent dynamic data: 137 time steps of 9 element and Targets 'flash point ' is a 137x1 matrix, and targets represents dynamic data: 137 time steps of 1 element.
5. Randomly dividing up 137 target timesteps to three kinds of target timesteps as:
  - a) Training, the targets that are displayed on the network during training, its error is modified, and the data are divided by up to 70% which means 95 target timesteps.

- b) Validating, the targets that are utilized to measure network generalization. Training should be discontinued when generalization stops improving, and the data are divided by up to 15% which means 21 target timesteps.
- c) Testing, the target which has no effect on training and providing an independent of measuring training network performance which mean the training during and after; the data are divided by up to 15% which means 21 target timesteps.

Levenberg -Marquardt backpropagation is used for network training. The training discontinues automatically when the generalization process stops to improve and is explained by increasing the root mean square error that is used to validate the samples.

The network is created and trained to form an open loop as shown below. Generally, training is more efficient of an open loop (single step) comparing with a closed loop (multi step).The open loop allows us to provide the network using the correct previous outputs and neural network training to predict output.

The function in hidden layer is known as Sigmoid Activation Function. It uses the function to determine activation. The equation of sigmoid function can be given as [16]:

$$f(x) = 1/(1 + e^{-x})$$

The function in the output layer is known as Linear Activation Function. It may be helpful in the case when you need a whole set of numbers to be output. Generally, you will need to think about your neurons as non-active or active; the linear function is defined as follows [16] :

$$f(x) = x$$

After training, the open loop network may turn into a closed loop formation or any other form required when appropriate.

6. The MSE will be obtained. MSE is the average squared difference between outputs and targets, zero means no error, lower values are better.

R values measure the relationship between outputs and targets that means a close relationship.

### 3.3.4 Model Design for Sample (B)

The flash point and fuel physical properties of all selected feeds were used as input and target parameter, the input (5) and target (1).

The following steps were applied in training the network using Matlab R 2017b software.

1. In this model, the thumb rules were applied to determine the number of nodes in the hidden layer. The most common rule for the number of nodes in the hidden layer is taking the average of input and output.
2. Choosing the best method of prediction. The solution of a Neural Network NN is applied [14].

**Table 3.4** The best prediction for sample (B)

	Nonlinear input – output		NN		NARX	
	MSE	R	MSE	R	MSE	R
<b>Training</b>	8.108	0.9206	<b>0.549</b>	<b>0.9942</b>	6.621	0.9344
<b>Validation</b>	11.287	0.8887	<b>0.586</b>	<b>0.9944</b>	32.822	0.7639
<b>Testing</b>	27.756	0.7361	<b>0.721</b>	<b>0.9958</b>	6.87	0.9392

3. Getting the data and the target time series (FP) from workspace input time series  $x(t)$  defining the desired out  $y(t)$ , ( five parameters are inputted such as density, viscosity, sulfur content, distillation end point for distillation residue %).
4. Selecting the time series format matrix means Inputs 'properties' are a 192x5 matrix, and inputs represent dynamic data: 192 time steps of 5 element and Targets 'flash point ' is a 192x1 matrix, and targets represents dynamic data: 192 time steps of 1 element.

5. Dividing up 192 target timesteps randomly to three kinds of target timesteps.
  - a) Training: the data are divided by up to 70% which means 134 target timesteps.
  - b) Validating: the data are divided by up to 15% which means 29 target timesteps.
  - c) Testing: the data are divided by up to 15% which means 29 target timesteps.

Levenberg-Marquardt backpropagation is used for network training. the training discontinues automatically when the generalization process stops to improve and is explained by increasing the mean square error that is used to validate the samples.

6. The MSE will be obtained which is the average squared difference between outputs and targets, zero means no error, lower values are better.  
R values measure the relationship between outputs and targets that means a close relationship.

## CHAPTER 4

### RESULTS AND DISCUSSION

In this chapter, the results for samples (A) and (B) are discussed.

#### 4.1 Sample (A)

##### 4.1.1 Hidden Nodes

The previous chapter was devoted to explain how to calculate hidden nodes and how to apply the thumb rules to determine the number of nodes in hidden layer; where the most popular rule is that the number of nodes in the hidden layer is the average of the input and output [14].

It was started from one node lower down the average input and output.

$$\text{Mean} = (9 + 1) / 2 = 5.$$

$$\text{Start point} = 4 \text{ nodes.}$$

For each node, we trained three-times and recorded the regression between the target and output for training, validation and testing results. Finally, the averages were calculated , as suggested in the literature [14].

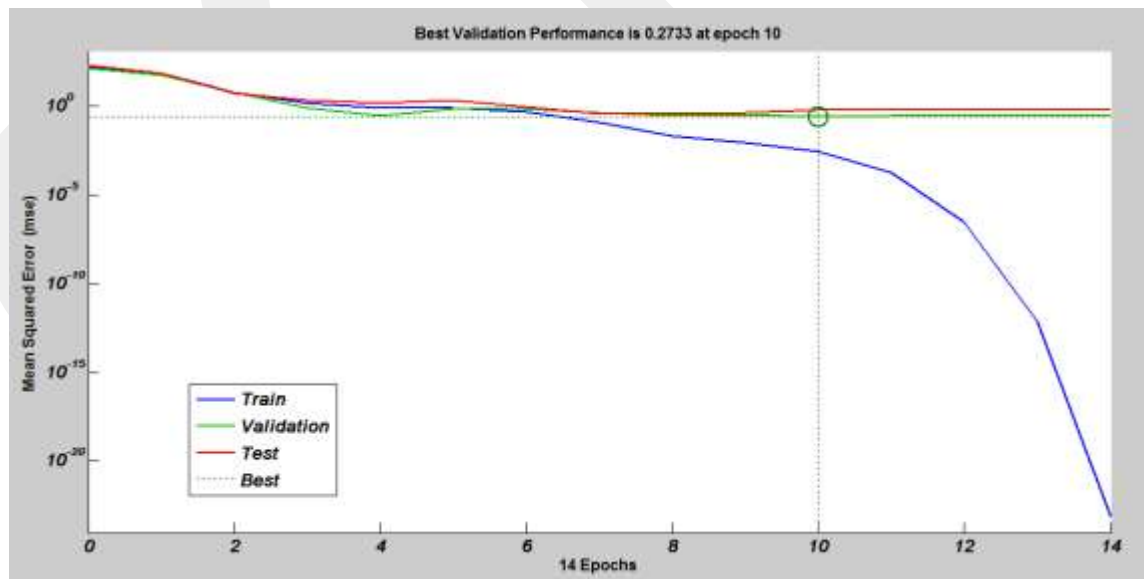
**Table 4.1** Hidden nodes regression results of sample (A)

Node	Training	Validation	Testing	All
4	0.995	0.982	0.994	0.9890
5	0.993	0.991	0.960	0.9915
6	0.986	0.983	0.989	0.9919
7	0.997	0.970	0.960	0.9860
8	0.994	0.976	0.963	0.9840
9	0.986	0.957	0.898	0.9541
<b>10</b>	<b>0.999</b>	<b>0.985</b>	<b>0.983</b>	<b>0.9933</b>
11	0.994	0.954	0.884	0.9844
12	0.997	0.985	0.985	0.9899

Table 4.1 also shows the result of 4 to 12 nodes. The results are still reasonable but 10 nodes have a regression coefficient over 0.993341. Because of the study has a big data set; it is preferred to choose a hidden number 10.

#### 4.1.2 The Network Performance for Sample (A)

After finding the hidden nodes, the evaluation of the network performance was achieved.



**Figure 4.1** Network performance of sample (A)

The MSE is a squared difference between output and target. This means that a better prediction can be reached to when MSE value is low and the zero value of MSE means the training performance without error. Figure 4.1 shows the performance plot of the training. Note that the performance of the training reached to the minimum value at the iteration 10 and the training continued to the iteration 14 then halts. Also this figure does not show any major problems during training, and the best validation performance is 0.2733 at epoch 10

### 4.1.3 The Regression Plots for Sample (A)

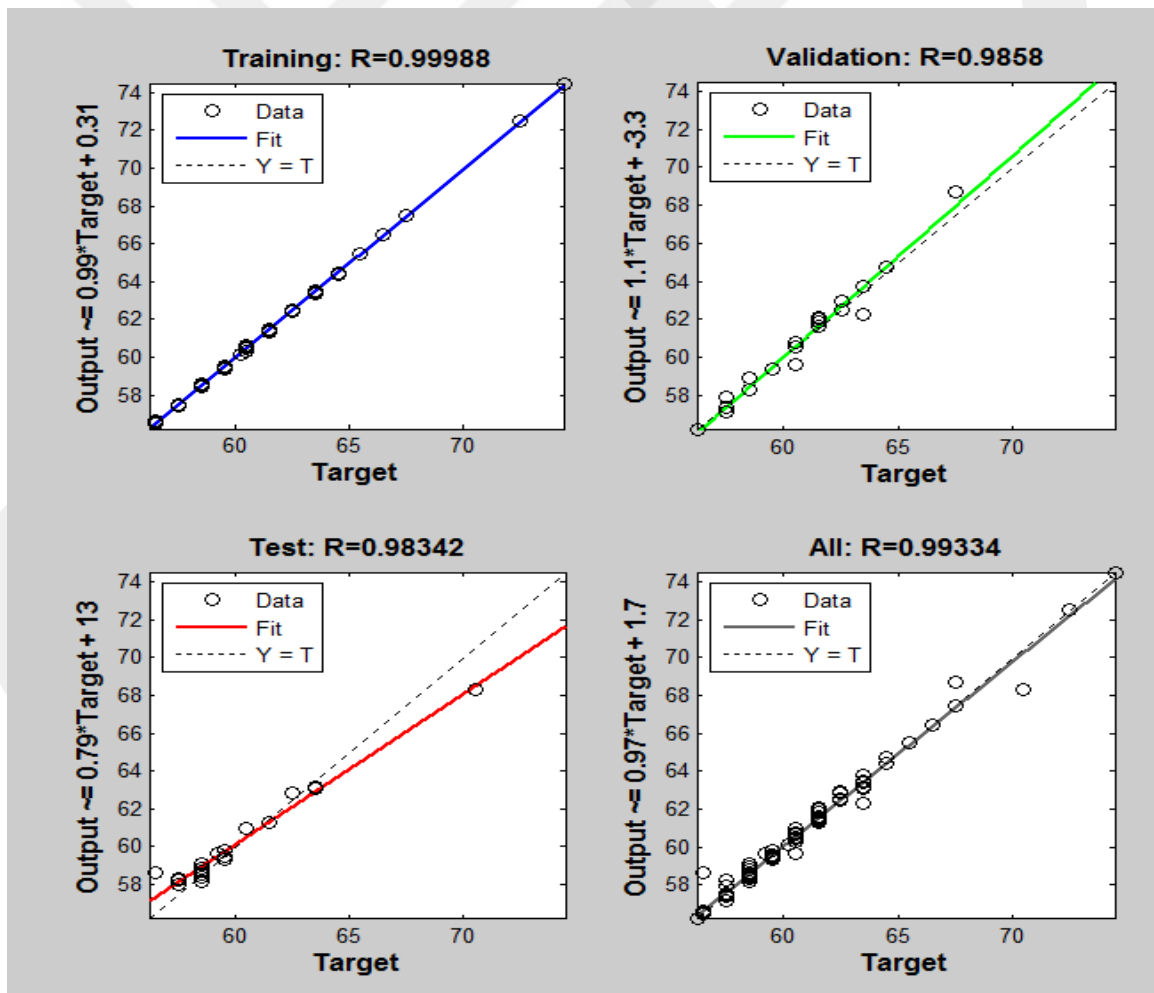


Figure 4.2 Regression plots for sample (A)

Figure 4.2 shows the regression plots, indicating the relationship between the output, and the target values.

