NOVEMBER 2017

M.Sc. in Industrial Engineering

AL MOTHANA AL SHAREEF

UNIVERSITY OF GAZIANTEP

GRADUATE SCHOOL OF

NATURAL & APPLIED SCIENCES

EMPIRICAL MODELLING, CONTROL AND OPTIMIZATION OF LUBRICANT RECYCLING UNIT

M. Sc. THESIS

IN

INDUSTRIAL ENGINEERING

BY

AL MOTHANA AL SHAREEF

NOVEMBER 2017

Empirical Modelling, Control and Optimization of Lubricant Recycling Unit

M.Sc. Thesis

in

Industrial Engineering

University of Gaziantep

Supervisor

Prof. Dr. Serap ULUSAM SEÇKINER

by

Al Mothana AL SHAREEF

November 2017

REPUBLIC OF TURKEY
UNIVERSITY OF GAZIANTEP
GRADUATE SCHOOL OF NATURAL & APPLIED SCIENCES
INDUSTRIAL ENGINEERING DEPARTMENT

Name of the thesis: Empirical Modelling, Control, and Optimization of Lubricant Recycling Unit
Name of the student: Al Mothana AL SHAREEF

Exam date: 01/11/2017

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Ahmet Necmeddin YAZICI
Director

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

Prof. Dr. Serap ULUSAM SEÇKİNER
Head of Department

This is to certify that we have read this thesis and that in our consensus opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Serap ULUSAM SEÇKİNER
Supervisor

Examining Committee Members	Signature
Prof. Dr. Serap SEÇKİNER	.....
Asst. Prof. Dr. Eren ÖZCEYLAN	.....
Asst. Prof. Dr. Yunus EROĞLU	.....

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original of this work.

Al Mothana AL SHAREEF

ABSTRACT

EMPIRICAL MODELLING, CONTROL, AND OPTMIZATION OF LUBRICANT RECYCLING UNIT

AL SHAREEF, Al Mothana
M.Sc. in Industrial Engineering
Supervisor: Prof. Dr. Serap ULUSAM SEÇKİNER
November 2017
111 pages

This study consists of two phases. In the first phase, four different forecasting approaches were applied to the chemical processes found in the lubrican recycling unit. The predictive methods used are Multiple Linear Regression, Multiple Nonlinear Regression, Artificial Neural Networks and Ridge Regression. The input variables of models are temperature and pressure, and the output variables of the unit are quantity and quality. Both quantities and qualities of useful end products are classified as light gas oil, medium gas oil, and heavy gas oil and are kept in the daily record of the chemical unit. Four different forecasting models have been compared in term of their performance. The artificial neural network approach showed best predictive model with the least error rate. In the second phase of the study, the non-linear mathematical optimization model generated the optimal temperature and pressure values in order to maximize the quantity and quality of light, medium, and heavy gas oils. Thus, optimum parameters have been obtained to help eliminate at least 4300 litres of wasted oil per week and production has been increased from 57% to over 70%. As a result, it has been observed that controlling the chemical processes by nonlinear mathematical modelling increases the production quantity.

Key Words: Empirical modelling, Multiple Linear Regression, Multiple Nonlinear Regression, Artificial Neural Networks, Ridge Regression.

ÖZET

KARTER YAĞ GERİ DÖNÜŞÜM SİSTEMİ İÇİN AMPİRİK MODELLEME, KONTROL VE OPTİMİZASYON

AL SHAREEF, Al Mothana
Yüksek Lisans Tezi, Endüstri Mühendisliği Bölümü
Tez Yöneticisi: Prof. Dr. Serap ULUSAM SEÇKİNER
Kasım 2017
111 sayfa

Bu çalışma, iki aşamadan oluşmaktadır. İlk aşamada, karter yağ geri dönüşüm sisteminde bulunan kimyasal işlemler üzerine dört farklı öngörü yaklaşımı uygulanmıştır. Kullanılan öngörüsel yöntemler, Çoklu Doğrusal Regresyon, Çoklu Doğrusal Olmayan Regresyon, Yapay Sinir Ağları ve Ridge Regresyonu'dur. Modellerin girdi değişkenleri sıcaklık ve basınç, sistemin çıktı değişkenleri ise miktar ve kalitedir. Yağların hem miktarı hem de kalitesi, hafif yağ, orta yağ ve ağır yağ olarak sınıflandırılarak kimyasal birimin günlük kaydında tutulmaktadır. Dört farklı öngörü modeli performansları açısından karşılaştırılmıştır. Yapay sinir ağları yaklaşımı, en az hata oranı ile en iyi öngörü modeli olmuştur. Çalışmanın ikinci aşamasında, doğrusal olmayan matematiksel optimizasyon modeli ile geri dönüşüm sonrasında ortaya çıkacak hafif yağ, orta yağ ve ağır yağların miktar ve kalitesini maksimize edecek sıcaklık ve basınç değerleri bulunmuştur. Böylece, haftada en az 4300 litre atığın azaltılmasına yardımcı olmak için optimum parametreler elde edilmiş ve üretim % 57'den % 70'in üzerine çıkarılmıştır. Sonuç olarak, doğrusal olmayan matematiksel modelleme ile kimyasal süreçlerin kontrol edilmesinin üretim değerini arttırdığı gözlenmiştir.

Anahtar Kelimeler: Ampirik modelleme, Çoklu Doğrusal Regresyon, Çoklu Doğrusal Olmayan Regresyon, Yapay Sinir Ağları, Ridge Regresyonu.

With the deepest gratitude and warmest affection, I would love to dedicate this work to my beloved mother; Safa'a Wazzan, my great father; Abdul Salam Al Shareef and to the rest of the family.

In memorial of my late uncle; Osama Al Shareef and his family, and to the victims of Syria, a credit is paid to them all!

ACKNOWLEDGEMENTS

Accomplishing this work was not easy and has become a source of pride to me, to my family, and to the beloved ones. I, therefore, acceptingly owe those who readly, with all ease, lent me their hand, and heart and were part of the contribution to the success of this work.

First and foremost, I thank the almighty Allah for His guidance all over the academic and dessertional trip.

I, then, would like to pleasedly thank my dessertation supervisor Prof. Dr. Serap SEÇKİNER for her nice follow-ups and valuable advices during this work. She remained a guide to me in every aspect of the work and a heartfelt gratitude and indebtness is paid to her hereafter.

I would definetly not forget to thank my company where I work; KAYALAR ENERGY/Gaziantep-Turkey; for facilitating me to conduct this work on its newly constructed chemical system with full privileges and references.

Finally, a special thank to all teachers and colleagues at my department without whom I wouldn't have arrived at this achievement..Thank you all!

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS/ABREVIATIONS

AI	Artificial intelligence
ANN	Artificial neural network
B	Regression coefficient
b_k	Neuron bias
e	Error of residuals
EM	Empirical modelling
EPA	Environmental Protection Agency
HGO	High gas oil
LGO	Light gas oil
LSM	Least square method
MAPE	Mean absolute percentage error
MGO	Moderate gas oil
MLR	Multiple linear regression
MNLR	Multiple nonlinear regression
MSE	Mean square error
NORA	National Oil Recycling Association
PA	Predictive analytics
$\varphi_{(.)}$	Activation function
R	Independent value of correlation
RR	Ridge regression
T	Temperature
V	Volume
v_k	Induced local field
WO	Waste oil
$\omega_{(km)}$	Synaptic weigh

CHAPTER 1

INTRODUCTION

1.1 General Introduction

Nearly all processes and reactions existing in any chemical plant show complexity in their behaviour while they exhibit nonlinear and no equilibrium state of work. This is because of certain phenomena that happens naturally and undergo physical and chemical laws.

In chemical industries, stakeholders seek to continuously increase productivity and profitability of the chemical processes to survive in the market. To achieve this goal, it is crucial to ensure chemical plants are run with optimal efficiency despite of numerous constraints influencing their processes.

To enhance productivity of a process, design and safety with minimal capital and operating cost, it is vital to understand the dynamical characteristics or performance of the processes in any chemical plant. The process characteristics can be realized through implementation of predictive analysis by building up predictive models either based on first principle or empirical approaches. However, empirical approach is preferred in this work over the fundamental or first-principle approach for their efficacy and easiness where these models should correctly capture the exact characteristics of a system. However, empirical modelling approaches are many and vary in term of their prediction capability.

The term “predictive analytics” is the heart of every modelling operation that is intended to forecast upcoming events of a system. Predictive analytics helps reveal unseen relationships in the information and data that can help one predict with greater assurance what may happen in the future, and deliver you with valued perceptions for your organization (Bari et al., 2017). However, Siegel (2016) defined predictive analytics as “technology that leans from experiential data to forecast the imminent or future behaviour of entities so as to drive better decisions. In short, predictive analytics is the learning, understanding and use of information (data) to predict future outcomes.

In simple words, predictive analytics is to learn when to lie, buy, or click through prediction ability according to Siegel (2016).

Predictive analytics has been growing in popularity in recent years. However, early attempts to prediction took place during the World War II. The story starts with the father of so called "cybernetics"; Norbert Wiener. He started to determine the performance of a plane pilot who belongs to the German airplane fleet in 1940. His aim was to shoot all enemy warplanes from above. His methodology was to consider the trajectory parameter as the input parameter and predicts where the plane would be located in order to hit it with machinegun shells. Unfortunately, Wiener couldn't synchronize with the exact time and only could predict a single second after a plane's movement. Nonetheless twenty seconds of forthcoming trajectory were necessary to shoot down a plane.

Numerous number of applications related to predictive analytics have been found involving building forecasting models which are helpful in every field of human activity. For example, in the scope of mechanical engineering, detecting maintenance time when trained worker shortage is noticed is helpful to monitor heat exchanger performance. In chemical process industry, forecasting need is more efficiently sought by applying statistical analysis. Governments monitor mortgage insurance using fully or partially (semi-automated) systems and organize raw data form multiple sources. In the area of human resources management, predictive labour force analytics has been used extensively in order to obtain different multivariate regression models to analyse why employee repeatedly turnover. Predictive analytics has been also useful in generating procedures to predict industries, training, controlling production schedules, and detecting risk at certain areas. Predictive analytics has been introduced scholastically to monitor issues encountered in the health sector. Moreover, predictive models have applied regression and functional estimate to more precisely predicted clinical patient stay period (Olson and Wu, 2017).

Predictive analytics covers a diversity of statistical tools starting from machine learning, predictive modelling, and data mining that studies current and historical truths to make estimates about future or otherwise unknown events. However, machine learning which include algorithmic tool like neural networks, regression analysis, and regularization algorithm tools are among the most popular techniques that are followed for construction of the predictive models.

1.2 Research Background (Literature Review)

Chemical process control and optimization always pose a great deal of tedious and (sometimes impossible) work because it necessitates the need for modelling these processes. However, this task won't be easy due to the nature of the chemical processes. As stated before, chemical processes are complex in nature, nonlinear and interactive. In addition, traditional control approaches do not handle nonlinearities issues, interaction, and the multiple process constraints very well. It typically reacts to process disturbances after they have already occurred. Despite of this fact, chemical processes traditionally used to be controlled by applying linear system and tools. The basic reason behind using linear system theory is being an analytical solution and more rigorous and stable performance (Bequette, 1991).

However, in spite of the vital importance of process nonlinearities, a number of papers and articles discussed about nonlinear systems techniques in the early 1980s. Ray (1983) explored the importance of multivariable process control in the nonlinear systems as they play a minor role. This "disregard" was to the extent that the first three engineering foundation conferences on chemical process control the only papers that focused on nonlinearities were Ray (1982), Shinnar (1986), and Morshedi (1986).