If the training is very good, the network output and target will have the same value. The R all data set is found to be 0.99334; this indicates the suitability of perfect data (fit).

Generally, it can be considered that the results are acceptable with reasonable errors and very close to goals.

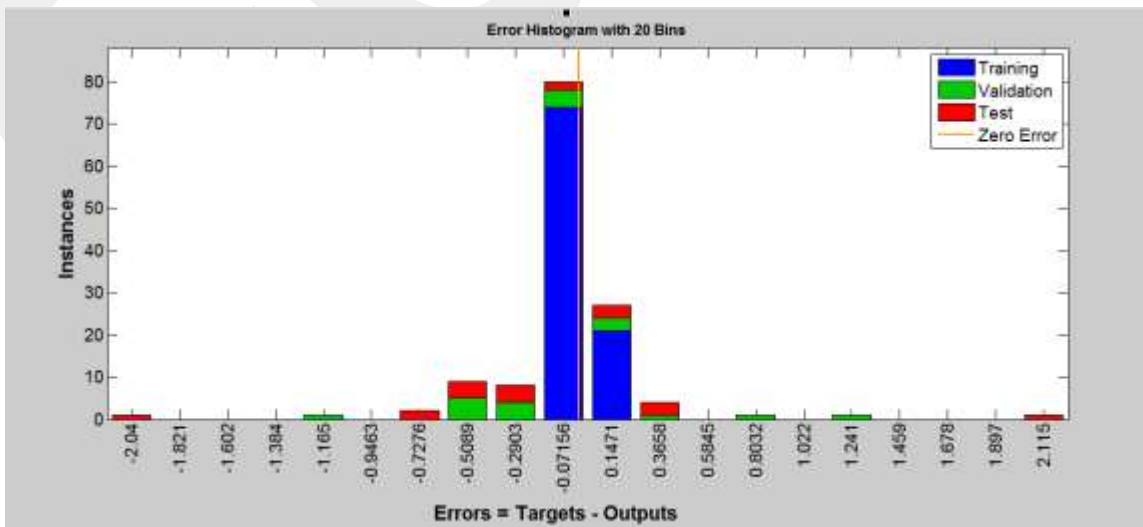
#### 4.1.4 Error Histogram of Sample (A)

Error Histogram is another way used to obtain more information about the neural network. Figure 4.3 exhibits the distribution of the training. It is clear from the figure that the validation and test results are acceptable and the error distribution results are reasonable.

As stated previously, the error differences between target and output which are shown in Figure 4.3 are calculated using the following equation:

$$\text{Errors } ^\circ\text{C} = (\text{Targets} - \text{Outputs}) ^\circ\text{C}$$

In addition, the figure gives a good indication of the values of outliers. Nevertheless, these outliers do not affect the results because five data out of 137 total data have approximate errors within 2  $^\circ\text{C}$  and most errors fall between 0.3658 $^\circ\text{C}$  and -0.7276 $^\circ\text{C}$ . These errors are considered acceptable and satisfied the training distribution results due to nature of the data of this study.



**Figure 4.3** Error histogram of sample (A)

#### 4.1.5 Test of Network for Sample (A)

The regression plots and error histogram explain the situation of the neural network behaviour with data training. The test network is considered the best indicator to see whether the network performance is good enough or no and the database used outside to the training database. Therefore, 60 samples outside the data are used. Satisfactory results are obtained; where, these results are R 0.937778 and MSE 0.22365.

#### 4.1.6 The Simplest Formula to Predict the Flash Point for Sample A

The equation for predicting the flash point created by the Polymath 6.10 program (explained in APPENDIX C POLYMATH) which shows the report that reveals the way of obtaining the simplest formula to predict the FP. The formula can be described as multiple linear regression (MLR). Multiple linear regression is the most popular form of linear regression analysis. The analysis of predictive, multiple linear regression is used to interpret the relationship between independent variables two or more and continuous dependent variable one.

$$FP = a_1*CN + a_2*Density + a_3*Gross\ heating\ combustion + a_4*Net\ heating\ combustion + a_5*PNA + a_6*ARO + a_7*Naphthalene + a_8*Smoke\ point + a_9*ARO-JET$$

Where,  $a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9$  are constants that can be determined in the regression analysis. The value of  $R^2$  equals to 0.9663 and the value of **RMSE** is 0.06380. This method was used by Alex Oliveira and others [40].

**Table 4.2** Constants of the equation for sample (A)

Variable	Value
a1	3.153108
a2	503.7732
a3	-54.93183
a4	45.42566
a5	-0.015655
a6	3.243207
a7	-6.422627
a8	-0.3260662
a9	-0.0687018

**4.1.7 Summary of the Results for Sample (A)****Table 4.3** Summary of the results for sample (A).

	Training	Validation	Testing	All	Test network
<b>R</b>	0.9998	0.9858	0.9884	0.9933	0.9377
<b>MSE</b>	0.0029	0.2730	0.6300	0.1364	0.1223
<b>Network performance of sample A</b>	The best validation performance is 0.2733 at epoch 10.				
<b>Error Histogram of Sample (A)</b>	Most errors fall between 0.3658 °C and -0.7276 °C (five databases of the total 137 data) with an error nearly 2 °C				
	Errors °C = Targets °C – Outputs °C				
<b>Simple equation to predict flash point MLR</b>	$FP = a_1 * CN + a_2 * Density + a_3 * Gross\ heating\ combustion + a_4 * Net\ heating\ combustion + a_5 * PNA + a_6 * ARO + a_7 * Naphthalene + a_8 * Smoke\ point + a_9 * ARO-JET$ ( R <sup>2</sup> equals to 0.9663, RMSE is 0.06380)				

## 4.2 Sample (B)

The output is a calculated flash point and inputs are the physical properties of the diesel experiment. Moreover, in this study, a big data set is used to train the model. This model can predict flash point which can be considered as great achievement and accomplishment.

### 4.2.1 Hidden Nodes for Sample (B)

The previous chapter involves the explanation of how to calculate hidden nodes and how to apply the thumb rules to determine the number of nodes in hidden layer. It was started from one node lower down the average input and output.

$$\text{Mean} = (5+1) / 2 = 3$$

Start point = 2 nodes.

For each node, a three-times and recorded the regression between the target and output for training, validation and testing results were trained. Finally, the averages were calculated.

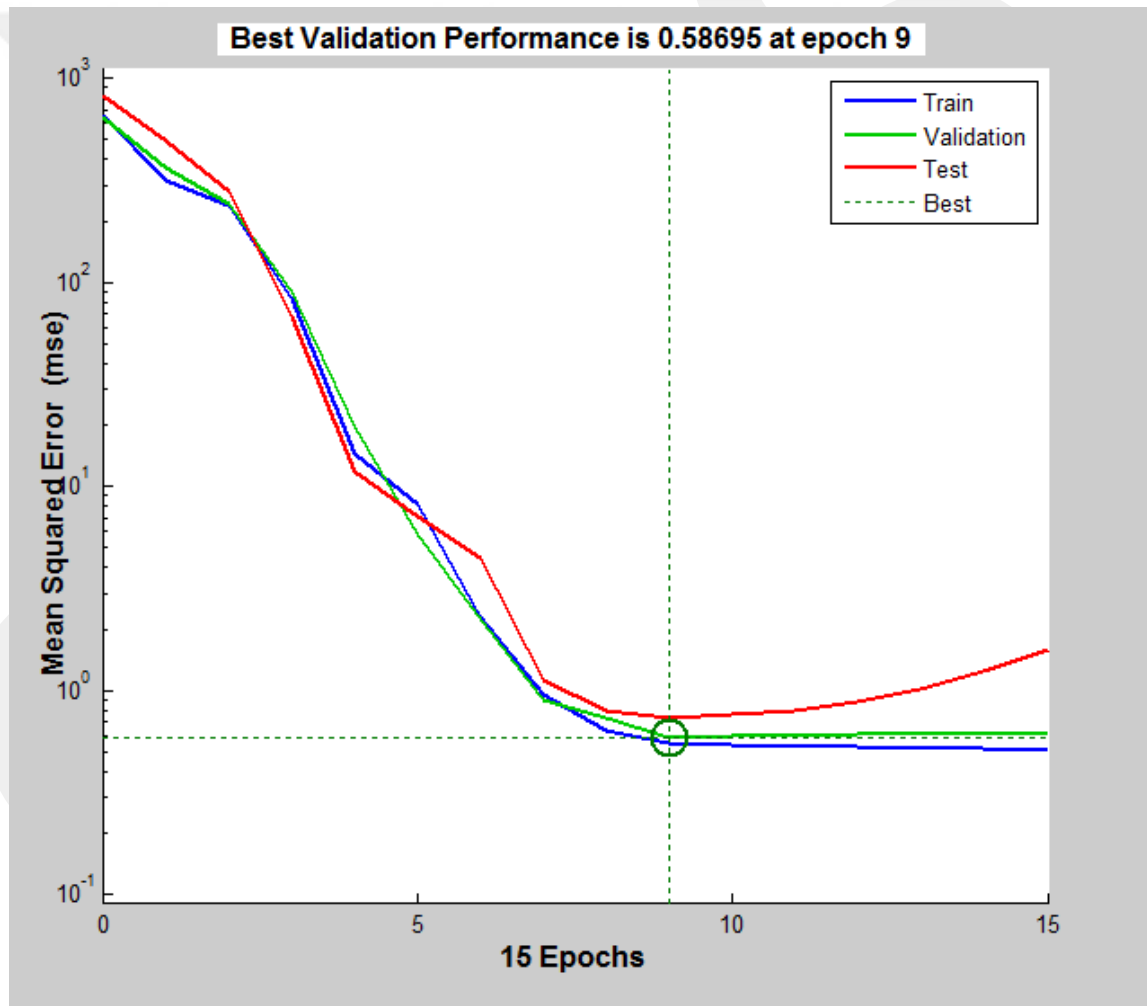
**Table 4.4** Hidden nodes regression results of sample (B)

<b>Node</b>	<b>Training</b>	<b>Validation</b>	<b>Testing</b>	<b>All</b>
2	0.988	0.978	0.970	0.976
3	0.981	0.980	0.969	0.976
4	0.982	0.977	0.970	0.976
5	0.978	0.973	0.975	0.975
6	0.978	0.980	0.965	0.974
7	0.980	0.976	0.962	0.972
8	0.984	0.972	0.972	0.977
9	0.980	0.977	0.979	0.978
<b>10</b>	<b>0.9942</b>	<b>0.994</b>	<b>0.9958</b>	<b>0.9958</b>

Table 4.4 shows the result of 2 to 10 nodes. The results are reasonable but 10 nodes have a regression coefficient over 0.9958. Because of the study has a big data set; it is preferred to choose a hidden number 10.

#### 4.2.2 The Network Performance for Sample (B)

After finding the hidden nodes, the evaluation of the network performance was achieved.



**Figure 4.4** Network performance of sample (B)

Figure 4.4 shows the performance plot of the training; note that the performance of the training reached to the minimum value at the iteration 9 and the training continued to the

iteration 15 then halts. Also this figure does not show any major problems during training, and the best validation performance is 0.5869 at epoch 9.