In our research, chemical processes are manipulated based on nonlinear fashion considering the dynamic and unsteady-state behaviour of the system.

In general, there are two main types of model; fundamental and empirical model. The fundamental model; also known as first-principle or parametric model; is built on concept of physical-chemical relationships. Actually, these models are obtained by applying conservation principle, reaction kinetics, transport phenomena, and thermodynamic (e.g., phase equilibrium) relationships. On the one hand, fundamental models offer some potential disadvantages that they are time consuming to develop and they often have a large number of equations with many parameters that need to be estimated. On the other hand, the empirical model has certain advantages over the fundamental one. The empirical model is generally developed to use when the actual process is too complex and the underlying phenomena are not well understood or when the numerical solution of the fundamental model is quite difficult or when the empirical model provides satisfactory predictions of the process characteristics. Experimental plant data are used to develop a relationship between the process input

and process output as an empirical model using a mathematical framework such as artificial network (Rojas, 1996), least square method (LSM) (Jain et al., 1995), etc.

In this research, empirical modelling approach is considered and different predictive models are generated based on the historical dataset collected from the lube oil unit. Lube oil recycling unit; as a case study; is a typical chemical process plant as it was designed, manufactured, and tested personally. The system's input (independent) variables are temperature and pressure, which can stimulate the process and can bring internal change in process conditions. The system's output (dependent) variables are the final product of the distillation process; mainly are light, moderate and heavy gas oil. These types of variables cannot be manipulated but can represent the internal state of the process.

Four well-known empirical approached were followed to construct predictive models for the lube oil recycling unit: (1) Multiple Linear Regression (MLR), (2) Multiple Nonlinear Regression (MNLR), (3) Artificial Neural Network (ANN) and (4) Ridge Regression (RR) where each hold discrete properties and has pros and cons. In our research, we compared the prediction performance of each technique against some performance metrics. Meanwhile, we found out the optimal parameters of the system's process through construction of solved mathematical model.

There are many recent studies which have been conducted in the field of process modelling and optimization. In these studies, predictive models were generated and tested for the purpose of predicting possible future events of the chemical processes.

Huyck B (2013) has contributed in the field of model predictive control in the chemical process industry hosted by industrial controllers. Huyck employed the recent online quadratic optimization problem (QP) in order to minimize the use of computational power and memory to solve the linear model predictive controllers (MPC). Two practical set-ups were used, i.e., a mini set-up consisting of a fan and heating resistor, and a pilot-scale distillation column set-up. On the mini set-up, he proved that a programmable logic controller (PLC) is able to solve the quadratic optimization problem accompanied by a model predictive controller. Nazario D, Ramirez B and Jackson H (1999) used artificial neural network to control the pH of the erythromycin acetate salt. Experiments were mainly conducted to determine the time delay of chemical reactions. The suggested methodology includes three main steps: (1) the cross-correlation function is used to detect time delay, (2) a feedforward neural network is used to model the input and output variables of a nonlinear dynamic

process, and (3) an optimization technique is used to solve the control equation and implement the corrective action. The selected neural network algorithm works as an adaptive procedure. The implemented algorithm reads the last 60 observations from four variables to generate a recommendation for controlling the pH of the erythromycin acetate salt. Dellana and West (2009) compared the multi-period predictive ability of linear ARIMA models to nonlinear time delay neural network models in water quality applications. Comparisons are made for a variety of artificially generated nonlinear ARIMA data sets that simulate the characteristics of wastewater process variables and watershed variables, as well as two real-world wastewater data sets. While the time delay neural network model was more accurate for the two real-world wastewater data sets, the neural networks were not always more accurate than linear ARIMA for the artificial nonlinear data sets. In some cases of the artificial nonlinear data, where multi-period predictions are made, the linear ARIMA model provides a more accurate result than the time delay neural network. This study suggests that researchers and practitioners should carefully consider the nature and intended use of water quality data if choosing between neural networks and other statistical methods for wastewater process control or watershed environmental quality management.

1.3 Problem Statement

Chemical processes, unlike of discrete processes, are very nonlinear and complicated in nature, and they form the continuous production system. They include numerous and various nonlinear parameters which describe the system's dynamical behaviours. However, modelling the chemical processes is uneasy task that requires full understanding of chemical reactions and processes existing in the intended system, as well as the physical properties and characteristics of the system, such as reactor temperature, pumps vacuum, pressure vessels, and so on.

Chemical process control and optimization are characterized by a broad invasion of distributed control systems into chemical plants. The integration from process control up to business management is a great challenge of today which follows from the overall computerization of production. However, without process control, it would not be possible to operate most modern processing facilities safely, reliably, and profitably, while satisfying quality standards (Jose A. Romagnoli and Ahmet Palazoglu, 2006). Moreover, the art of design, optimize, and control of chemical processes should be considered simultaneously (Segovia and Petriciolet, 2016).

In today's competitive world it is necessary to have the highest-quality product at the lowest cost. In addition, products must be safe and environmentally friendly. Almost, all manufacturing processes with different level of superiority, can be enhanced in order to reduce cost and enhance product quality (Girish K. Malhotra, 2011). Complex chemical plants require continuous monitoring and control of their processes in order to reduce waste, and improve productivity and quality. As part of the growing interest in waste recycling, different treatments have been explored with the aim of recuperating both the energetic and chemical value of the waste oil (WO).

In the last decades, the process of recycling waste lube oil where after recycled oil is utilized as a fuel has become increasingly an important issue for the researchers and technologists. Most of the lube oils are commonly attained from petroleum resources. The waste oil quantities can be refined and treated to produce fuels or lubricating oil. On the other hand, the waste oils expose potential threat to the environment because of their metallic contents and other dangerous contaminants.

In 2008, the US Environmental Protection Agency (EPA) estimated that Americans purchase 2.5 billion gallons of lube oil and produce 1.37 billion gallons of used oil each year. However, the National Oil Recycling Association (NORA) stated that the 1.37 billion gallons of used oil is formed from 856 million gallons from transport means and 521 million gallons from manufacturing and industrial usage.

In our research, chemical processes in the lube oil recycling unit are control manually with controllable parameters: temperature and vacuum, which are adjusted in accordance to the quality and quantity of the final products (mainly light, moderate and heavy gas oil).

After laborious research work and experimental tests on the lube oil recycling unit, we could determine the system's production capacity to be maximally 57% of the total treated used oil. This percentage is considered modest, yet this system is deemed less economic and lees friendly to the environment with waste of over 40%. Subsequently, this amount of waste is generally dumped beneath the ground. Moreover, control and optimization of the chemical process requires the establishment of predictive models that can estimate coming situations and events which, in return, will enable us to take preventive or correction action beforehand.

In the proposed system, it was noticed that there is weekly loss of around 4300 litres of gas oil that constitutes 3% of the final products generated by the system. The reason behind this loss is because processes are not yet optimized. That is to say, they are

uncontrollable and optimal parameters were still not determined yet. Therefore, the design and modelling of chemical processes is currently driven by a large number of requirements posed by energy and material costs, and the demand for robust, fault-tolerant systems. These concepts induce persistent need for effective process modelling tools and techniques. Interestingly, prior to the 1940s, most chemical processing plants were essentially run manually, as the plant operators adjusted material and energy flows by manipulating rather large valves by hand (Jose A. Romagnoli and Ahmet Palazoglu, 2006). Nevertheless, in the last decade, interest in integrating computer modelling and process simulation in the design of chemical processes has grown exponentially due to the apparent economic, environmental, and operational benefits.

1.4 Research Objectives

This work aims to:

1. Implement four different empirical approaches and generate four relevant predictive models on the processes of the lube oil recycling unit.
2. Compare the models' performances
3. Control system's processes.
4. Optimize system's processes by:

    a. Maximizing useful final product quantity above 70% (maximizing quantity)
    b. Minimizing waste and pollution [energy + by-products (till 30%)]
    c. Standardizing final distillates' compositions (quality)

1.5 Scope of the Research

This study focused on the introduction of different predictive modelling techniques which can be deployed in the modelling of chemical processes. Particularly, the study focused on four basic empirical approaches. These approaches are the most applicable tools in the modelling of chemical processes: Multiple Linear Regression (MLR), Multiple Nonlinear Regression (MNLR), Artificial Neural Network (ANN), and Ridge Regression (RR). Accordingly, with these approaches four different predictive models were generated and were compared based on their predictive performance and then, used to control system's processes. Finally, Mathematical model was also constructed for the purpose of optimizing the system's processes.

1.6 Outline of Research

The research's thesis is divided into five chapters and us outlined according to the following:

Chapter 1 introduces the gist of the research including a concise literature review on the work. It also includes the problem statement, research objectives, and the scope, along with a reference section and appendices. The significance of this research with regard to the importance of implementing empirical modelling on chemical processes was also elaborated in this chapter.

Chapter 2 provides a research on the techniques concerning the elements of the research, so as to better understand the issues of the research, starting with the introduction of the chemical process in general, the lube oil recycling processes and their systems and then introducing concepts and techniques used in the work like Empirical Modelling, Machine Learning, Multiple Linear Regression, Multiple Nonlinear Regression, Artificial Neural Networks, Ridge Regression, mathematical modelling and etc.

Chapter 3 explains the research methodology including description of the real-life chemical unit and the approaches that have been adopted in this work. The work was planned and subdivided into four phases as: phase (1): Data Collection and Processing, phase (2): Construction of Predictive Models, phase (3): Mathematical Modelling, Phase (4): Obtaining Optimized Parameters. However, each phase was allocated with its own techniques, tools and software as well.

Chapter 4 introduces all results obtained and discusses them in detail. Readings, graphs and tables for each methodological phase is of course described in chapter 3.

Chapter 5 draws conclusions in the light of results obtained in chapter 4. The chapter also provides recommendations as a guidance for future work. The thesis structure is illustrated in Figure (1.1).