#### 4.2.3 The Regression Plots for Sample (B)

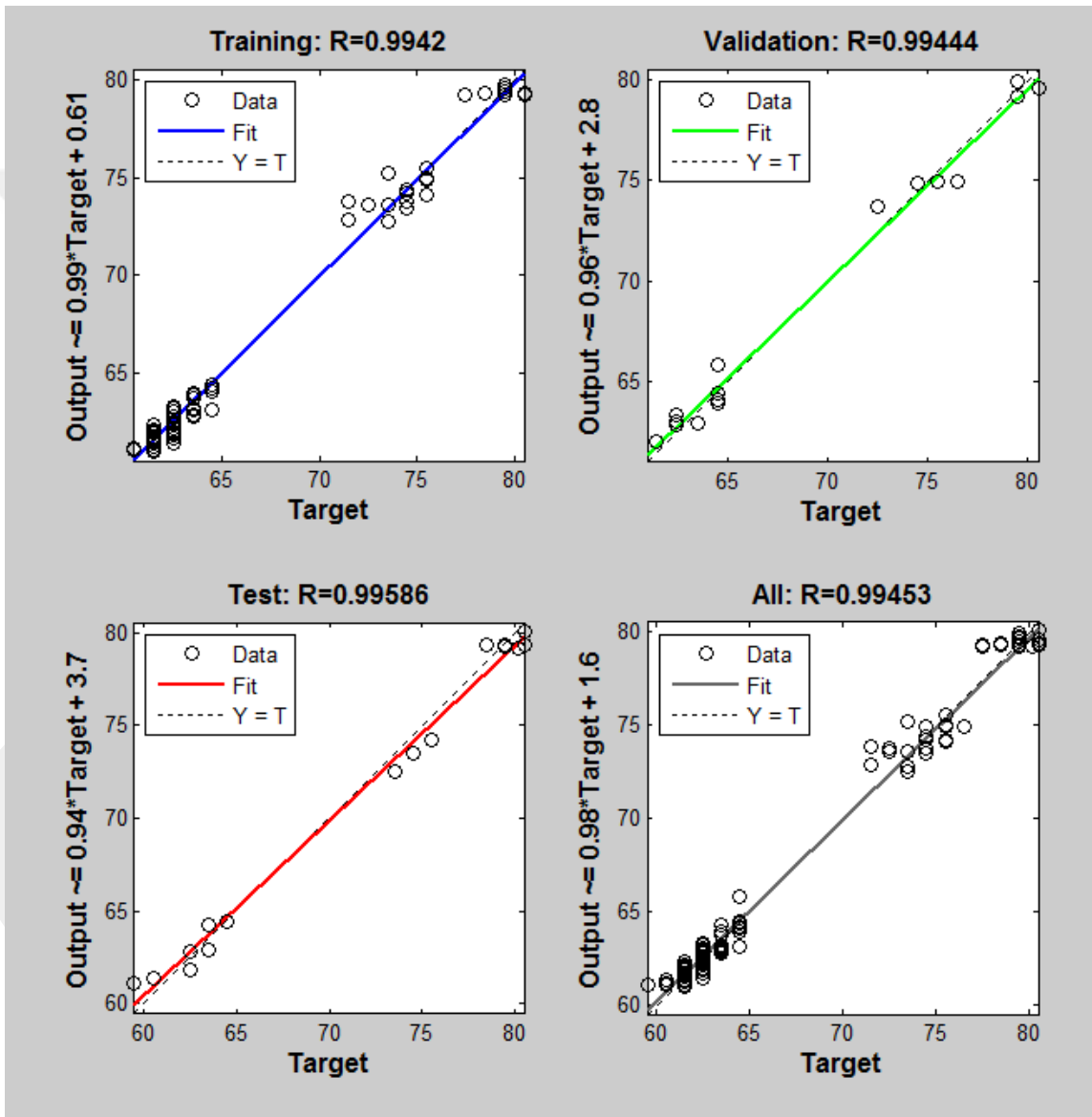


Figure 4.5 Regression plots of sample (B)

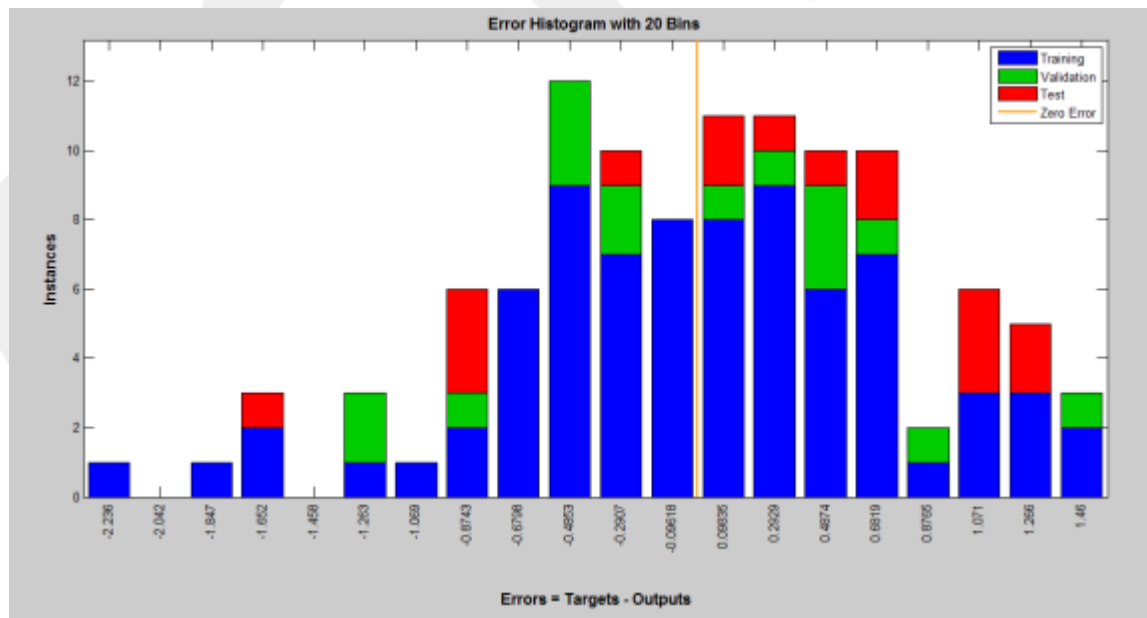
Figure 4.5 shows the regression plots, indicating the relationship between the output and the target values. If the training is very good, the network output and target will have the same value. The R value is found 0.9945 for all dataset. This indicates the suitability of perfect data (fit data).

#### 4.2.4 Error Histogram of Sample (B)

Figure 4.6 shows the distribution of the training. It is clear from the figure that the validation and test results are acceptable and the error distribution results are reasonable. As stated previously, the error differences between target and output which are shown in Figure 4.6 are calculated using the following equation:

$$\text{Errors } ^\circ\text{C} = (\text{Targets} - \text{Outputs}) ^\circ\text{C}$$

In addition, the figure gives a good indication of the values of outliers. However, these outliers do not affect the results because 12 databases out of 109 (the total data) have approximate errors nearly between 1.46  $^\circ\text{C}$  and -2.236  $^\circ\text{C}$ , and most errors fall between 1.266  $^\circ\text{C}$  and -1.263  $^\circ\text{C}$ .



**Figure 4.6** Error histogram of sample (B)

#### 4.2.5 Test of Network for Sample (B)

The regression plots and error histogram explain the situation of the neural network behavior with data training. The test network is considered the best indicator to see whether the network performance is good enough or not and the database is used outside the training database. Therefore, 60 samples outside the training data are used. Satisfactory results are obtained; these results are  $R = 0.99331$  and  $MSE = 0.7372$ . Figure 4.7 gives a good indication of the test network.

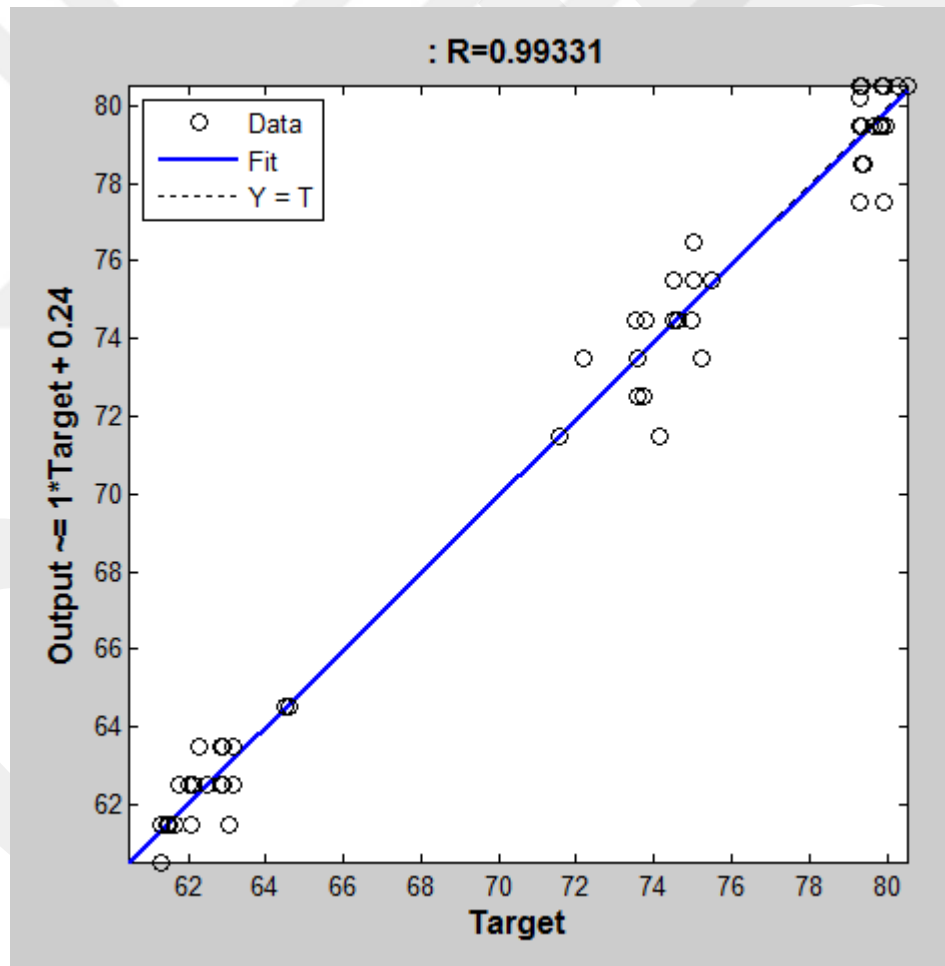


Figure 4.7 Test network of sample (B)

#### 4.2.6 The Simplest Formula to Predict the Flash Point for Sample (B)

The equation for predicting the flash point created by the Polymath program as explained in APPENDIX D shows the report that shows the way of obtaining the simplest formula to predict the FP. The formula can be described as multiple linear regression. MLR is the most popular form of linear regression analysis. The analysis of predictive, MLR is used to interpret the relationship between two or more independent variables and one continuous dependent variable.

$$\text{The equation } FP = a_1 * \text{Sulfur content} + a_2 * \text{Density} + a_3 * \text{Viscosity} + a_4 * \text{End point} + a_5 * \text{Distillation Residue}.$$

Where  $a_1, a_2, a_3, a_4, a_5$  are the constants that can be determined in the regression analysis. The value of  $R^2$  equals to 0.9797 which means that the relationship is multi-linear, and the value of **RMSE** is 0.0987.