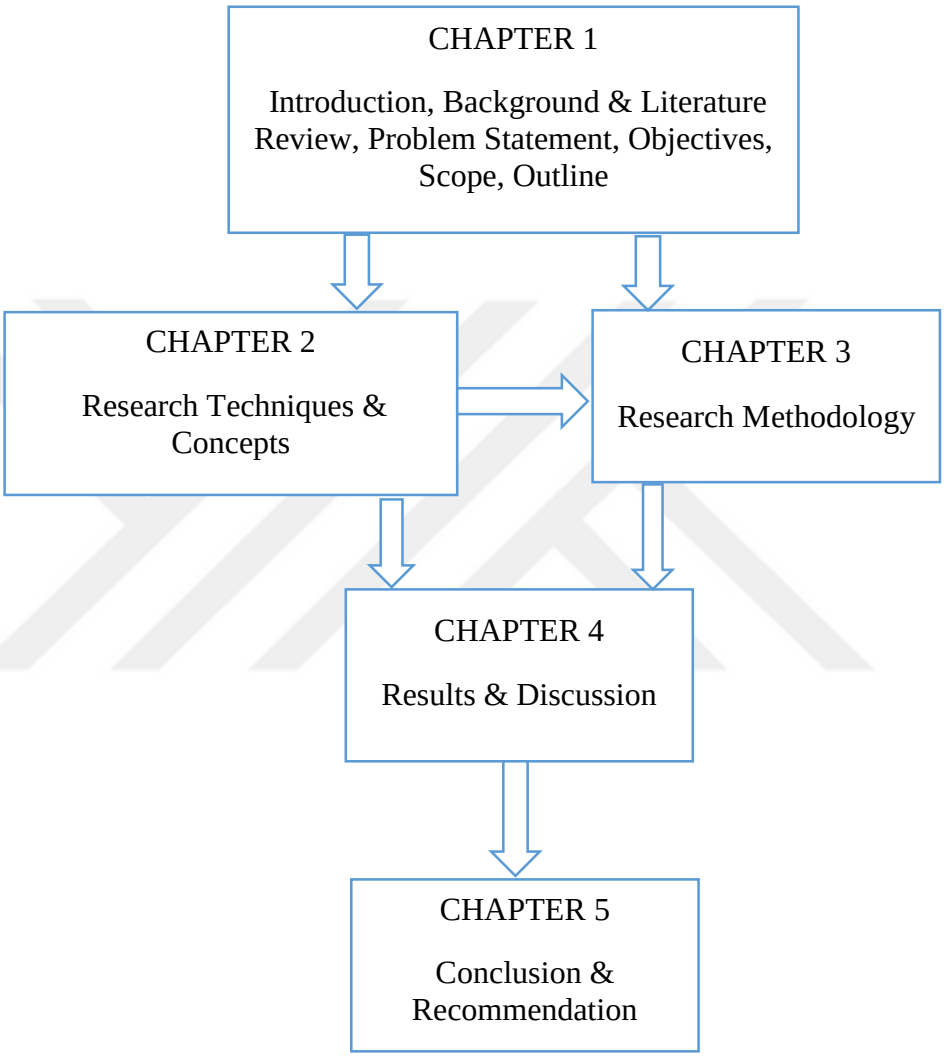


Figure 1.1 Thesis structure

CHAPTER 2

RESEARCH TECHNIQUES

2.1 Introduction

This chapter reviews the academic research techniques and tools of the empirical modelling techniques. However, the chapter starts with the introduction of chemical processes and their nature. It later explains a typical unit of chemical process plant called the lubricant recycling unit and the associated processes involved within it as far as our case study is concern. Next, the chapter describes the process modelling and some related concepts. Then, it compares the empirical modelling to the fundamental modelling showing all pros and cons. Finally, the chapter introduces in detail the four proposed empirical techniques deployed in our case study. These are the Artificial Neural Network (ANN), Multiple Linear Regression (MLR), Multiple Nonlinear Regression (MNLR) and Ridge Regression (RR) tools respectively.

2.2 Chemical Processes and Modelling

Most chemical processes are complex in nature as they exhibit non-linear and non-equilibrium behaviour. This complexity is because of instantaneous and often coupled momentum-, heat-and mass-transfer phenomena and kinetic processes taking place at different scales. Nonlinear dynamics can be defined as a nonlinear-behaviour system with key variables like concentrations in a chemical reaction, evolve in time. Nearly all systems of interest, including living ones, are nonlinear, often extremely so, but scientists often find it convenient, e.g., in the case of relaxation kinetics, to utilize conditions where the system under consideration behaves linearly. Linear mathematics is accustomed and manageable, but it cannot generate even a small portion of the rich phenomenology of which the “simplest” of nonlinear systems is capable. In any chemical system processes, the phenomena of nonlinearities naturally rise from determined equations of mass kinetics; if it undergoes unusual circumstances, any bimolecular basic steps lead to quadratic form. Considering an example of relaxation kinetics behaviour which is explained as: if a unit or any system is near to steadiness

or equilibrium, the nonlinear effects may be negligible. However, equilibrium is not approachable if when the chemical system remains under the interest of the study of nonlinear dynamics.

2.2.1 Lubricant Recycling Processes

Lubricant oil or in short form (lube oil) is an automotive oil that can be defined as an engine or any lubricating or greasing oil that has been used over time. Different kinds of lubes are derived from the petroleum distillation process under certain vacuumed environment. They are physically viscous and temperature absorbent agents which help for lubricating moving parts inside engines and machines e.g. engine crankcase, brakes, power steering, hydraulic machines, transmission parts, heavy machine tools, refrigerators, etc. Lube oil gets unfitted for additional usage for two reasons: accumulation or gathering of physical contaminants or impurities in the oil e.g. fuel, lead, water, abrasives, organic acids, etc. Lube oil as other petroleum products is obtained after distillation process takes place in the vacuum column. Lube oil is more viscous than most of the other petroleum products and can be produced by fractioning crude oil by a high boiling point (>400°C, >750°F) and, in fact, lubricating oil is identified by viscosity.

2.2.1.1 Lubricant Recycling Processes Main Steps:

Before recycling process begins, all solid materials, degraded additives and dissolved metals are removed.

Step1: Dehydration

The oil is gradually heated to 120C^{o} and boiled in an elliptical vessel to extract the water that has been mixed into it.

Step2: Diesel stripping

The dry or dehydrate is fractionated in the vacuum distillation column in the same way that the crude oil does. The fractions occur as follows:

1. Heavy and light fuel oil including diesel. Local oil provides enough diesel extracted from the used oil in order to use it in energy plants like burners and boilers where they run efficiently.

2. Lubricating or lube oil. The majority of the feed will refine in the refinery to generate lube oil portion.

3. Residue or remainder parts are the distillates which was not fractionated yet in the feeds. This encompasses all the carbon products, wears from different reacting to environment items like metal and most of the lead and oxidation products. Figure (2.1) Shows the general schematic diagram of the lube oil production unit.

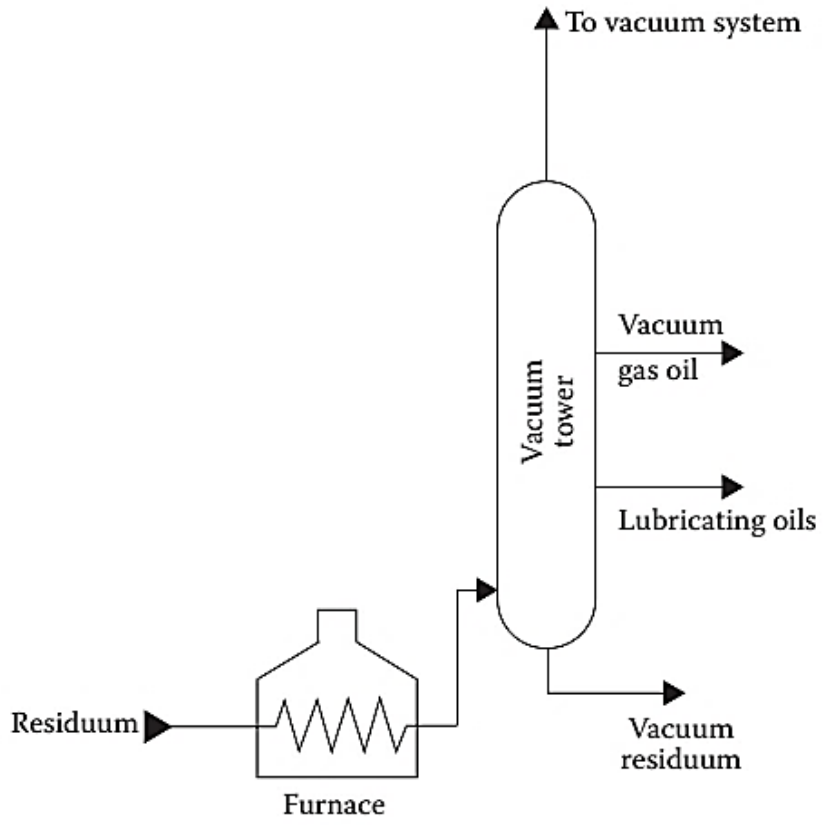


Figure 2.1 General schematic diagram of the lube oil production unit (Source: OSHA, 2006)

Step3: Lube oil distillation and condensation

The used oil is pushed into a filter and then followed to a polisher, where used oil is polished till all particles that are larger than one-micron diameter is removed. The base oil that is produced, however, varies from the original one in respect to its chemical and physical properties. Figure (2.2) shows the main steps of the lube oil production processes.

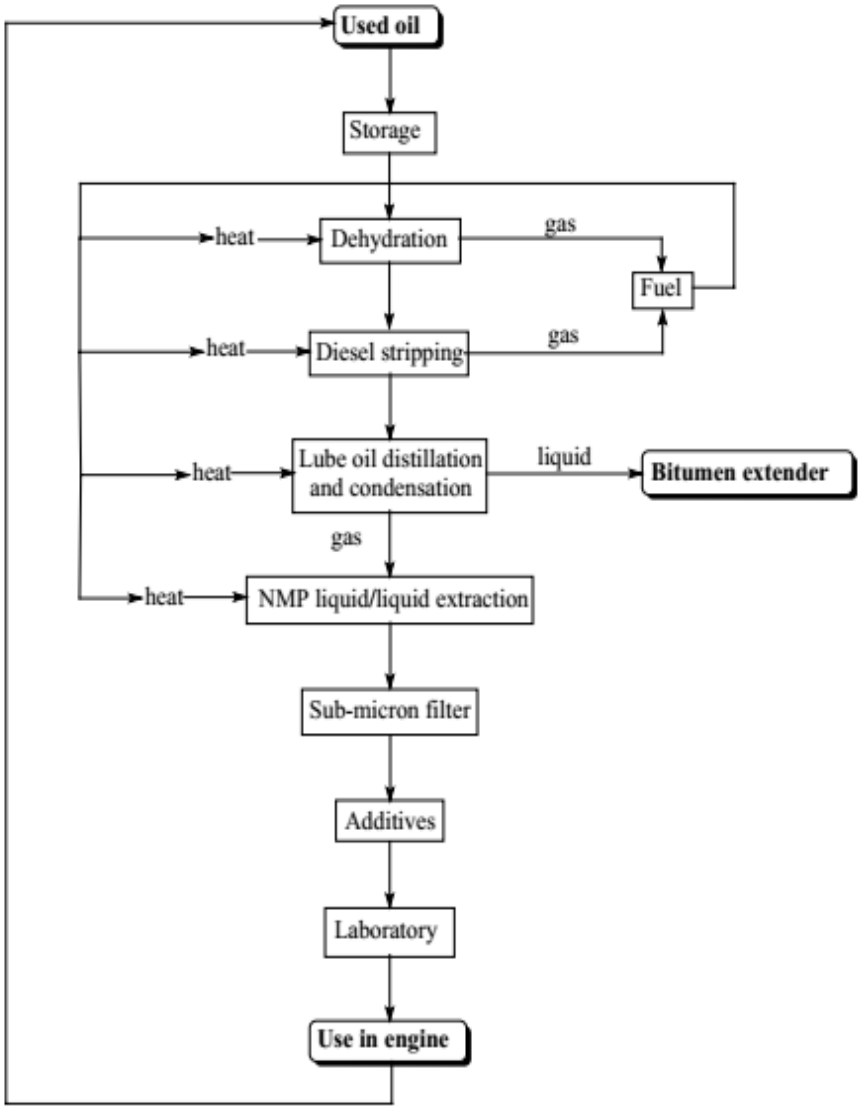


Figure 2.2 Main steps of the lube oil production processes (Source: International Journal of Scientific & Engineering Research, 2015)

2.2.2 Why Lube Oil is Harmful to the Environment?

The lube oil contains mainly S, Ca, Zn and P chemical elements. However, after lube oil gets polluted (contaminated) or "waste", it consists of several harmful elements and parts generated form the wearing off processes of machine and engine parts like lead, zinc, copper, steel and iron particles, sulphur and other toxic metals. In the year 2008, the American Environmental Protection Agency (EPA) which is responsible for environmental protection issues, reported that in a naval search that Americans are thought of buying over 2.5 billion gallons of lube oil. Consequently, they dispose 1.37

billion gallons of used oil each year. When waste oil dumped or casted onto the rivers or under the ground, it causes serious environmental pollution. Unfortunately, tons of waste oil are discharged in the water nevertheless, it is strictly illegal and prohibited. Used lube oils pose real threats to the environment when directly disposed to the nature rather than proper treatment i.e. one gallon of used engine oil casted under waterways brings a potential threat to the environment and can pollute one million gallons of water.

2.2.3 Importance of Lube Oil Recycling

In recent years, reprocessing of the waste lube oils and exploiting off the products as fuels have become vital topics for researchers. The greatest amount of the lube oils is mostly obtained from petroleum resources. Waste oil can be treated to produce fuels or lubricating oil. However, the waste oil poses environmental danger and threat because of both their metal content and other contaminants. The high quantity of waste oils can be converted into usable final products like fuel or diesel throughout refining and treating processes. Converting waste oil into diesel-like product to be utilized as a fuel for vehicle engines without harmful by-product is very crucial. Utilization of such diesel-like oils, and unification of the produced fuels helps decrease in fuel consumption, protecting environment from lethal and harmful chemicals. It also protects foreign exchange, reduces saves of foreign exchange and reduces greenhouse gas emissions.

It is well understood that one of the properties acquired by any chemical processes is the presence of nonlinearity which incur a challenging in controlling of these processes. Despite of this fact, chemical processes used to be controlled conventionally by adopting linear system analysis. Also, the computational demands for linear system modelling and simulation (and implementation) are usually quite small when compared to a nonlinear modelling and simulation. Obviously, the usage of any linear system method or technique is far unsuitable and limiting for chemical processes because they are highly nonlinear. The advancement experienced in the area of nonlinear control theory, associated with computer hardware technological progression; now allow advanced, nonlinear control strategies to be successfully implemented on chemical processes. As part of the growing interest in lube oil recycling, alternative way of treatments has been tried with the aim of improving chemical value of the WO.

2.2.4 Usage and Benefits of Recycled Lubricant

Waste oil exposes great threat to the environment and economic of the state. However, recycling of used oil again and again can will definitely help in the preserving the stability of the eco system and reduces pollution. Oil recycling is part of energy recovery which increases option for energy resources. The energy saved by recycling of lube oil can also help reduce dependence on foreign oil imports.

Recycled lube oil can be distilled and used as diesel or marine fuel in process similar to crude oil refining. Industrial heating application is another area wherein recycled oil can be used widely in most of the chemical plants as fuel agent like furnaces, boilers, space heaters, and cement kilns. Used lube can be recycled to generate grease and refined lube which is applied on the moving parts inside the vehicle in order to prevent excessive friction and reduce high heat.

2.3 Process Modelling

Process modelling is essentially an effort that involves relating together the properties of a system influenced by a process. Represented as mathematically symbols, the properties are associated with each other using relevant relations under one or more assumptions. The outcome is a set of mathematical equations, which is a process model.

2.3.1 Process Model

The term “process model” involves two terms, process which is defined as a set of activities taking place in a system, and resulting in certain effects on its properties. A process is either natural, or man-made. Natural processes-such as blood circulation in a human body, photosynthesis in plants, or planetary motion in the solar system-happen without human volition, and are responsible for certain effects on the associated systems. The other term is model that constitutes of a set of mathematical equations that involve system properties. A simple example of a model is the ideal gas law as in Equation (2.1):

$$PV = \mathrm{n}RT \tag{2.1}$$

Where P is gas pressure, V is the volume occupied, T is temperature in Kelvin, R is the gas constant and n is the number of mole.

However, process model is a set of equations or relations involving properties of the system under the influence of a process. The properties represent observable occurrences or phenomena classifiable into the categories of (i) initiating events, (ii) specifications, and (iii) effects. Thus, a process model is a scheme according to which a process with given specifications and initiating events would generate effects in the system.

Figure (2.3) shows the general conception of the model as a triangle, each side of which represents the relations between the two ends or vertices denoting the categories. If we know these relations, i.e., the process model, we can use it to unravel one or more unknown phenomena when the remaining ones are known.

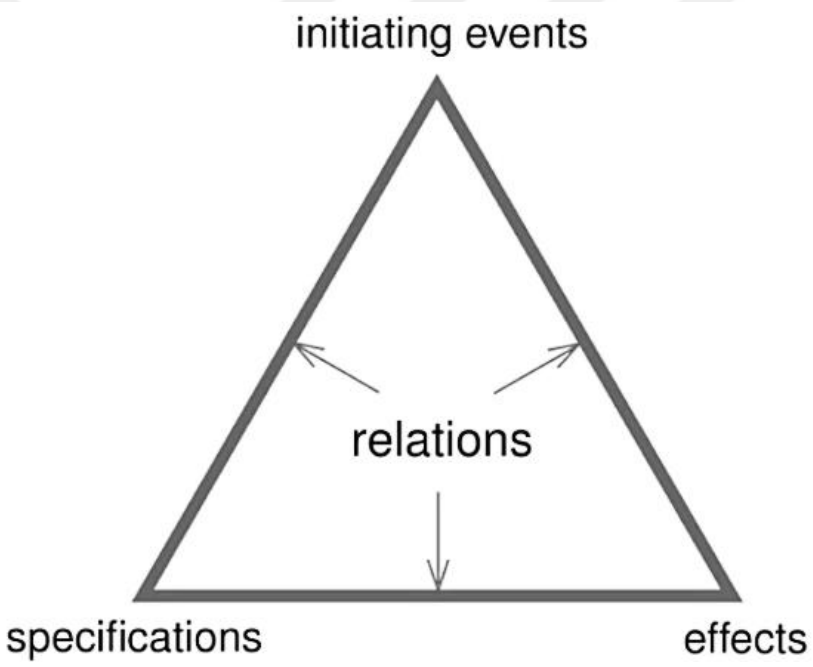


Figure 2.3 General conception of the model
(Source: Simant R. Upreti, 2017)

However, we can do this without having to execute the process in the real world but of course, the better the process model the better is our ability to explain the involved phenomena.

In particular, we show the effects based on its model, and the knowledge of initiating events, and specifications. As a matter of fact, we can predict the effects for different initiating event, and specification. Doing that enables us to isolate desirable effects,

and the related initiating events, and specifications. We can then apply the latter two in the real world to achieve the desirable effects from the process. This exercise basically is process optimization and control.

2.3.2 Linear Vs. Nonlinear Model

Although majority of chemical processes are thought to be nonlinear systems, conventionally linear input-output systems were being modelled as dynamic system for decades. The existence of a well-developed linear control theory enhances the use of linear models. Linear models may provide sufficient accuracy in the area of the linearization point, but they have inadequate predictive capability when the system has been subjected to large instabilities. Lately, the interest in describing chemical processes by nonlinear input-output models has increased significantly. This is mainly attributed to "bad" representation of linear models. Advance process monitoring and model-based control techniques are expected to give better results when process models that are more accurate over a wide range operating conditions are used. Model-predictive control methods which are becoming prevalent in chemical process industries permit straight use of nonlinear models in control algorithms. Another significant reason of the importance of the nonlinear models is the availability of new tools which provide automated techniques for building nonlinear models like neural networks. So far, other examples and works on the development of nonlinear modelling has been expanded over the last two decades (Haber and Unbehauen, 1990).

2.4 Predictive Analytics

Predictive analytics consists of variety of statistical tools and techniques like data mining tools, machine learning and predictive modelling approaches. These approaches unleash the secret behind data. In short, this approach allows us to predict the future. Data science algorithms take historical data and spit out a statistical model. Which can predict who will buy cheat, lie, or die in the future.

2.4.1 Scope of Predictive Modelling

Predictive modelling is an ensemble of statistical algorithms coded in a statistical tool, which when applied on historical data, outputs a mathematical function (or equation). It can in-turn be used to predict outcomes based on some inputs (on which the model operates) from the future to drive a goal in business context or enable better decision making in general.

Statistics are important to understand data. It tells volumes about the data. How is the data distributed? Is it cantered with little variance or does it vary widely? Are two of the variables dependent on or independent of each other? Statistics helps us answer these questions with mean of mean, variance, co-variance, and correlation.

Algorithms, on the other hand, are the blueprints of a model. They are responsible for creating mathematical equations from the historical data. They analyse the data, quantify the relationship between the variables, and convert it into a mathematical equation. There is a variety of them: linear regression, logistic regression, clustering, Decision Trees, Time-Series modelling, Naive Bayes Classifiers, Natural Language Processing, and so on.

2.5 Mathematical Function

Most of the data science algorithms have underlying mathematics behind them. In many of the algorithms, such as regression, a mathematical equation (of a certain type) is assumed and the parameters of the equations are derived by fitting the data to the equation.

For example, the goal of linear regression is to fit a linear model to a dataset and find the equation parameters of the Equation (2.2):

$$Y = \alpha_0 + \beta_1.X_1 + \beta_2.X_2 + \cdots + \beta_n.X_n \tag{2.2}$$

The purpose of modelling is to find the best values for the coefficients. Once these values are known, the previous equation is good to predict the output. The equation above, which can be considered as a linear function of Xi's (or the input variables), is the linear regression model.

2.6 Empirical Modelling

Empirical modelling is considered as one of the most frequently used methods in analysis and science, technology, industry, engineering, and management, and is of course is considered the most diverse in term of the variety of difficulties and problems. The initiation of fast, low-priced desktop computing integrated with current advancement in empirical modelling algorithms; like those of machine learning; caused this to make it possible to tackle and deal with new kind of problems and

exploring new tools and techniques. Empirical modelling is best used when theoretical modelling basis is not available. In recent years, scientists and researchers began to prove against this common perception by taking unusual and new approaches of solving such problems. As a result, they have developed the empirical models that are smoothly applicable, extendable, and practically supplemented with other techniques. The empirical modelling essential flexible point in that it resembles of the white-box models where correlation between input and output is approached.

Basically, empirical models are based on input-output relations that do not take into consideration the description of the processes taking place within the system. This why it is sometimes referred to as the "black-box" approach, because the system is considered as a non-transparent box with attention focused only on input and output variables, without much concern about what is going on inside the box. "Empirical modelling" is modelling directly from data, rather than by some analytic process (Hundley, 2014). However, an empirical relationship can be computed quantitatively by the mean of using curve-fitting approach (Shiavi, 2007). Techniques for empirical modelling have been obtainable for fairly some time and are reciprocally called curve fitting in engineering and numerical methods, regression analysis and machine learning.

2.6.1 Black-Box and White-Box Models

The term "black-box" defines input-output relation in which this the relationship in form of correlation between input and output is not identified. It resembles an "opaque box" in some way decodes the input data (signal) into some specific output but does not show how the translation occurs. Concerning modelling and simulation, black-box is implemented when modelling approaches are based on experimental data and observations but not on given formulas or laws. Knowing about the input and output data along with little knowledge on the behaviour of the system, an interconnected network can be developed like in the case of neural network. If these models were constructed properly, they will definitely work consistently (Rogsch and Klingsch, 2011). White-box built-on models, i.e. models with verified formulas, laws and adages, are favoured on black-box one. However, it is possible to reach a combination between the two approaches. The slow and steady evolution between white-box and black-box methods can be illustrated as in Figure (2.4). Most real-life problems which

can be defined by recognized laws can be simulated taking the white-box as the unique exemplary model.