**Table 4.5** Constants of the equation for sample (B)

Variable	Value
$a_1$	-12.23854
$a_2$	121.99
$a_3$	21.43645
$a_4$	-0.1727325
$a_5$	-5.279361

#### 4.2.7 Summary of the Results for Sample (B)

**Table 4.6** Summary of the results for sample (B)

	Training	validation	Testing	All	Test network
<b>R</b>	0.9942	0.9944	0.9958	0.9945	0.9933
<b>MSE</b>	0.549	0.586	0.7208	0.58109	0.7372

<b>Network performance of sample (B)</b>	The best validation performance is 0.5869 at epoch 9.
<b>Error Histogram of Sample (B)</b>	Most errors fall between 1.46 °C and -2.236 °C
	12 datasets out of 109 (the total data) have approximate errors nearly between 1.46 °C and -2.236 °C
	Errors = Targets – Outputs
<b>Simple equation to predict flash point (MLR)</b>	<p><i>The equation <math>FP = a_1 * \text{Sulfur content} + a_2 * \text{Density} + a_3 * \text{Viscosity} + a_4 * \text{End point} + a_5 * \text{Distillation Residue}</math>.</i></p> <p>(R<sup>2</sup> equals to 0.9797, RMSE is 0.0987)</p>

## CHAPTER 5

### CONCLUSIONS

In this chapter, the conclusions of this work are presented and some considerations related to data analysis and performance model policy, and future recommendations are given.

The general findings of this study are:

ANN is useful tool for predicting the flash point. ANN modules were developed to expect the flash point at perfect prediction based on the value of R which was greater than 0.98 (means it is a perfect fit).

For sample (A), the linear correlation coefficient R is calculated. The value of R equals to 0.9933, and the value of MSE equals to 0.1364. In addition, the best validation performance at epoch 10 is obtained and it is 0.2733. An equation is developed to predict the FP based on the composition and physical properties. FP has an accuracy of 99%. The predicted results have very small errors and they are reasonable, and very close to the actual value.

From applying the simplest formula for predicting the FP of sample (A), the coefficient of determination  $R^2$  is 0.966 and RMSE equal 0.0638. This formula can be described as MLR. Some positive changes in variable values of cetane number, density, net heating combustion, ARO, and ARO-JET are discovered and some negative changes in variable values of gross heat of combustion, PNA, naphthalene and smoke point are found. These results indicate that the decrease in smoke point and the increase in aromatic content are permitted, i.e., a relatively small increase in aromatics levels from 17.6 wt% (flash point is 57.5 °C and 14 mm of the smoke point) to 19.9 wt% (flash point is 74.5 °C - 11.8 mm of the smoke point).

For sample (B) ANN is used again. The input data are density, viscosity, sulfur content, distillation residue and distillation end point. The straight line shows linear relationship between the target and output. The high correlation coefficient (R) value is 0.9945 and the Mean Square Error (MSE) is 0.58109. The results are close to the actual value with small acceptable errors.

From applying the simplest formula for predicting the FP in sample (B), the coefficient of determination  $R^2$  is 0.9797 and RMSE equals to 0.0987. This formula can be described as a MLR. Some positive changes in variable values of density and viscosity are found and some negative changes in variable values of sulfur, distillation residue, and distillation end point are discovered. Overall, this study supports the idea which states that the increase in fuel flash point occurs when there is an increase in fuel viscosity and fuel density due to the increase in tension surface of diesel. The equation is developed to predict the FP based on the diesel physical properties in order to be able to predict FP with 97.9% accuracy.

This study serves laboratories in diesel power stations where a supervisor of the laboratory should use the simplest formula to predict FP, and these equations reduce the time and cost.

It is recommended that using a spreadsheet, e.g., Excel in MS-Office, to enter the simplest formula, then input the values of variables, so that the user can calculate the flash point.

In the future, prediction of the other parameters of diesel can be investigated and a study for the prediction of sulfur content in diesel which causes corrosion can be proposed.

## REFERENCES

- [1] J.Bacha, J.Freel, A. Gibbs, "Diesel fuels technical review," Freightliner, 2007.
- [2] T. Karhu, "The effect of renewable diesel oil on engine performance," Master's thesis , Turku University of Applied Sciences, Turkey, 2015.
- [3] "Annual energy review,2011," Indendent Statstics ans Analysis US Energy Information Administration, 2012.
- [4] "Short term energy outlook," Indendent Statstics ans Analysis US Energy Information Administration, 2013.
- [5] H. Aleme, P. Barbeira, "Determination of flash point and cetane index in diesel using distillation curves," *Fuel*, vol. 102, pp. 129–134, 2012.
- [6] D. Jones, P.Pujado, Handbook of petroleum processing, USA: Springer, 2006.
- [7] H.Heinemann, J. Speight, Fundamentals of petroleum and petrochemical engineering, United States of America: CRC press, 2011.
- [8] "Air pollutant emissions trends data," United States Environmental Protection Agency, 2005.
- [9] M. R. Riazi, Characterization and properties of petroleum fractions, USA: Printed in Baltimore, 2005.
- [10] R.Lee, J. Pedley, and C.Hobbs, "Fuel quality impact on heavy duty diesel emissions a literature review," *Journal of Fuels and Lubricants*, vol. 107, no. 4, 1998.
- [11] Smith, H.David , Wauquier, J.Pierre, Crude oil, petroleum products, process flowsheet, vol. 45, paris: Petroleum Refining, 1995.
- [12] "Automotive diesel specification," East African Community, 2011.
- [13] S. Haykin, Neural networks and learning machines (3rd edition), Canada: Pearson Education, 2009.

- [14] A. Reza, Predict the flow of well fluids a big data approach, University of Stavanger, June 2014, Master thesis.
- [15] D. Graupe, Principles artificial neural networks (3<sup>rd</sup> edition), USA: World Scientific, 2013.
- [16] J. Heaton, Introduction to neural networks with Java (Second Edition), United States of America: Heaton Research, 2008.
- [17] L. Phoon, A. Mustaffa, H. Hashim, R. Mat, "A review of flash point prediction models for flammable liquid mixtures," *Industrial & Engineering Chemistry Research*, vol. 53, pp. 12553–12565, 2014.
- [18] D. White, C. Beyler, C. Fulper, J. Leonard, "Flame spread on aviation fuels," *Fire Safety Journal*, vol. 215, pp. 1-31, 1997.
- [19] S. Lee, D. Myeong Ha, "The lower flash points of binary systems containing non-flammable component," *Korean Journal of Chemical Engineering*, vol. 20, no. 5, pp. 799–802, September 2003.
- [20] K. Satyanarayana, P. G. Rao, "Improved equation to estimate flash points of organic compounds," *Journal of Hazardous Materials*, vol. 32, no. 1, pp. 81-85, 1992.
- [21] F. Carroll, J. Godinho, and F. Quina, "Development of a simple method to predict boiling points and flash points of acyclic alkenes," *Industrial & Engineering Chemistry Research*, vol. 50, no. 24, p. 14221–14225, 2011.
- [22] L. Catoirea, "A unique equation to estimate flash points of selected pure liquids," *Journal of Physical and Chemical Reference Data*, vol. 33, no. 4, pp. 1083-1111, 2005.
- [23] E. M. Valenzuela, R. Vazquez-Roman, S. Patel, M. Sam Mannan, "Prediction models for the flash point of pure components," *Journal of Loss Prevention in the Process Industries*, vol. 24, pp. 753-757, 2011.
- [24] F. Gharagheizi, A. Eslamimanesh, A. Mohammadi, D. Richon, "Empirical method for representing the flash point temperature of pure compounds," *Industrial & Engineering Chemistry Research*, vol. 50, pp. 5877–5880, 2011.

- [25] J.Lazzus, "Neural network particle swarm modeling to predict thermal," *Mathematical and Computer Modelling*, vol. 57, pp. 2408–2418, 2013.
- [26] F.Gharagheizi, P. Ilani-kashkouli, S.Mirkhani, A. Mohammadi, "Computation of upper flash point of chemical compounds using a chemical structure-based model," *Ind. Eng. Chem. Res.*, vol. 51, pp. 5103–5107, 2012.
- [27] M. Keshavarz , M.Ghanbarzadeh, "Simple method for reliable predicting flash points of unsaturated hydrocarbons," *Journal of Hazardous Material*, vol. 193, pp. 335–341, 2011.
- [28] L. Catoirea, S. Paulmier, "Estimation of closed cup flash points of combustible solvent blends," *Journal of Physical and Chemical Reference Data*, vol. 35, no. 1, 2005.
- [29] L. Catoire, S. Paulmier, V.Naudet, "Experimental determination and estimation of closed cup flash points of mixtures of flammable solvent," *Process Safety Progress*, vol. 25, no. 1, pp. 33-39, 2006.
- [30] M.Hristova, D.Damgaliev, J. Hristov, "Practical data correlation of flash points of binary mixtures by a reciprocal function," *Thermal Science*, vol. 15, no. 3, pp. 905-910, 2011.
- [31] S. Y. Kim, B.Lee, "A prediction model for the flash point of binary liquid mixtures," *journal of Loss Prevention in the Process Industries*, vol. 23, pp. 166–169, 2010.
- [32] J.MejJa, N.Salgado, C.Orrego, "Effect of blends of diesel and palm castor biodiesels on viscosity, cloud point and flash point," *Industrial Crops and Products*, vol. 43, pp. 791-797, 2013.
- [33] D.Korres, G.Anastopoulos, E.Loiz, A.Alexandridis, H.Sarimveis, G.Bafas, "A neural network approach to the prediction of diesel fuel lubricity," *Fuel*, vol. 10, p. 1243–1250, 2002.
- [34] R. Balabin, R. Safieva, "Near-Infrared (NIR) Spectroscopy for biodiesel analysis, fractional composition, iodine value, and cold filter plugging point from one vibrational spectrum," *Fuel and Energy*, vol. 25, no. 5, pp. 2373-2382, 2011.

- [35] B. Basu, G. S. Kapur, A.Sarpal, R. Meusinger, "A neural network approach to the prediction of cetane number of diesel fuels using nuclear magnetic resonance (NMR) Spectroscopy," *Energy & Fuels*, vol. 17, pp. 1570-1575, 2003.
- [36] F. Oliveira, L. Carvalho, L. Teixeira, R. Sales, "Predicting cetane index, flash point and content sulfur of diesel-biodiesel blend using an artificial neural network (ANN) model," *Energy Fuels*, vol. 31, no. 4, pp. 3913-392, March 21, 2017.
- [37] J.Kumar, A.Bansal, "Application of artificial neural network to predict properties of diesel-biodiesel blends," *Kathmandu University Journal of Science, Engineering and Technology*, vol. 6, no. 2, pp. 98-103, November 2010.
- [38] M. Agarwal , K.Singh , S.Chaurasia, "Prediction of biodiesel properties from fatty acid composition using linear regression and ANN techniques," *Journal Indian Chemical Engineer*, vol. 52 (4), no. 4, pp. 347-361, 2010.
- [39] G.Liu, L Wang, H. Qu , H. Shen, X. Zhang, S. Zhang, Z. Mi, "Artificial neural network approaches on composition property relationships of jet fuels based on GC-MS," *Fuel*, vol. 86, no. 16, pp. 2551-2559, November 2007.
- [40] A. Filho, A. Barros, S.Labidi, I. Viegas, D.Marques, A.Romariz, R, Sousa, , A.Marques, E.Marques, "Application of artificial neural networks to predict viscosity, iodine value and induction period of biodiesel focused on the study of oxidative stability," *Fuel*, vol. 145, pp. 127-135, 2015.
- [41] J.Kumar, A.Bansal, M.K.Jha, "Comparison of statistical and neural network techniques in predicting physical properties of various mixtures of diesel and biodiesel," in *Proceedings of the World Congress on Engineering and Computer Science*, San Francisco, USA, October 24-26-2007.
- [42] H.Yang, Z. Ring, Y.Briker, N. McLean, W.Friesen, C.Fairbridge, "Neural network prediction of cetane number and density of diesel fuel from its chemical composition determined by LC and GC-MS," *Fuel*, vol. 81, no. 1, pp. 65-74, January 2002.