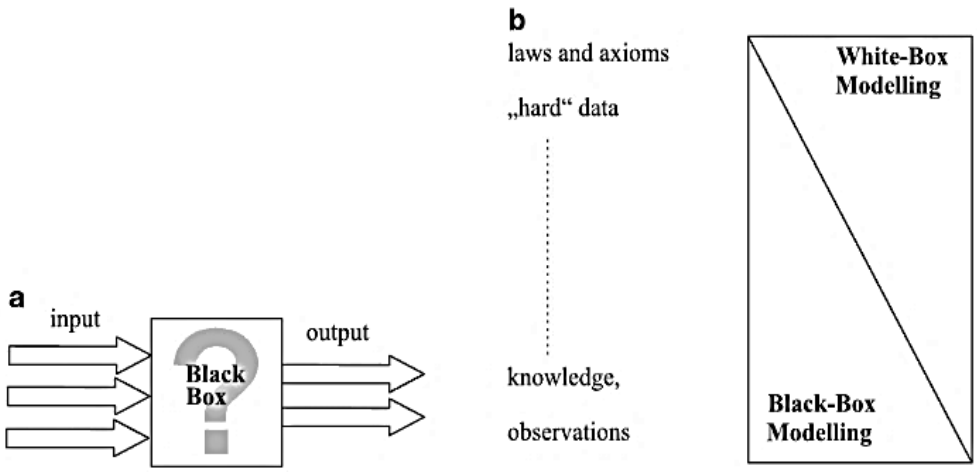


Figure 2.4 Black-Box (a) vs.White-Box (b) (Source: Gabriel Wurzer, Kerstin Kowarik, Hans Reschreiter, 2015)

2.6.2 Empirical Vs. Fundamental Models

The fundamental models which is also known as first-principle models, are built on physical-chemical relationships. Actually, these models are derived by applying the conservation principle may also include transport phenomena, reaction kinetics, conservation of energy, and thermodynamic (e.g., phase equilibrium) relationships. The fundamental models are beneficial pose great potential for solving complex system. Since these models include detailed physical-chemical relationships, they can dazzlingly mimic the behaviour of nonlinear systems. The other advantage of preferring the empirical method over the first-principles method is that the system's all over states are generally physical variables such as pressure, temperature or concentration which can be directly measured. Furthermore, the fundamental models are known for their time-consuming characters during development stage and they frequently require a large number of equalities that needed to be measured and estimated.

The empirical model is generally developed to use when the actual process is too complicated or complex and phenomena often cannot be interpreted with mathematical or numerical solution or when the empirical model provides satisfactory predictions of the process characteristics. Experimental plant data are used to develop

a relationship between the process input and process output as an empirical model using a mathematical framework such as artificial neural network (ANN) (Rojas, 1996), least square method (LSM) (Jain et al, 1995), etc. nonetheless, the time required to acquire this kind of models has been significantly reduced while the empirical models mostly can be used confidently as long as the operating range is considered in the operation (Parker, 1999).

2.7 Artificial Intelligence and Machine Learning

Artificial Intelligence (AI) is the study of intelligent behaviour (in humans, animals, and machines) and the attempt to find ways in which such behaviour could be engineered in any type of artefact (Whitby, 2003). Winston (1992) defined Al as "AI is the study of the combination that make it conceivable, reason and act." Another definition was made by (Nilsson, 1998) "AI..is concerned with intelligent behaviour in artefacts." Learning is a multiple-faceted phenomenon. Leaning mechanism involves the process of acquiring new informative knowledge, the design of cognitive skills through training or practice, the organization of new knowledge into general, effective representation, and the discovery of new facts and theories through observation and experimentation. Since the start of the computer ear, scientists were so motivated to perform such capabilities in computers. Solving this problem has been, and remains, a most challenging and fascinating long-range goal in artificial intelligence (AI). The computer modelling of learning processes in their manifold appearances constitutes the subject matter of machine learning.

Machine learning is defined as a group of approaches which are able to automatically detect any pattern in the data, and then used to unleash concealed or uncover patterns in order to be used in prediction of future data, or to perform other kinds of decision making under uncertainty (such as planning how to collect more data) (Murphy, 2014). Another definition introduced by Stephen Marsland (2015) as "machine learning is a technique of causing computers and processing units adjust and adapt processes to the nearest accuracy.

2.7.1 Types of Machine Learning

a) Supervised learning: it is also called leaning from exemplars which means that training a set of examples with true responses (targets). The algorithm performs generalization in order to allow for true reaction to all inputs.

b) Unsupervised learning: unlike to supervised learning, here provision of correct responses is not required, instead the algorithm tends to "pick" all similarities among inputs which have something in common and are categorized altogether.
c) Reinforcement learning: this type of machine learning combines between the supervised and unsupervised learning. Here, the algorithm is informed whenever the answer is wrong. However, it is not informed when the answer is correct. It is ought to find out rainbow of possibilities until the right answer is reached.

2.8 First Tool: Artificial Neural Networks

The term "neural network" comes as a result of thorough Artificial Intelligent (AI) researches, which tries to understand and mimic the human brain behaviour. According to Haykin (2009), Neural network is immensely in-parallel configured processor built on ordinary processing units that has ordinary tendency for storing up empirical data and making it available for use. It looks like the brain in two respects; knowledge is attained by the network from its atmosphere through a learning process and interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

Rojas (1996) described neural network as a modelling of biological nervous system composing of many inter-connected nonlinear computational units called (neurons). A neural network is an interrelated assemblage of simple processing entities, units or nodes, whose functionality is slackly based on the animal neuron. The processing capability is observed in the inter-unit connection strengths and weight, and obtained adaptation to the real mapping from learning from different set of trainings patterns. While Kevin (1997) gave a comprehensive definition to neural network as an interrelated assemblage of primary processing units or nodes which emulate nearly the animal neuron. The processing capability of the network is noticed in between the processing units in form of weights or strength attained by the pattern learning processes. Nielson (1990) has described neural networks as a computing unit consists of a number of simple but extremely interconnected nodes or processing elements, which can process data and information using its dynamic state response to any external input signal.

The ANN is extremely a simplified model of the brain and essentially a function approximate that transforms inputs to outputs to the best of its ability and according to Kevin (1997), artificial neural networks can be defined as a simplified models or systems that resemble the work of the neurons which happens obviously in the visceral brain. The application of neural networks aims at mapping a set of inputs onto a relative set of output patterns. The network performs this mapping task by first learning from a sequence of previous cases defining patterns of input and output belonging to the intended system. The network then runs on new set of inputs in order to predict the relative outputs.

In order to better understand the functioning of neural networks, let us imagine how the human brain works. In the nervous system of the brain, neurons interact in a complicated way and human external senses detect stimuli, and send this external signal "input" to the brain via neurons to the human brain. In the brain, other neurons receive the input signal and they excitedly interact with each other. According to the incoming inputs, the brain makes a decision and sends an "output" signal in the form of reaction to the response received by the human body sensual externals. This illustrated in Figures (2.5) and (2.6).

Figure 2.5 Human brain input/output signal interaction

Figure 2.6 ANN input/output signal interaction

2.8.1 ANN Biological Paradigm

In biological point of view, the ideal neuron structure consists mainly of cell body joined with tubular axon, and a number of hair-like dendrites, shown in Figure (2.7), the tiny dendrites create some very fine filamentary brushes around the body of the neuron. The axon like long thin tube that splits into many branches that almost approach and touch the dendrites of the adjacent cells. Synapse is that slit located between end bulb and the dendrites, across which information is propagated.

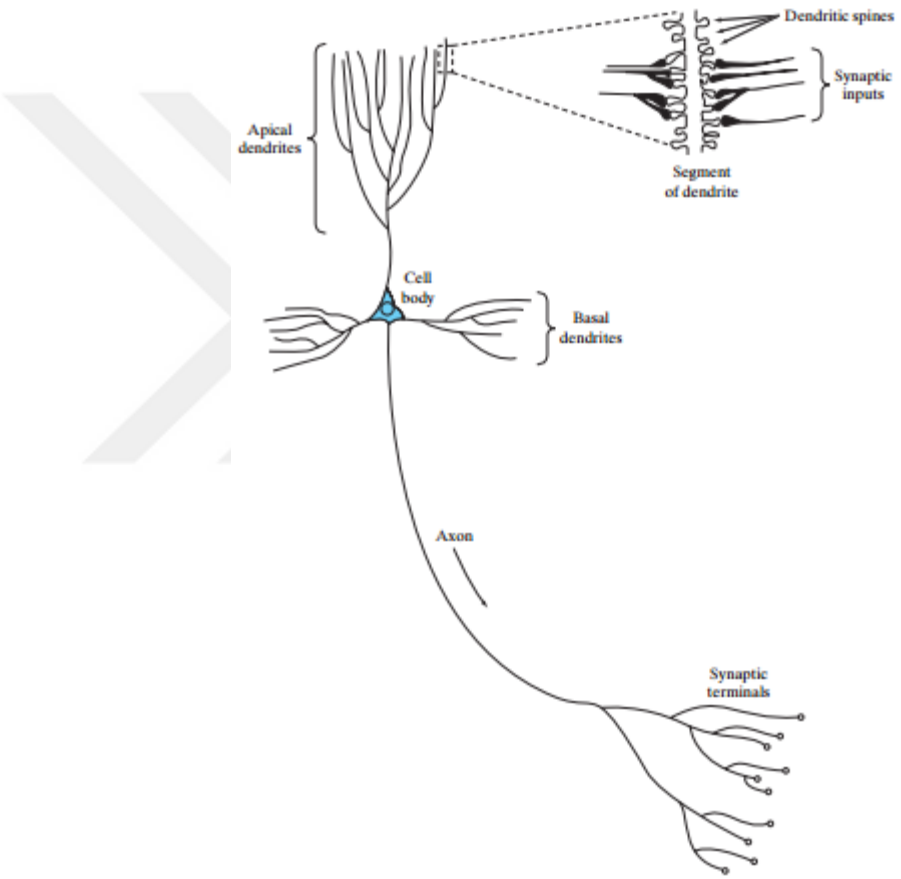


Figure 2.7 Humanoid neuron cell structure (Source: Simon Haykin, 2008)

In reality, brain in human body constitutes of around 10^{11} (100 billion) nerve cells and neurons. Neurons transfer electric pulses and communicate through electrical signals and the interneuron connections are interceded by electrochemical joints called synapses, which are positioned on branches of the cell referred to as dendrites. Each neuron naturally accepts many thousands of contacts from other neurons and

consequently receives numerous of incoming signals, which ultimately arrives at the cell body. Later, these signals are combined or added up together repeatedly. If, however, the yielding signal surpasses a specified threshold value. then the neuron would trigger or generate a voltage impulse as a response. This trip repeats itself over and over to other neurons via axon.

2.8.2 Artificial Model of Neuron

A neuron is a data-processing unit that is essential to the process of a neural network. The block diagram of Figure (2.8) shows a typical model consisting of a collection of neurons and its main fundamentals of the neural model:

1. In this model, the output of a neuron t signals, weighted by the corresponding synaptic strengths of the neuron.
2. A transfer function or activation function for restraining the output's amplitude of a neuron and it limits the allowable range of output to certain limited value.