## APPENDIX

### A. DATA RELATIVE TO CHAPTER 3 (SAMPLE A)

No	FP (°C)	CN	Density g/cc	Gross MJ/Kg	Net MJ/Kg	PNA wt%	ARO wt%	Naphthalene vol%	Smoke point m.m	ARO-JET wt%
1.	56.5	51.6	0.8177	45.499	43.056	3.8	18.2	4.2	14	17.3
2.	56.5	51.4	0.8185	45.507	43.063	3.7	18	4.1	14.2	17.1
3.	56.5	52.4	0.8233	45.467	43.047	4	18.2	4.9	13.5	17.1
4.	56.5	52.1	0.821	45.5	43.069	3.9	18.1	4.5	14	17.1
5.	56.2	53.5	0.8236	45.464	43.042	4.4	18.1	4.8	12	16.6
6.	56.5	51.5	0.8203	45.51	43.069	3.9	18.4	4.3	14.5	17.5
7.	56.5	52.6	0.8227	45.468	43.041	4	17.8	4.7	13	16.6
8.	56.5	52.2	0.8204	45.493	43.061	3.8	17.9	4.4	13.6	16.8
9.	56.5	52	0.8201	45.499	43.061	3.6	17.7	4.3	14	16.7
10.	57.5	51.8	0.8203	45.501	43.064	3.6	17.6	4.3	14	16.6
11.	57.5	51.8	0.8188	45.484	43.042	3.7	17.8	4.2	13.6	16.8
12.	56.5	51.7	0.8185	45.499	43.059	3.7	17.8	4.2	13.9	16.8
13.	57.5	52.1	0.8213	45.449	43.021	3.9	18	4.6	12.9	16.9
14.	57.5	52.6	0.8235	45.471	43.05	3.9	17.8	4.7	13.1	16.5
15.	57.5	52.6	0.8247	45.451	43.027	4.1	18.1	4.8	12.6	16.9
16.	57.5	52.2	0.8251	45.445	43.022	4.1	18.3	4.9	12.8	17.2
17.	57.5	52.3	0.8237	45.448	43.023	4	18.2	4.8	12.8	17.1
18.	57.5	51.7	0.8232	45.5	43.063	3.8	18.1	4.5	14.1	17.2
19.	57.5	52.5	0.8229	45.479	43.053	3.9	17.7	4.6	13.2	16.5
20.	57.5	52.1	0.8214	45.452	43.02	4	18.1	4.6	12.8	17.1
21.	57.5	52	0.8196	45.482	43.042	3.9	18	4.2	13.3	16.9
22.	57.5	52.2	0.8225	45.446	43.019	4	18.1	4.7	12.9	17

23.	58.5	52.7	0.8194	45.446	43.019	4.1	18.2	4.6	12.3	17
24.	58.5	51.9	0.8211	45.516	43.079	3.6	17.8	4.3	14.4	16.8
25.	58.5	52.1	0.8242	45.441	43.011	4.1	18.2	4.7	12.5	17.1
26.	58.5	52.3	0.8247	45.451	43.023	4	18.2	4.8	12.8	17.2
27.	58.5	52.5	0.8267	45.473	43.048	3.9	17.7	4.8	13	16.6
28.	58.5	52.5	0.8234	45.476	43.051	3.9	17.9	4.8	13.4	16.7
29.	58.5	54.4	0.8186	45.449	43.035	4.6	18	4.8	11.1	16.2
30.	58.5	52.1	0.8224	45.476	43.048	3.9	18.1	4.6	13.5	17.1
31.	58.5	52	0.8223	45.465	43.034	3.9	18.1	4.6	13.2	17.1
32.	58.5	52.4	0.8223	45.437	43.012	4.1	18.1	4.8	12.5	17.1
33.	58.5	51.7	0.8221	45.495	43.055	3.8	18.1	4.4	13.8	17.1
34.	58.5	52.1	0.8199	45.496	43.067	4	18.1	4.5	13.9	17.1
35.	58.5	52.2	0.8221	45.477	43.02	4	18.2	4.7	12.8	17.1
36.	58.5	51.9	0.82	45.508	43.073	3.9	17.9	4.4	14.1	16.9
37.	58.5	51.5	0.8206	45.528	43.087	3.7	18.2	4.3	14.9	17.3
38.	58.5	52.1	0.821	45.504	43.074	3.8	17.9	4.4	14.1	16.8
39.	58.5	53	0.8202	45.477	43.052	4.1	17.9	4.5	12.7	16.6
40.	58.5	52.3	0.8261	45.457	43.033	4.1	18.2	4.9	12.9	17.1
41.	58.5	52.2	0.8255	45.459	43.036	4	18.2	4.9	13.1	17.1
42.	58.5	52.5	0.8158	45.513	43.07	3.7	18	4.1	14.9	17.2
43.	58.5	52.4	0.8207	45.465	43.039	3.8	17.6	4.5	12.9	16.4
44.	58.5	52.5	0.8223	45.471	43.047	3.9	17.7	4.6	13.2	16.5
45.	59.5	52.4	0.8196	45.453	43.028	4.1	18.1	4.6	12.7	17
46.	59.5	52.1	0.8206	45.448	43.018	3.9	18	4.6	12.8	17
47.	59.5	52	0.8214	45.497	43.059	3.7	17.8	4.4	13.8	16.8
48.	59.5	52	0.8228	45.508	43.073	3.8	18	4.5	14	16.9
49.	59.5	52	0.823	45.451	43.019	3.9	18.1	4.6	12.9	17.1
50.	59.5	52.1	0.8221	45.472	43.044	3.9	18	4.6	13.1	16.9
51.	59.2	53.4	0.8178	45.494	43.067	4.2	17.9	4.5	12.8	16.6
52.	59.5	52.3	0.818	45.47	43.039	3.8	17.6	4.4	13.3	16.5
53.	59.5	52.6	0.825	45.441	43.013	4.1	18.3	4.8	12.3	17.1

54.	59.5	52.9	0.8209	45.464	43.045	4.2	18.1	4.7	12.8	16.9
55.	59.5	52.3	0.8183	45.45	43.025	4.1	18.2	4.5	12.6	17.1
56.	59.5	52.5	0.8183	45.462	43.036	3.9	17.9	4.5	12.8	16.7
57.	59.5	52.7	0.8206	45.485	43.062	4.1	17.9	4.6	13.2	16.7
58.	59.5	53	0.8212	45.484	43.058	4	17.8	4.6	13.3	16.5
59.	59.5	53.1	0.8207	45.492	43.064	4	17.9	4.6	13.2	16.6
60.	59.5	53.5	0.8172	45.524	43.096	4.3	18	4.3	13	16.5
61.	59.5	53.5	0.8172	45.524	43.096	4.3	18	4.3	13	16.5
62.	59.5	52.5	0.8236	45.475	43.054	4	18	4.9	13.5	16.9
63.	59.5	52.2	0.8235	45.441	43.012	4	18.1	4.7	12.6	17.1
64.	59.5	52.2	0.8255	45.459	43.036	4	18.2	4.9	13.1	17.1
65.	59.5	52.3	0.8201	45.453	43.027	4	18.2	4.7	12.9	17.1
66.	59.5	52.2	0.8239	45.449	43.01	4	18.2	4.7	12.7	17.1
67.	59.5	53.1	0.8224	45.461	43.038	4.2	18.1	4.8	12.4	16.8
68.	59.5	52.8	0.8208	45.461	43.029	4	17.8	4.5	12.6	16.6
69.	60.5	52.9	0.822	45.486	43.06	3.9	17.7	4.6	13.2	16.5
70.	60.5	52	0.8213	45.448	43.016	4	18.2	4.6	12.9	17.2
71.	60.5	52.2	0.8218	45.461	43.037	3.9	18	4.7	13.3	16.9
72.	60.5	52.9	0.8214	45.459	43.037	4.2	18.1	4.7	12.4	16.8
73.	60.5	52.3	0.822	45.463	43.035	4	18.1	4.6	12.8	16.9
74.	60.5	52.8	0.8203	45.474	43.048	4.1	18.2	4.7	13.1	17
75.	60.5	53.9	0.8201	45.459	43.034	4.5	18.1	4.8	11.7	16.5
76.	60.5	52.2	0.8206	45.462	43.033	4	18.2	4.5	12.9	17.1
77.	60.5	52.5	0.8198	45.463	43.033	4.1	18.4	4.6	12.9	17.2
78.	60.5	52.5	0.8218	45.452	43.03	4.1	18.1	4.7	12.7	16.9
79.	60.5	54.2	0.822	45.478	43.045	3.9	18	4.5	13.3	16.9
80.	60.2	52.1	0.8189	45.489	43.058	3.9	18	4.4	13.6	16.9
81.	60.5	53	0.8174	45.498	43.077	4.2	18.1	4.6	13.5	16.8
82.	60.5	52.6	0.8216	45.548	43.031	4.1	18.1	4.7	12.6	16.9
83.	60.5	52.2	0.8242	45.449	43.021	4	18.2	4.8	12.8	17.1
84.	60.5	52.2	0.8225	45.454	43.026	4.1	18.2	4.6	12.7	17

85.	60.5	52.3	0.824	45.471	43.04	4.1	18.3	4.7	13	17.1
86.	60.5	52.3	0.8221	45.455	43.029	3.9	17.9	4.6	13	16.7
87.	60.5	52.7	0.8216	45.482	43.053	3.9	17.8	4.6	13.1	16.6
88.	61.5	52.3	0.8231	45.439	43.014	4	18.1	4.8	12.6	17
89.	61.5	53.3	0.8201	45.462	43.041	4.3	18.1	4.3	12.3	16.7
90.	61.5	53.4	0.8218	45.477	43.058	4.3	18.1	4.8	12.6	16.6
91.	61.5	53.1	0.8203	45.457	43.036	4.2	18	4.7	12.2	16.6
92.	61.5	52.5	0.8202	45.46	43.036	4	18.1	4.6	12.9	17
93.	61.5	52	0.8216	45.477	43.037	3.8	18.1	4.3	13.4	17.1
94.	61.5	52.6	0.8211	45.477	43.043	4	18	4.5	12.9	16.9
95.	61.5	52.3	0.8211	45.454	43.016	3.9	18.1	4.4	12.7	17
96.	61.5	52.7	0.8269	45.445	43.023	4	17.8	4.9	12.6	16.7
97.	61.5	53	0.8202	45.478	43.054	4.1	17.9	4.6	12.9	16.7
98.	61.5	52.5	0.8211	45.452	43.029	3.9	17.7	4.6	12.6	16.4
99.	61.5	52.2	0.824	45.442	43.014	4.1	18.2	4.8	12.5	17.2
100.	61.5	54.1	0.8191	45.48	43.066	4.5	18	4.7	12.1	16.4
101.	61.5	52.5	0.8213	45.461	43.038	4	17.9	4.7	13	16.7
102.	61.5	53.7	0.8219	45.464	43.046	4.4	18.1	4.8	11.8	16.5
103.	61.5	52.7	0.8208	45.458	43.038	3.9	17.6	4.8	12.8	16.4
104.	61.5	52.9	0.82	45.467	43.042	4.1	18	4.6	12.7	16.7
105.	62.5	53.4	0.8197	45.46	43.04	4.3	18.2	4.7	12.2	16.7
106.	62.5	53.3	0.8202	45.46	43.036	4.3	18.2	4.7	12	16.8
107.	62.5	52.5	0.824	45.457	43.034	4	18	4.7	12.7	16.8
108.	62.5	52.3	0.8232	45.453	43.02	3.8	17.9	4.6	12.9	16.8
109.	62.5	52.6	0.8232	45.47	43.046	3.9	17.6	4.7	12.9	16.4
110.	62.5	52.7	0.8256	45.466	43.036	4	18	4.7	12.8	16.8
111.	62.5	53.1	0.8224	45.48	43.057	4.3	18.2	4.7	12.8	16.8
112.	62.5	52.5	0.8196	45.444	43.018	4.1	18.2	4.7	12.6	17.1
113.	62.5	53.1	0.8222	45.498	43.077	4.1	17.8	4.7	13.4	16.5
114.	62.5	52.6	0.8228	45.458	43.034	4.1	18.1	4.7	12.6	16.8
115.	63.5	52.6	0.8206	45.46	43.033	3.9	17.9	4.6	12.9	16.8

116.	63.5	52.4	0.8207	45.437	43.008	4	18	4.6	12.4	16.9
117.	63.5	52.5	0.8237	45.495	43.067	3.7	17.6	4.6	13.9	16.6
118.	63.5	52.1	0.8237	45.474	43.047	4	18.2	4.7	13.3	17.2
119.	63.5	53	0.825	45.449	43.029	4.2	18.1	4.9	12.3	16.8
120.	63.5	54.3	0.8157	45.505	43.086	4.5	17.9	4.5	12.3	16.2
121.	63.5	54	0.8186	45.501	43.075	4.4	18	4.5	12.3	16.4
122.	63.5	52.9	0.8225	45.472	43.05	4.2	18.1	4.7	12.6	16.8
123.	63.5	52.7	0.823	45.472	43.043	4.2	18.2	4.6	12.7	16.9
124.	63.5	52.3	0.8253	45.472	43.046	4	18.1	4.8	13.2	17
125.	63.5	52.3	0.823	45.485	43.051	3.9	18	4.5	13.3	16.9
126.	63.5	53	0.8248	45.474	43.046	4.1	18.1	4.7	12.7	16.9
127.	64.5	53	0.8195	45.476	43.052	4.2	18.2	4.6	12.8	16.8
128.	64.5	53.2	0.8228	45.466	43.042	4.3	18.3	4.8	12.3	16.9
129.	64.5	53	0.818	45.49	43.066	4.3	18.3	4.5	12.9	17
130.	65.5	52.8	0.8229	45.464	43.041	4	18	4.7	12.9	16.8
131.	66.5	53.8	0.8214	45.449	43.029	4.5	18.2	4.8	11.3	16.6
132.	67.5	52.9	0.8231	45.458	43.04	4.3	18.3	4.9	12.3	16.9
133.	67.5	54.2	0.8212	45.478	43.065	4.5	17.9	4.7	12	16.2
134.	70.5	53	0.8193	45.486	43.064	4.2	18.1	4.6	13.1	16.8
135.	72.5	53.6	0.8221	45.464	43.074	4.3	18.1	4.8	12.1	16.6
136.	74.5	53.9	0.8201	45.456	43.042	4.4	18	4.8	11.8	16.4
137.	74.5	52.5	0.8309	45.559	43.062	4.5	19.9	3.3	11.6	19.6

**B. DATA RELATIVE TO CHAPTER 3 (SAMPLE B)**

<b>NO</b>	<b>FP (°C)</b>	<b>Sulfur content wt%</b>	<b>Density at 16 °C</b>	<b>Viscosity at 40 °C</b>	<b>Distillation End Point</b>	<b>Distillation Residue wt %</b>
1	62.5	0.4342	0.8276	2.685	360	5.5
2	63.5	0.4275	0.8278	2.685	360	5.5
3	62.5	0.4408	0.828	2.685	360	5.5
4	61.5	0.4605	0.828	2.685	360	5.5
5	64.5	0.4331	0.828	2.685	360	5.5
6	64.5	0.4254	0.828	2.685	360	5.5
7	63.5	0.4292	0.828	2.685	360	5.5
8	60.5	0.4349	0.876	2.685	360	5.5
9	62.5	0.4419	0.8277	2.685	360	5.5
10	61.5	0.4394	0.8277	2.685	360	5.5
11	62.5	0.4348	0.828	2.685	360	5.5
12	64.5	0.4356	0.8281	2.685	360	5.5
13	63.5	0.4364	0.8281	2.685	360	5.5
14	64.5	0.4886	0.8277	2.863	360	5.5
15	64.5	0.4486	0.828	2.863	360	5.5
16	62.5	0.4578	0.828	2.863	360	5.5
17	64.5	0.4375	0.8279	2.863	360	5.5
18	64.5	0.4351	0.8279	2.863	360	5.5
19	64.5	0.4378	0.8279	2.863	360	5.5
20	64.5	0.4371	0.8279	2.863	360	5.5
21	63.5	0.4392	0.8279	2.863	360	5.5
22	60.5	0.438	0.8279	2.863	360	5.5
23	61.5	0.7371	0.8293	2.452	360	4
24	61.5	0.7493	0.8295	2.452	360	4
25	61.5	0.7359	0.8296	2.452	360	4
26	62.5	0.7639	0.8279	2.452	360	4
27	62.5	0.7525	0.8298	2.452	360	4
28	64.5	0.4429	0.8279	2.685	360	5.5
29	63.5	0.4357	0.828	2.685	360	5.5
30	62.5	0.4309	0.828	2.685	360	5.5
31	64.5	0.4622	0.828	2.685	360	5.5
32	62.5	0.427	0.8281	2.685	360	5.5
33	60.5	0.7372	0.8293	2.452	360	4
34	61.5	0.7565	0.8292	2.452	360	4
35	62.5	0.7382	0.8295	2.452	360	4
36	63.5	0.4352	0.828	2.685	360	6
37	63.5	0.4299	0.828	2.685	360	6

38	76.5	0.4563	0.8248	2.768	360	2.5
39	79.5	0.4591	0.8248	2.768	360	2.5
40	80.5	0.4531	0.8247	2.768	360	2.5
41	77.5	0.4518	0.8248	2.768	360	2.5
42	74.5	0.7626	0.8233	2.246	310	2
43	75.5	0.7756	0.8232	2.246	310	2
44	76.5	0.7759	0.8232	2.246	310	2
45	74.5	0.7741	0.8232	2.246	310	2
46	73.5	0.7772	0.8231	2.246	310	2
47	75.5	0.7821	0.8231	2.768	360	2.5
48	79.5	0.4648	0.8247	2.768	360	2.5
49	81.5	0.4573	0.8247	2.768	360	2.5
50	81.5	0.4482	0.8248	2.768	360	2.5
51	80.5	0.4561	0.8247	2.768	360	2.5
52	78.5	0.4562	0.8247	2.768	360	2.5
53	67.5	0.7747	0.8231	2.246	310	2.5
54	82.5	0.466	0.8249	2.768	360	2.5
55	79.5	0.4492	0.8248	2.768	360	2.5
56	80.2	0.4491	0.8248	2.768	360	2.5
57	78.5	0.4508	0.8247	2.768	360	2.5
58	80.5	0.4534	0.8248	2.768	360	2.5
59	76.5	0.7727	0.8231	2.199	310	2
60	75.5	0.776	0.8231	2.199	310	2
61	76.5	0.7737	0.8232	2.199	310	2
62	74.5	0.7776	0.8231	2.199	310	2
63	71.5	0.7775	0.8231	2.199	310	2
64	79.5	0.4541	0.8249	2.778	360	2.5
65	79.5	0.4509	0.8249	2.778	360	2.5
66	75.5	0.4507	0.8248	2.778	360	2.5
67	80.5	0.4514	0.8248	2.778	360	2.5
68	80.5	0.4527	0.8249	2.778	360	2.5
69	77.5	0.447	0.8248	2.778	360	2.5
70	72.5	0.7764	0.8233	2.186	310	2
71	74.5	0.7778	0.8233	2.186	310	2
72	73.5	0.6993	0.8231	2.186	310	2
73	71.5	0.7708	0.8231	2.186	310	2
74	79.5	0.4676	0.8246	2.749	360	2.2
75	80.5	0.4343	0.8246	2.749	360	2.2
76	79.5	0.4523	0.8246	2.749	360	2.2
77	80.5	0.449	0.8264	2.749	360	2
78	71.5	0.7155	0.8233	2.158	310	2

79	74.5	0.7799	0.8232	2.158	310	2
80	73.5	0.7828	0.8232	2.158	310	2
81	72.5	0.7843	0.8232	2.158	310	2
82	70.5	0.7516	0.8234	2.158	310	2
83	74.5	0.7854	0.8234	2.158	310	2
84	76.5	0.4469	0.8248	2.749	360	2.2
85	79.5	0.4447	0.8247	2.749	360	2.2
86	79.5	0.455	0.8247	2.749	360	2.2
87	77.5	0.445	0.8246	2.749	360	2.2
88	76.5	0.7802	0.8235	2.158	310	2
89	61.5	0.636	0.828	2.3333	360	3.5
90	60.5	0.6221	0.8288	2.3333	360	3.5
91	59.5	0.6302	0.8288	2.3333	360	3.5
92	60.5	0.6372	0.8289	2.3333	360	3.5
93	61.5	0.6389	0.829	2.3333	360	3.5
94	62.5	0.639	0.8291	2.3333	360	3.5
95	61.5	0.6233	0.8266	2.36	360	3.5
96	64.5	0.6702	0.8254	2.36	360	3.5
97	60.5	0.6235	0.8262	2.36	360	3.5
98	62.5	0.6377	0.8264	2.36	360	3.5
99	63.5	0.6485	0.8248	2.36	360	3.5
100	64.5	0.6677	0.8254	2.36	360	3.5
101	63.5	0.62	0.8239	2.36	360	3.5
102	62.5	0.5289	0.8312	2.42	360	3.8
103	62.5	0.6386	0.8309	2.42	360	3.8
104	63.5	0.737	0.8302	2.454	360	3.6
105	61.5	0.7621	0.8292	2.454	360	3.6
106	62.5	0.7409	0.8287	2.454	360	3.6
107	60.5	0.7524	0.8286	2.437	360	3.3
108	57.5	0.7573	0.8282	2.437	360	3.3
109	62.5	0.7505	0.828	2.437	360	3.3
110	59.5	0.7517	0.8292	2.437	360	3.6
111	58.5	0.7581	0.8273	2.437	360	3.4
112	58.5	0.7634	0.8275	2.437	360	3.4
113	59.5	0.7493	0.8291	2.437	360	4
114	58.5	0.7572	0.8279	2.437	360	3.4
115	58.5	0.7525	0.8291	2.437	360	4
116	60.5	0.7458	0.8291	2.437	360	4
117	58.5	0.7456	0.8293	2.437	360	3.5
118	59.5	0.7287	0.8285	2.437	360	3.5
119	59.5	0.7455	0.8291	2.437	360	3.5