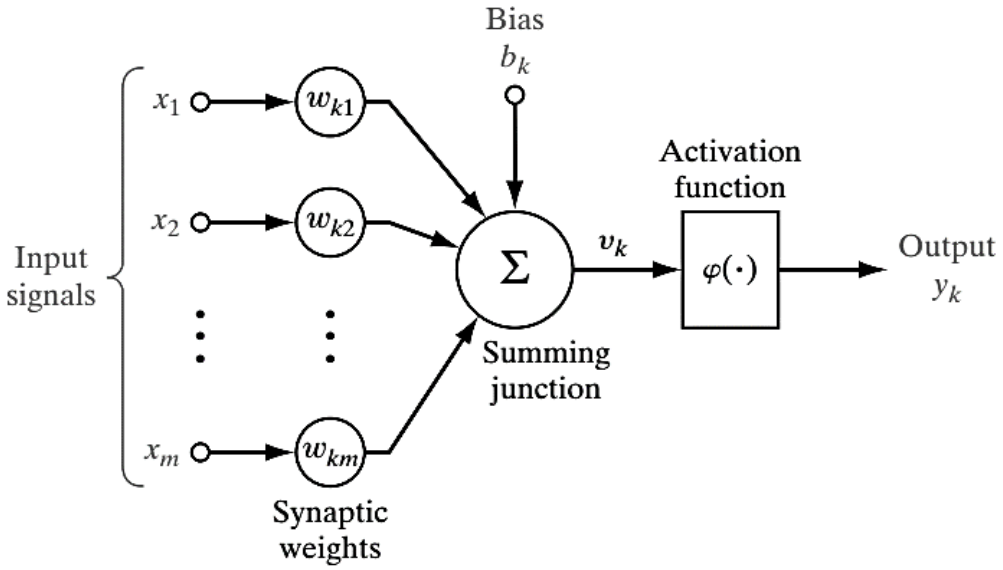


Figure 2.8 Block depiction of artificial neuron structure

Normally the output range of normalized amplitude is written in form of closed interval [0,1] or otherwise, [-1,1]. The graphical neural model representation of Figure (2.8) also contains an exterior applied bias, referred by b_k. The bias b_k has the influence of growing of dropping the absolute value of input of the transfer or activation function.

Mathematically we may describe neuron k depicted in Figure (2.8) by the pair Equations (2.2) and (2.3):

$$u_k = \sum_{j=1}^{m} w_{kj}\, x_j \tag{2.2}$$

and

$$y_k = \varphi(u_k + b_k) \tag{2.3}$$

Where x_1, x_2..x_m are the input signals; w_{k1} , w_{k2} ... w_{km} are the particular synaptic weights of neuron k; b_k is the bias; $\varphi(\cdot)$ is the transfer or activation function whereas y_k is the output signal. Interestingly, bias b_k is capable of granting strength to deploy an affine transformation to the output u_k of Figure (2.8) as shown by Equation (2.4).

$$v_k = u_k + b_k \tag{2.4}$$

2.8.3 Activation Functions

The activation function is used to calculate the output response of a neuron. The sum of the weighted input signal is applied with an activation to obtain the response. For neurons in same layer, same activation functions are uses. The nonlinear activation functions are used in a multilayer net. A few linear and nonlinear activation functions are discussed below.

1. Identity function

This function is represented by diagonal line indicating equal symmetric relations between dependent and independent variables. The function is given by

$f(x) = x$; for all x. This shown in Figure (2.9).

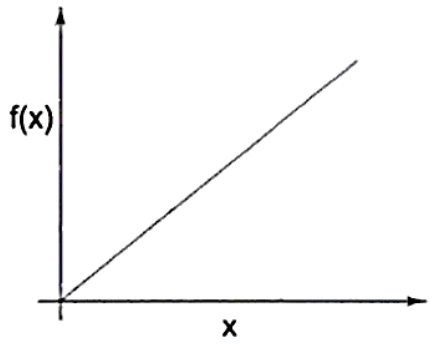


Figure 2.9 Identity function

2. Threshold function or binary step function

Heaviside Function is referred to threshold function and the output of neuron k is expressed as in equation (2.5).

$$y_k = \begin{cases} 1 & if \ \ v_k \geq 0 \\ 0 & if \ \ v_k < 0 \end{cases} \tag{2.5}$$

where v_k is the induced local field of the neuron; that is

$$\sum_{j=1}^{m} w_{kj} x_j + b_k \tag{2.6}$$

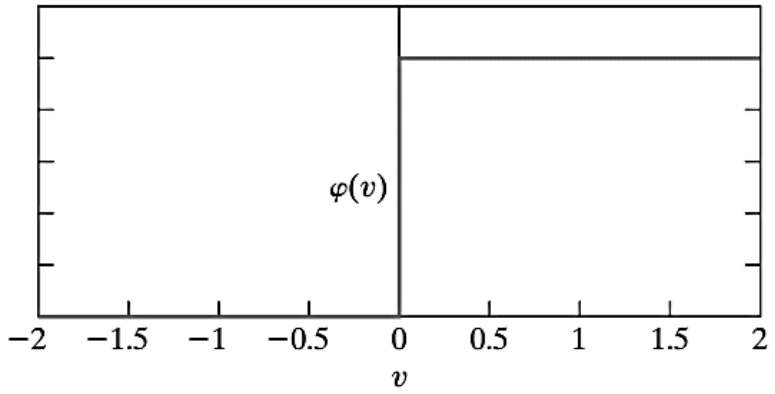


Figure 2.10 Binary step function

3. Sigmoid function

The sigmoid function which is graphically represented in shape of "S" letter, which is shown in Figure (2.11), is absolutely the most widely used activation function which

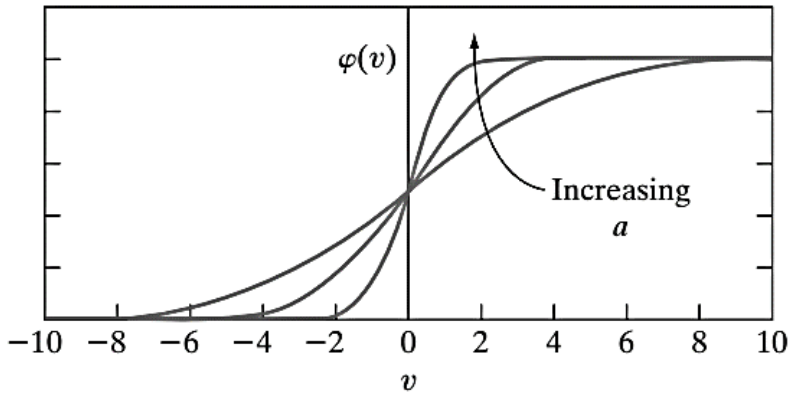


Figure 2.11 Sigmoid function

is also used in the construction of the neural networks. This is due to the powerful ability to balance between behaviour of linear and nonlinear system modelling. However, there are two main types of sigmoidal functions:

1. logistic function or binary sigmoidal function where the logistic function ranges between 0 and 1, and is defined by the Equation (2.7).

$$\varphi(v) = \frac{1}{1+exp(-av)} \tag{2.7}$$

Where *a* is called the slope parameter of the above sigmoid function. By changing the value of parameter, *a*, we get sigmoid functions of changed slopes, as illustrated in Figure (2.12). The slope at the origin equals $a/4$ and as the slope parameter pursues infinitude, the sigmoid function changes into a threshold function. In case a threshold function assumes the value of 0 or 1, a sigmoid function keeps *a* constant range of values from 0 to 1. The activation functions well-defined in Equation (2.7) range from 0 to +1.

2. Hyperbolic tangent function or Bipolar Sigmoidal Function shown in Figure (2.13) where the hyperbolic function ranges between +1 and -1 and is defined by

$$\varphi(v) = tanh(v) \tag{2.8}$$

Occasionally it is preferable to have range from -1 to +1 of the activation function where the activation function is considered as an odd function that induces local field. The threshold function of Equation (2.9) is now defined as:

$$y_k = \begin{cases} 1 \ if \ v_k > 0 \\ 0 \ if \ v_k = 0 \\ -1 \ if \ v_k < 0 \end{cases} \tag{2.9}$$

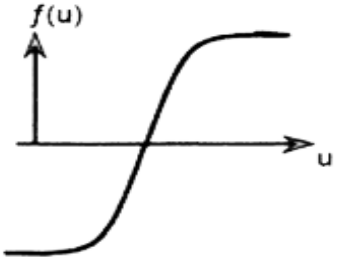


Figure 2.12 Hyperbolic tangent function

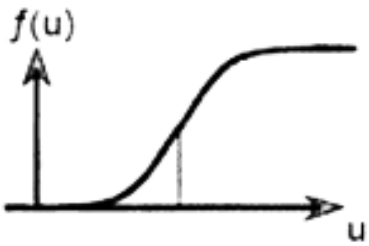


Figure 2.13 Logistic function

2.8.3.1 Signal-Flow and Mathematical Definition of Neuron

A neural network as shown in Figure (2.14) is a directed graph containing of nodes with interconnecting synaptic and activation links and is considered by four properties:

1. Each neuron is signified by a set of linear synaptic links, an externally imposed bias, and a possibly nonlinear activation link. The bias is represented by a synaptic link connected to an input fixed at +1.

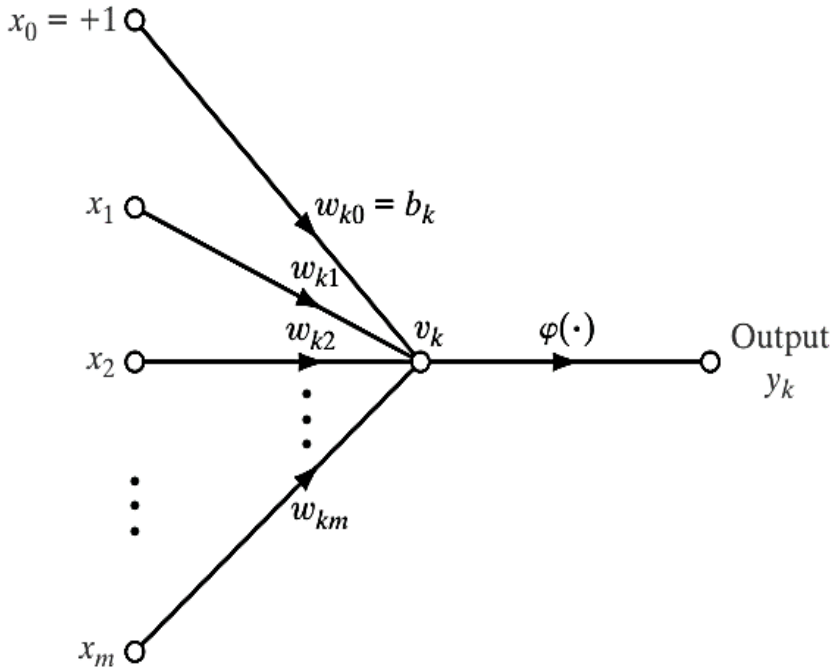


Figure 2.14 Neuron signal flow

2. The synaptic links lastly, when compared to empirical neuron weight their respective input signals.
3. The weighted sum of the input signals defines the induced local field of the neuron in question.
4. The activation link squashes the induced local field of the neuron to produce an output.

2.8.4 Topology of ANN

The topology of a neural network refers to how its nodes are interconnected. With interlayer topology, there are two options: (1) feedback connections, and (2) feedforward connection as illustrated in Figure (2.15).