120	60.5	0.7402	0.8292	2.437	360	3.5
121	58.5	0.7489	0.8288	2.437	360	3.5
122	59.5	0.7517	0.8289	2.437	360	3.5
123	57.5	0.7999	0.8264	2.618	360	4
124	58.5	0.7983	0.8263	2.618	360	4
125	58.5	0.8147	0.8263	2.618	360	4
126	57.5	0.8062	0.8268	2.618	360	4
127	59.5	0.7786	0.8277	2.605	360	3.8
128	60.5	0.7825	0.8293	2.605	360	3.8
129	58.5	0.8303	0.8291	2.605	360	4
130	64.5	0.7675	0.8291	2.614	360	3.8
131	62.5	0.4454	0.8276	2.833	360	5.3
132	63.5	0.3969	0.8278	2.833	360	5.3
133	62.5	0.3915	0.828	2.833	360	5.3
134	63.5	0.4347	0.8281	2.833	360	5.3
135	61.5	0.4259	0.828	2.833	360	5.3
136	61.5	0.4327	0.8278	2.833	360	5.4
137	63.5	0.4325	0.828	2.833	360	5.4
138	61.5	0.4269	0.8279	2.833	360	5.4
139	64.5	0.436	0.8279	2.833	360	5.4
140	64.5	0.4297	0.828	2.833	360	5.4
141	63.5	0.4302	0.8278	2.833	360	5.5
142	63.5	0.433	0.8279	2.833	360	5.5
143	61.5	0.4332	0.8279	2.833	360	5.5
144	64.5	0.4292	0.8279	2.833	360	5.5
145	62.5	0.4358	0.8279	2.833	360	5.5
146	62.5	0.4333	0.8279	2.684	360	5.5
147	63.5	0.4304	0.8281	2.684	360	5.5
148	62.5	0.4339	0.828	2.684	360	5.5
149	64.5	0.4349	0.8281	2.684	360	5.5
150	62.5	0.429	0.8281	2.684	360	5.5
151	59.5	0.8366	0.8288	2.45	360	3.6
152	57.5	0.8175	0.8279	2.45	360	3.6
153	59.5	0.8369	0.829	2.45	360	3.6
154	60.5	0.7826	0.8309	2.416	360	4
155	63.5	0.4256	0.8279	2.683	360	5.3
156	64.5	0.4253	0.8279	2.683	360	5.3
157	62.5	0.4248	0.828	2.683	360	5.3
158	63.5	0.4239	0.828	2.683	360	5.3
159	62.5	0.4288	0.828	2.683	360	5.3
160	62.5	0.4232	0.828	2.683	360	5.3

161	63.5	0.7402	0.8308	2.425	360	4
162	62.5	0.7486	0.8302	2.425	360	4
163	63.5	0.7486	0.8304	2.425	360	4
164	63.5	0.7247	0.8278	2.438	360	3.5
165	58.5	0.7827	0.8279	2.438	360	3.5
166	59.5	0.7348	0.8281	2.438	360	3.5
167	60.5	0.7308	0.8281	2.438	360	3.5
168	59.5	0.7318	0.828	2.438	360	3.5
169	59.5	0.7216	0.828	2.438	360	3.5
170	60.5	0.7274	0.828	2.438	360	3.5
171	59.5	0.7301	0.8281	2.438	360	3.5
172	59.5	0.7467	0.8287	2.452	360	3.8
173	61.5	0.748	0.8289	2.452	360	3.8
174	61.5	0.7364	0.8294	2.426	360	3.8
175	60.5	0.6568	0.8307	2.426	360	4
176	61.5	0.7345	0.8298	2.426	360	3.8
177	62.5	0.744	0.8291	2.426	360	3.8
178	61.5	0.7356	0.8298	2.426	360	3.8
179	61.5	0.7391	0.8284	2.438	360	3.8
180	62.5	0.7427	0.8289	2.438	360	3.8
181	58.5	0.7071	0.8289	2.438	360	3.8
182	61.5	0.7257	0.8286	2.438	360	3.8
183	61.5	0.7538	0.8286	2.438	360	3.8
184	64.5	0.64	0.8246	2.311	360	3
185	63.5	0.6511	0.8247	2.311	360	3
186	61.5	0.6424	0.8246	2.311	360	3
187	75.5	0.7658	0.8234	2.235	310	2
188	76.5	0.7803	0.8232	2.235	310	2
189	77.5	0.7822	0.8232	2.235	310	2
190	75.5	0.7958	0.8234	2.235	310	2
191	75.5	0.7858	0.8232	2.235	310	2
192	73.5	0.7249	0.8235	2.232	310	2.2

## C. POLYMATH REPORT SAMPLE (A) RELATIVE TO CHAPTER 4

**POLYMATH Report**  
Multiple linear regression

No Title  
06-Nov-2017

**Model:**  $FP = a1*CN + a2*DENSITY + a3*GROSS + a4*NET + a5*PNA + a6*ARO + a7*NAPHTHALENE + a8*SMOKE\ POINT + a9*ARO-JET$

Variable	Value
a1	3.153108
a2	503.7732
a3	-54.93183
a4	45.42566
a5	-0.0156549
a6	3.243207
a7	-6.422627
a8	-0.3260662
a9	-0.0687018

### General

Number of independent variables = 9  
Regression not including a free parameter  
Number of observations = 55

### Statistics

R <sup>2</sup>	0.966343
R <sup>2</sup> adj	0.9604896
Rmsd	0.0638076
Variance	0.2677397

### Source data points and calculated data points

	CN	DENSITY	GROSS	NET	PNA	ARO	NAPHTHALENE	SMOKE	ARO-JET	FP	FP calc	Delta FP
1	51.4	0.8185	45.507	43.063	3.7	18	4.1	14.2	17.1	56.5	56.97253	-0.4725328
2	51.8	0.8203	45.501	43.064	3.6	17.6	4.3	14	16.6	57.5	57.03491	0.465094
3	51.8	0.8188	45.484	43.042	3.7	17.8	4.2	13.6	16.8	57.5	57.61975	-0.1197475
4	51.7	0.8185	45.499	43.059	3.7	17.8	4.2	13.9	16.8	56.5	57.00374	-0.5037436
5	51.9	0.8211	45.516	43.079	3.6	17.8	4.3	14.4	16.8	58.5	58.11512	0.3848828
6	52.5	0.8234	45.476	43.051	3.9	17.9	4.8	13.4	16.7	58.5	59.53226	-1.032262
7	52	0.8223	45.465	43.034	3.9	18.1	4.6	13.2	17.1	58.5	59.20447	-0.7044709
8	51.7	0.8221	45.495	43.055	3.8	18.1	4.4	13.8	17.1	58.5	58.55422	-0.0542189
9	52.2	0.8221	45.477	43.02	4	18.2	4.7	12.8	17.1	58.5	58.25012	0.2498843
10	51.9	0.82	45.508	43.073	3.9	17.9	4.4	14.1	16.9	58.5	57.49618	1.003821
11	52.1	0.821	45.504	43.074	3.8	17.9	4.4	14.1	16.8	58.5	58.90416	-0.404162
12	52.5	0.8158	45.513	43.07	3.7	18	4.1	14.9	17.2	58.5	58.83404	-0.3340361
13	52.4	0.8207	45.465	43.039	3.8	17.6	4.5	12.9	16.4	58.5	59.05494	-0.5549413
14	52.4	0.8196	45.453	43.028	4.1	18.1	4.6	12.7	17	59.5	59.65893	-0.1589269
15	52	0.8228	45.508	43.073	3.8	18	4.5	14	16.9	59.5	58.93828	0.5617156
16	52	0.823	45.451	43.019	3.9	18.1	4.6	12.9	17.1	59.5	59.74259	-0.2425927
17	52.1	0.8221	45.472	43.044	3.9	18	4.6	13.1	16.9	59.5	59.21079	0.2892128
18	52.3	0.8183	45.45	43.025	4.1	18.2	4.5	12.6	17.1	59.5	59.70955	-0.2095493

19	52.5	0.8183	45.462	43.036	3.9	17.9	4.5	12.8	16.7	59.5	59.17311	0.3268924
20	52.7	0.8206	45.485	43.062	4.1	17.9	4.6	13.2	16.7	59.5	60.10422	-0.6042224
21	52.5	0.8236	45.475	43.054	4	18	4.9	13.5	16.9	59.5	59.45837	0.0416288
22	52.2	0.8235	45.441	43.012	4	18.1	4.7	12.6	17.1	59.5	60.31043	-0.8104313
23	52.2	0.8255	45.459	43.036	4	18.2	4.9	13.1	17.1	59.5	60.29618	-0.7961827
24	52.3	0.8201	45.453	43.027	4	18.2	4.7	12.9	17.1	59.5	59.16162	0.3383829
25	52.2	0.8239	45.449	43.01	4	18.2	4.7	12.7	17.1	59.5	60.27335	-0.7733487
26	52.9	0.822	45.486	43.06	3.9	17.7	4.6	13.2	16.5	60.5	60.66257	-0.1625733
27	52.3	0.822	45.463	43.035	4	18.1	4.6	12.8	16.9	60.5	60.29716	0.202838
28	52.8	0.8203	45.474	43.048	4.1	18.2	4.7	13.1	17	60.5	60.57939	-0.0793876
29	52.2	0.8206	45.462	43.033	4	18.2	4.5	12.9	17.1	60.5	60.16089	0.3391143
30	52.5	0.8198	45.463	43.033	4.1	18.4	4.6	12.9	17.2	60.5	60.64681	-0.1468107
31	52.5	0.8218	45.452	43.03	4.1	18.1	4.7	12.7	16.9	60.5	60.59293	-0.0929293
32	53	0.8174	45.498	43.077	4.2	18.1	4.6	13.5	16.8	60.5	59.94774	0.5522624
33	52.2	0.8242	45.449	43.021	4	18.2	4.8	12.8	17.1	60.5	60.24929	0.2507064
34	52.2	0.8225	45.454	43.026	4.1	18.2	4.6	12.7	17	60.5	60.66778	-0.167785
35	52.3	0.824	45.471	43.04	4.1	18.3	4.7	13	17.1	60.5	61.01824	-0.5182416
36	52.3	0.8221	45.455	43.029	3.9	17.9	4.6	13	16.7	60.5	59.81589	0.6841087
37	52.7	0.8216	45.482	43.053	3.9	17.8	4.6	13.1	16.6	60.5	60.08225	0.4177527
38	52.6	0.8211	45.477	43.043	4	18	4.5	12.9	16.9	61.5	60.66939	0.8306064
39	52.3	0.8211	45.454	43.016	3.9	18.1	4.4	12.7	17	61.5	60.78689	0.7131076
40	52.7	0.8269	45.445	43.023	4	17.8	4.9	12.6	16.7	61.5	61.64976	-0.1497624
41	53	0.8202	45.478	43.054	4.1	17.9	4.6	12.9	16.7	61.5	60.96758	0.5324171
42	52.9	0.82	45.467	43.042	4.1	18	4.6	12.7	16.7	61.5	61.00019	0.4998064
43	53.4	0.8197	45.46	43.04	4.3	18.2	4.7	12.2	16.7	62.5	62.88557	-0.385568
44	53.3	0.8202	45.46	43.036	4.3	18.2	4.7	12	16.8	62.5	62.69878	-0.1987842
45	52.7	0.8256	45.466	43.036	4	18	4.7	12.8	16.8	62.5	62.29291	0.2070944
46	53.1	0.8224	45.48	43.057	4.3	18.2	4.7	12.8	16.8	62.5	62.77091	-0.2709128
47	53	0.825	45.449	43.029	4.2	18.1	4.9	12.3	16.8	63.5	62.75213	0.7478669
48	54.3	0.8157	45.505	43.086	4.5	17.9	4.5	12.3	16.2	63.5	63.6361	-0.1360971
49	54	0.8186	45.501	43.075	4.4	18	4.5	12.3	16.4	63.5	64.1833	-0.6832978
50	53	0.8248	45.474	43.046	4.1	18.1	4.7	12.7	16.9	63.5	63.19911	0.3008869
51	52.5	0.8309	45.559	43.062	4.5	19.9	3.3	15.6	19.6	74.5	74.44528	0.0547194
52	52.2	0.8234	45.402	42.993	4.5	18.8	5.1	11.8	17.5	61.5	61.46605	0.0339534
53	52	0.8236	45.373	42.973	5	19.8	5.6	11.5	18.5	61.5	61.67387	-0.1738734
54	52.9	0.8223	45.439	43.026	4.5	18.6	5.1	12.1	17.2	62.5	61.85979	0.6402099
55	52.2	0.8286	45.389	42.984	4.6	19	5.4	11.4	17.7	63.5	63.22792	0.2720762