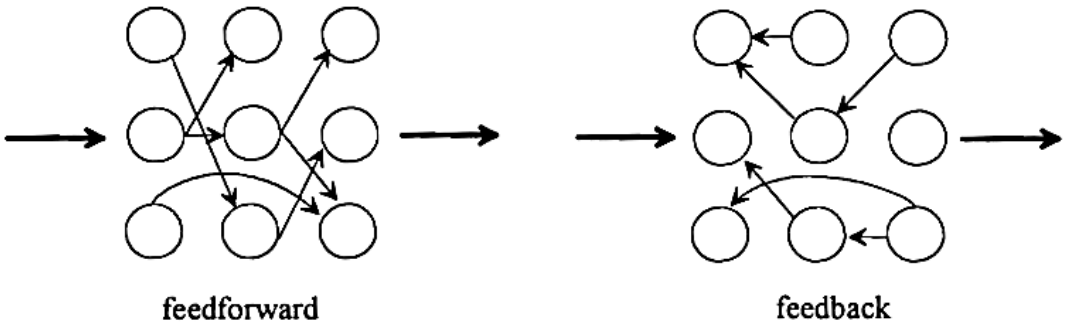


Figure 2.15 ANN feedforward and feedback connection

The type of problem we are trying to solve determines which topology we favour. And since we are dealing with dynamic modelling of a chemical process, where we are trying to map an output response based on an input signal, we would favour the feedforward connection.

2.8.4.1 Multilayer Feedforward Neural Networks

This is the second type of a feedforward neural network and it distinguishes itself from others by having single or multiple hidden layers where hidden neurons or units are considered as computation nodes. The term “hidden” indicates that this part of the neural network is not seen directly. The real function of hidden neurons is to link between two variables, e.g., the external input and the network output in some appropriate manner (Churchland and Sejnowski, 1992).

The main nodes in the layer of input of the network source particular elements of the transfer function. Typically, every neuron existing in each layer inside the network has its own input and output signals of the preceding layer. The established layer of the final output signals of the neurons forms the overall responses of the network. The architectural graph in Figure (2.16) illustrates the layout of a multilayer feedforward neural network. For the sake of brevity, the network in Figure (2.16) is referred to as

a 10-4-2 network configuration because it has 10 source nodes, 4 hidden neurons, and 2 output neurons.

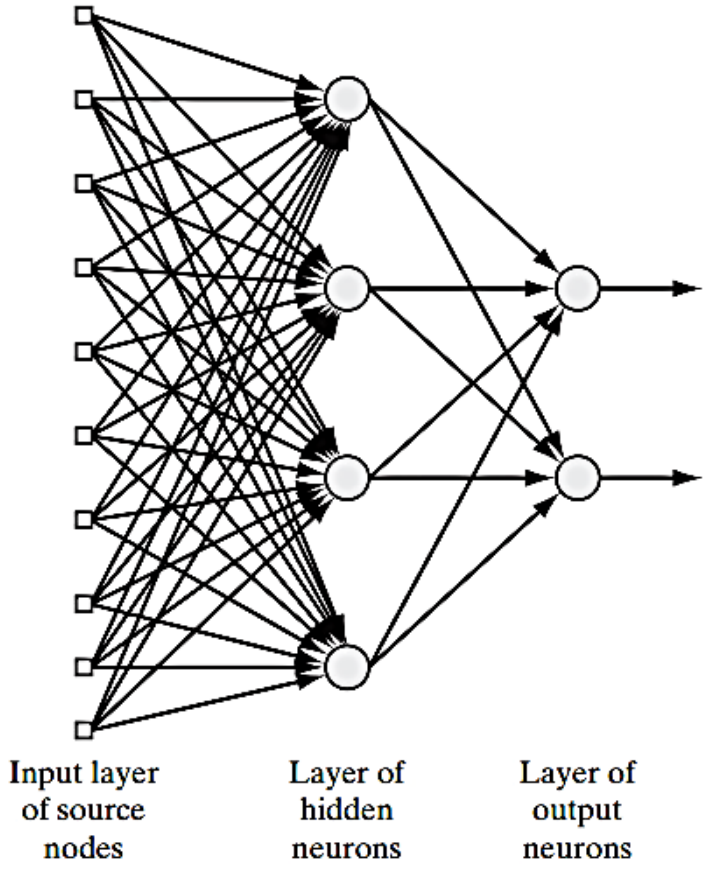


Figure 2.16 ANN 10-4-2 network

2.8.5 Potential Applications of Artificial Neural Networks

Applications of neural networks to bioprocessing and chemical engineering have increased significantly since 1988. One of the first application papers was by Hoskins and Himmelblau (1988) who implemented a neural network on the fault diagnosis of a chemical reactor system. Since then, the number of research journals on neural network applications in bioprocessing and chemical engineering has climbed incredibly. Ware (1991) provides a good overview of potential applications of neural networks, as listed below:

- Classification - predict categorical output by using input values; e.g., determining most likely to happen disease through certain symptoms.

- Prediction - use input variables to estimate an output value; e.g., determining rate of evaporation giving temperature, wind velocity and humidity as input parameters.
- Data association – determine typical data through error-free association. Also determine data which contain error; e.g., recognize input patterns with noisy among five ideal patterns.
- Data conceptualization – analyse data and determine relationships conceptually; e.g., group data with several attributes in order to infer certain relationships.
- Data filtering – make an input signal smooth and even; e.g., lessen a noisy electrocardiogram input signal and make it free of noisy.
- Optimization – find out optimum value; e.g., estimate minimum-length trip for a traveling salesperson.
- Neural networks hold much capacity for important applications to bioprocessing and chemical engineering problems that are complex, nonlinear, and uncertain. A well-structured and possibly a large neural network can effectively tackle complex modelling problematic issues.

  This text focuses on the bioprocessing and chemical engineering applications that primarily use classification, prediction and data-filtering networks. We concentrate on the following applications:
- Prediction and optimization networks – predicts the values of process performance variables from independent operating variables based on laboratory of plant data. For quality control applications. Note that this is not the same as using an optimization network, since the optimization is performed using statistical methods and numerical techniques rather than by the network itself. In addition, prediction networks can be used to as “software-based sensors” (soft sensors) for quantitative predictions of variables that are not easily measurable.
- Process-forecasting, modelling, and control networks – are prediction networks similar to those previously mentioned and used for process optimization and adaptive process control.

2.8.6 Some Previous Studies in Bioprocessing and Chemical Engineering:

- Estimation of so called “mass-transfer coefficients” in electrochemical refining of metals (Wizzard and Fehrman, 1991).
- Estimation of acid strength of mixed oxides in designing multicomponent catalyst (Kito et al, 1992).
- Prediction of tray temperatures in distillation columns (Schnelle et al, 1990; Chenng et al., 1992; Ponton and Klemes, 1993).
- Prediction of the silicon ration in the blast furnace data (Bulsari and Saxen, 1991).
- Estimation of fermentation efficiency and performance parameters (Oishi et al., 1992; Hofland et al., 1992).
- Prediction of complicated and complex kinetics in processes taking place in metallurgy and mineral mines (Reuter et al., 1993).
- Prediction of performance of a dying system (Huang and Mujumdar, 1993).
- Estimation of the performance of the carbon-in-leach gold-recovery process (Van der Walt et al., 1993).
- Prediction of quantitative structure-activity relationships (OSAR) of pharmaceuticals (Lie et al., 1992a,b).
- Prediction of autoclave performance data in composite manufacturing (Wu, 1990; Wu and Joseph, 1990; Joseph et al., 1992; Shiek, 1992; Joseph and Wang, 1993).

2.9 Regression Analysis

Regression analysis is a simple an old technique for investigating functional relationship among variables. It is used to model the relationship between a response variable and several predictor (explanatory) variables. Besides, regression analysis technique became widely followed and used statistical for analysing data and proving functional relationship among variables (Chatterjee and Hadi, 2006). The description of collected or measured data, especially by linear regression function with several parameters, is needed for modelling and predicting (Späth, 1992).

Once a model has been identified, various forms of inferences such as prediction, control, information extraction, knowledge discovery, and risk evaluation can be done within the framework of deductive argument. Thus, the key to solving various real-

world problems lies in the development and construction of suitable linear and nonlinear regression modelling. Douglas C, (2012) has defined regression analysis as a statistical technique for investigating and modelling the relationship between variables. Regression analysis allows one to relate one or more criterion variables to one or more predictor variables. The term "to relate" is an empty formula that can be filled with an ever-increasing number of concepts, assumptions, and procedures (Eye and Schuster, 1998).

Applications of regression analysis are plentiful and happen in nearly every field, including engineering, the physical, chemical, and biological sciences, management, economics, and the social sciences. In fact, regression analysis may be the most widely used statistical technique. There are two main types of regression; one is simple linear regression and the other is simple nonlinear regression. However, regression analysis can include multiple variables that differentiates itself from the simple regression to multiple regression. The multiple regression analysis (MRA) can be categorized into two kinds; Multiple Linear Regression (MLR) and Multiple Nonlinear Regression (MNLR). Regression analysis can be used in the following area:

1. Data description
2. Parameter estimation
3. Prediction and estimation
4. Control

Of course, there are other alternative methods of regression and according to Birkes and Dodg (1993), there are alternative methods of regression like: Least-Squares Regression, Least-Absolute-Deviations Regression, M-Regression, Nonparametric Regression, Bayesian Regression and Ridge Regression.

2.9.1 Multiple Regression

Allison (2004) defined multiple regression as a statistical method for studying the association relation between a single variable called "dependent variable" and one or multiple variables named "independent variables". The term "multiple" indicates that there are two or more independent variables.

Multiple regression does two things that are very desirable. For prediction studies, multiple regression makes it possible to combine many variables to produce optimal

predictions of the dependent variable. For causal analysis, it separates the impact of independent variables on the dependent variable so that you can inspect the unique involvement of each variable. The formula for multiple regression is defined as in Equation (2.10):

$$Y' = a + b_1X_1 + b_2X_2 + .. + b_nX_n \quad (2.10)$$

where

Y' = a predicted value of Y (which is dependent variable)

a = the "Y intercept"

b_1= the change in Y for each 1 increment change in X_1

b_2= the change in Y for each 1 increment change in X_2

X_n= an X score (X is independent variable)

Seber and Wild (2003) stated that the statisticians play main role to find relationships between variables. In regression problems and in typical condition, there is one of the variables, frequently known as the dependent or response variable and is denoted by y while the other variables $x_1, x_2, \ldots x_k$, typically termed as independent, explanatory or regressor variables, are used to predict the behaviour of y. Plotting this relationship between y and x_i can be expressed via some function f as in Equation (2.11):

$$y \approx f(x_1, x_2, \ldots, x_k) \quad (2.11)$$

2.9.1.1 Second Tool: Multiple Linear Regression

The term "multiple linear regression" and the approaches for exploring the relations associated with two kinds of variables; dependent and independent variables; originated to one century back. Francis Galton was first to introduce multiple linear regression in 1908 when he was busy with his study in heredity. During one of his experiment, he observed that kids of tall parents are likely to be taller than average but below their parents' average height. This "regression toward mediocrity" explains statistical relationships between all variables, particularly in the simple regression which describes relationship between one dependent variable (y) and one independent

variable (x). The simple linear regression model is depicted in Figure (2.17) and stated in the Equation (2.12):

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{2.12}$$

Where y variable is substituted with dependent value while β_0 is a constant which intercept with y and β_1 represents a slope of the regression line and finally x is the variable with independent values with ε is forming a random error.