## D. POLYMATH REPORT SAMPLE (B) RELATIVE TO CHAPTER 4

**POLYMATH Report**  
Multiple linear regression

No Title  
29-Dec-2017

**Model:** FP = a1\*sulfur content + a2\*density + a3\*viscosity + a4\*end point + a5\*Distillation residue

Variable	Value
a1	-12.23854
a2	121.99
a3	21.43645
a4	-0.1727325
a5	-5.279361

### General

Number of independent variables = 5  
Regression not including a free parameter  
Number of observations = 109

### Statistics

R <sup>2</sup>	0.9797291
R <sup>2</sup> adj	0.9789495
Rmsd	0.0987736
Variance	1.114555

### Source data points and calculated data points

	Sulfur content	Density	Viscosity	End point	Distillation residue	FP	FP calc	Delta FP
1	0.4342	0.8276	2.685	360	5.5	62.5	61.98165	0.5183474
2	0.4275	0.8278	2.685	360	5.5	63.5	62.08805	1.411951
3	0.4408	0.828	2.685	360	5.5	62.5	61.94967	0.5503258
4	0.4605	0.828	2.685	360	5.5	61.5	61.70858	-0.2085751
5	0.4292	0.828	2.685	360	5.5	63.5	62.09164	1.408359
6	0.4419	0.8277	2.685	360	5.5	62.5	61.89961	0.6003852
7	0.4394	0.8277	2.685	360	5.5	61.5	61.93021	-0.4302112
8	0.4348	0.828	2.685	360	5.5	62.5	62.02311	0.4768946
9	0.4364	0.8281	2.685	360	5.5	63.5	62.01572	1.484277
10	0.4886	0.8277	2.863	360	5.5	64.5	65.14376	-0.643764
11	0.4486	0.828	2.863	360	5.5	64.5	65.6699	-1.169902
12	0.4375	0.8279	2.863	360	5.5	64.5	65.79355	-1.293551
13	0.4351	0.8279	2.863	360	5.5	64.5	65.82292	-1.322924
14	0.4378	0.8279	2.863	360	5.5	64.5	65.78988	-1.28988
15	0.4371	0.8279	2.863	360	5.5	64.5	65.79845	-1.298447
16	0.7371	0.8293	2.452	360	4	61.5	61.40633	0.0936699
17	0.7493	0.8295	2.452	360	4	61.5	61.28142	0.218582
18	0.7359	0.8296	2.452	360	4	61.5	61.45761	0.0423867
19	0.7639	0.8279	2.452	360	4	62.5	60.90755	1.592449
20	0.7525	0.8298	2.452	360	4	62.5	61.27885	1.221148
21	0.4357	0.828	2.685	360	5.5	63.5	62.01209	1.487909
22	0.4309	0.828	2.685	360	5.5	62.5	62.07084	0.4291643
23	0.427	0.8281	2.685	360	5.5	62.5	62.13077	0.369235

24	0.7372	0.8293	2.452	360	4	60.5	61.40511	-0.9051062
25	0.7565	0.8292	2.452	360	4	61.5	61.1567	0.3432965
26	0.4591	0.8248	2.768	360	2.5	79.5	78.95265	0.5473513
27	0.4531	0.8247	2.768	360	2.5	80.5	79.01388	1.486119
28	0.4518	0.8248	2.768	360	2.5	77.5	79.04199	-1.54199
29	0.7626	0.8233	2.246	310	2	74.5	75.14175	-0.6417452
30	0.7756	0.8232	2.246	310	2	75.5	74.97045	0.5295548
31	0.7759	0.8232	2.246	310	2	76.5	74.96677	1.533226
32	0.7741	0.8232	2.246	310	2	74.5	74.9888	-0.488803
33	0.7772	0.8231	2.246	310	2	73.5	74.93866	-1.438665
34	0.7821	0.8231	2.768	360	2.5	75.5	74.79222	0.7077816
35	0.4648	0.8247	2.768	360	2.5	79.5	78.87069	0.6293099
36	0.4561	0.8247	2.768	360	2.5	80.5	78.97717	1.522835
37	0.4562	0.8247	2.768	360	2.5	78.5	78.97594	-0.4759415
38	0.4492	0.8248	2.768	360	2.5	79.5	79.07381	0.4261898
39	0.4491	0.8248	2.768	360	2.5	80.2	79.07503	1.124966
40	0.4508	0.8247	2.768	360	2.5	78.5	79.04203	-0.5420296
41	0.4534	0.8248	2.768	360	2.5	80.5	79.02241	1.477592
42	0.776	0.8231	2.199	310	2	75.5	73.94584	1.554163
43	0.7776	0.8231	2.199	310	2	74.5	73.92626	0.5737442
44	0.4541	0.8249	2.778	360	2.5	79.5	79.2404	0.259595
45	0.4509	0.8249	2.778	360	2.5	79.5	79.27957	0.2204317
46	0.4514	0.8248	2.778	360	2.5	80.5	79.26125	1.23875
47	0.4527	0.8249	2.778	360	2.5	80.5	79.25754	1.242461
48	0.447	0.8248	2.778	360	2.5	77.5	79.3151	-1.8151
49	0.7764	0.8233	2.186	310	2	72.5	73.68667	-1.186666
50	0.7778	0.8233	2.186	310	2	74.5	73.66953	0.8304678
51	0.6993	0.8231	2.186	310	2	73.5	74.60586	-1.105859
52	0.7708	0.8231	2.186	310	2	71.5	73.7308	-2.230804
53	0.4676	0.8246	2.749	360	2.2	79.5	80.00074	-0.5007387
54	0.4343	0.8246	2.749	360	2.2	80.5	80.40828	0.091718
55	0.4523	0.8246	2.749	360	2.2	79.5	80.18799	-0.6879884
56	0.449	0.8264	2.749	360	2	80.5	81.50383	-1.00383
57	0.7155	0.8233	2.158	310	2	71.5	73.83177	-2.331772
58	0.7799	0.8232	2.158	310	2	74.5	73.03141	1.468588
59	0.7828	0.8232	2.158	310	2	73.5	72.99592	0.5040802
60	0.7843	0.8232	2.158	310	2	72.5	72.97756	-0.477562
61	0.7854	0.8234	2.158	310	2	74.5	72.9885	1.511502
62	0.4447	0.8247	2.749	360	2.2	79.5	80.2932	-0.7932002
63	0.455	0.8247	2.749	360	2.2	79.5	80.16714	-0.6671433
64	0.636	0.828	2.3333	360	3.5	61.5	62.58023	-1.080232
65	0.6389	0.829	2.3333	360	3.5	61.5	62.66673	-1.166731
66	0.639	0.8291	2.3333	360	3.5	62.5	62.67771	-0.1777058
67	0.6233	0.8266	2.36	360	3.5	61.5	63.13723	-1.637229
68	0.6702	0.8254	2.36	360	3.5	64.5	62.41685	2.083146
69	0.6377	0.8264	2.36	360	3.5	62.5	62.9366	-0.4365962
70	0.6485	0.8248	2.36	360	3.5	63.5	62.60924	0.8907641

71	0.6677	0.8254	2.36	360	3.5	64.5	62.44745	2.05255
72	0.62	0.8239	2.36	360	3.5	63.5	62.84824	0.6517568
73	0.5289	0.8312	2.42	360	3.8	62.5	64.55608	-2.05608
74	0.6386	0.8309	2.42	360	3.8	62.5	63.17692	-0.6769156
75	0.737	0.8302	2.454	360	3.6	63.5	63.67196	-0.1719622
76	0.7409	0.8287	2.454	360	3.6	62.5	63.44125	-0.9412468
77	0.7493	0.8291	2.437	360	4	59.5	60.91108	-1.411075
78	0.7458	0.8291	2.437	360	4	60.5	60.95391	-0.45391
79	0.433	0.8279	2.833	360	5.5	63.5	65.20553	-1.705531
80	0.4292	0.8279	2.833	360	5.5	64.5	65.25204	-0.7520375
81	0.4333	0.8279	2.684	360	5.5	62.5	62.00783	0.4921722
82	0.4304	0.8281	2.684	360	5.5	63.5	62.06772	1.432282
83	0.4339	0.828	2.684	360	5.5	62.5	62.01268	0.4873163
84	0.429	0.8281	2.684	360	5.5	62.5	62.08485	0.4151485
85	0.7826	0.8309	2.416	360	4	60.5	60.27295	0.2270516
86	0.4256	0.8279	2.683	360	5.3	63.5	63.1365	0.3634998
87	0.4253	0.8279	2.683	360	5.3	64.5	63.14017	1.359828
88	0.4248	0.828	2.683	360	5.3	62.5	63.15849	-0.65849
89	0.4239	0.828	2.683	360	5.3	63.5	63.1695	0.3304953
90	0.4288	0.828	2.683	360	5.3	62.5	63.10954	-0.6095359
91	0.4232	0.828	2.683	360	5.3	62.5	63.17807	-0.6780717
92	0.7486	0.8302	2.425	360	4	62.5	60.79659	1.703406
93	0.7247	0.8278	2.438	360	3.5	63.5	63.71467	-0.2146729
94	0.748	0.8289	2.452	360	3.8	61.5	62.28001	-0.7800062
95	0.7364	0.8294	2.426	360	3.8	61.5	61.92562	-0.4256204
96	0.6568	0.8307	2.426	360	4	60.5	62.00252	-1.502523
97	0.7345	0.8298	2.426	360	3.8	61.5	61.99767	-0.4976696
98	0.744	0.8291	2.426	360	3.8	62.5	61.79601	0.7039895
99	0.7356	0.8298	2.426	360	3.8	61.5	61.98421	-0.4842072
100	0.7391	0.8284	2.438	360	3.8	61.5	62.02782	-0.5278238
101	0.7427	0.8289	2.438	360	3.8	62.5	62.04476	0.4552399
102	0.7257	0.8286	2.438	360	3.8	61.5	62.21622	-0.7162182
103	0.7538	0.8286	2.438	360	3.8	61.5	61.87232	-0.3723153
104	0.64	0.8246	2.311	360	3	64.5	64.27816	0.2218404
105	0.6511	0.8247	2.311	360	3	63.5	64.15451	-0.6545108
106	0.7658	0.8234	2.235	310	2	75.5	74.87898	0.6210201
107	0.7958	0.8234	2.235	310	2	75.5	74.51182	0.9881762
108	0.7858	0.8232	2.235	310	2	75.5	74.60981	0.8901888
109	0.7249	0.8235	2.232	310	2.2	73.5	74.27155	-0.7715535