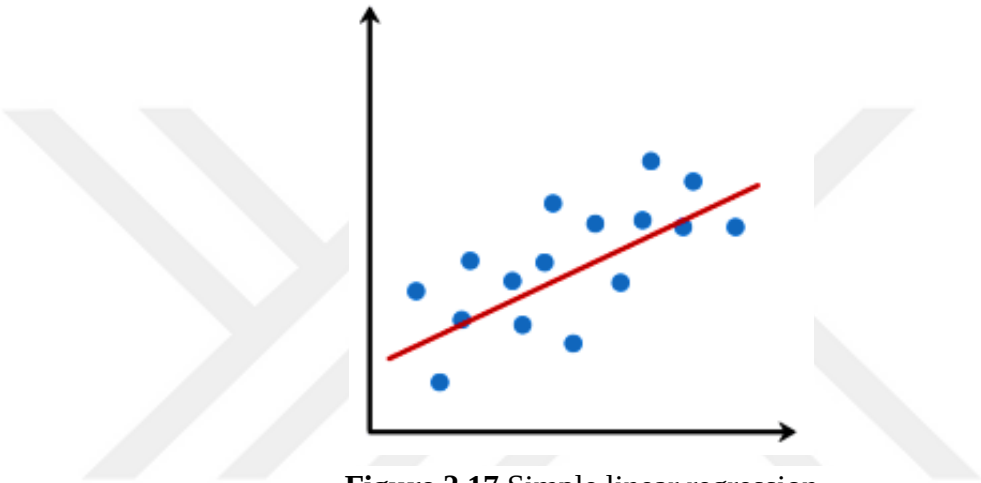


Figure 2.17 Simple linear regression

The dependent variable which sometimes is referred to response variable whereas the independent variable is sometimes called predictor or explanatory variable which is responsible for altering changes in the response data. Regression model can be described more generally as in Equation (2.13):

$$y = E(y) + \epsilon \tag{2.13}$$

Where $E(y)$ is the value of the response variable, like to say, when $E(y)$ is composed of exploratory variables in form of linear combination $x_1, x_2, \dots, x_k$. For example, if k = 1 the regression indicates a simple linear regression and when $E(y)$ represents a nonlinear function of $x_1, x_2, \dots, x_k$, then the model is nonlinear.

2.9.1.2 Third Tool: Multiple Nonlinear Regression

Multiple nonlinear regression is one of the simplest tool where relationships between more than two variables are uncovered in form of second-order (or higher) mathematical representations and formulas. Usually, two types of variables play role in this case, response (or predicted or dependent) variable, which depends on a set of variables called the explanatory variables (or regressors or independent variables). Figure (2.18) shows the graphical representation between the outcome and predictor variables.

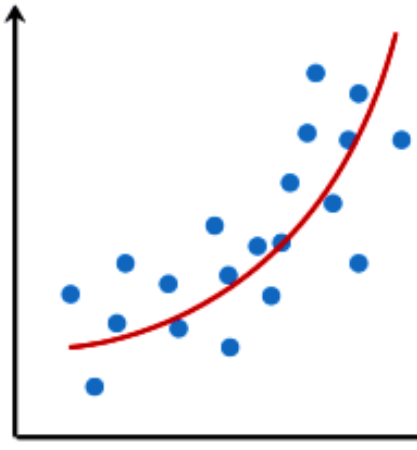


Figure 2.18 Non-linear regression

2.9.1.3 Polynomial Regression

Polynomial regression is a form of nonlinear regression in which the relationship between the independent variable x and the dependent variable y is modelled as an nth order polynomial. Polynomial regression reveals the possible nonlinear relationship between two parameters, e.g., the value of x and the corresponding provisional mean of y, denoted E(y x), and has been used to describe nonlinear phenomena of various chemical processes. Although polynomial regression fits a nonlinear model to the data, as a statistical estimation problem it is linear, in the sense that the regression function E(y x) is linear in the unknown parameters that are estimated from the data. In term of mathematics, a polynomial is described as an expression containing a group of variables and coefficients that includes any mathematical operations like addition, subtraction, multiplication, and non-negative integer exponents.

An example of a polynomial of a single indeterminate (or variable), x, is indicated in Equation (2.14).

$$x^2 - 4x + 7 \tag{2.14}$$

which is a quadratic polynomial. Polynomial can fit to a family of nth order functions, e.g., first-order function (linear regression), second-order function (quadratic regression), third-order function (cubic regression), and so on. Figure (2.19) shows different polynomial fitting to the motorcycle data.

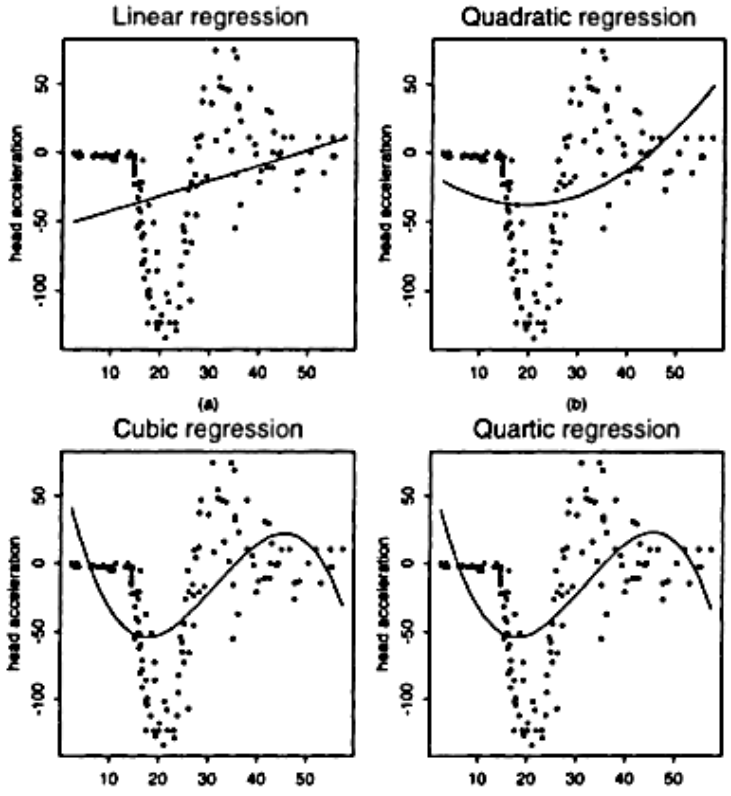


Figure 2.19 Different polynomial fitting to the motorcycle data

2.10 Fourth Tool: Ridge Regression

Ridge Regression is an analysing designed to deal with multiple regression data that experience multicollinearity. Ridge regression is a technique used to analyse data with highly correlated independent variables (Hurwitz, 2015). When multicollinearity happens, minimum estimates are unbiased. However, the variances usually are great and they could be at distance from the accurate value.

2.10.1 Multicollinearity

Multicollinearity (sometimes called collinearity) is statistical phenomena when here are somewhat a linear relation appears. For instance, assume that the three elements of a combination are examined by involving their ratios of the total. Simply, these variables will seemingly have a (perfect) linear relationship: P1+P2+P3=100. However, during regression processing, the previous logical relationship leads to a division by zero. Subsequently, causing all calculations to be ended. But when the relationship is not the same, the calculations are not ended and division by zero is not applicable though dividing by an infinitesimal quantity will still distract the results.

2.10.2 Effects of Multicollinearity

Multicollinearity can produce erroneous estimations in the coefficients of regression. It can also cause inflation in the standard error in coefficients of regression and cause deflation in the partial t-tests in the coefficients of regression.

2.10.3 Ridge Regression Models

The ridge regression model equation can be described in the form of matrix as in Equation (2.15):

$$Y = X\mathrm{B} + e \tag{2.15}$$

Where *Y* is dependent value while *X* variable holds for the independent value, and B is a constant value represents the regression coefficients that represents the error or residuals.

2.10.4 Standardization

Standardization of variables is the initial step in ridge regression analysis. Both dependent and independent variables' means are subtracted and divided by their standard deviations.

2.10.5 Ridge Regression Basics

The coefficient values of regression are estimated using the Equation (2.16):

$$\hat{B} = (X'X)^{-1}X'Y \tag{2.16}$$

$X'X = R$, where R is independent value of correlation in form matrix and the estimated value of population described in Equation (2.17):

$$E(\widehat{B}) = \sigma^2 R^{-1} \tag{2.17}$$

And assuming y's are in standard form, σ^2=1. Thus, we find that

$$V(\widehat{b_j}) = r^{jj} = \frac{1}{1-R_j^2} \tag{2.18}$$

Where R_j^2 is the R-squared value where in this situation, variance, as indicated in Equation (2.18), is the VIF and R-squared as a denominator becomes nearer to 1 and the variance (and hence VIF) will increase promptly.

Rule of thumb drop this value for VIF to 10. If we solve in backwards style, we can translate it into value equal to 0.90 of R-squared. henceforth, when the value of R-squared falls between larger than 0.90 (or equal) and a single independent variable, we will definitely have to expect multicollinearity.

2.10.6 Choosing K (Ridge Trace)

In order to choose k value or ridge trace, let suppose this example when the values of k are determined and seen on a logarithmic scale. As a result, a vertical line is drawn at the nominated value of k which is 0.006 shown in Figure (2.20). Thus, the vertical line or axis which contains the points for the least squares solution.

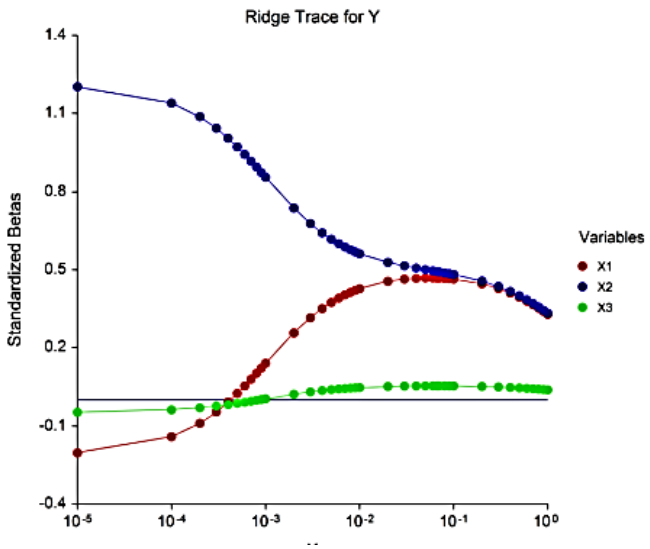


Figure 2.20 Ridge trace for Y variable

2.10.7 Analytic k

In 1976, the scientist Hoerl and Kennard introduced a different way for choosing k value. His technique was built on the Equation (2.19):

$$k = \frac{ps^2}{\tilde{B}'\tilde{B}} \tag{2.19}$$

In order to get the prime value of k, least squares coefficients are used to produce a value of k. The ridge trace, which is found in the plot of the coefficients belong to the ridge regression for different values of k, is an important part of ridge regression. The sums of squares of residuals should also be plotted. The ridge trace is examined for trends of the ridge coefficients as k is changed. The best estimates of the ridge coefficients are those where the trace shows that the coefficients have stabilized and the sums of squares of residuals is still small (Marquardt and Snee, 1975).

Hoerl and Kennard (1970b) discuss the use of the ridge trace to eliminate variables with the least predicting power. Thus, ridge regression can be used as a guide for selecting the best subset of variables; that is, ridge regression is an alternative for stepwise regression.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This chapter presents the case study description and design and methodology that were followed to fulfil the research objectives. For simplicity, the research work was divided into 4 phases. The first phase includes data collection, pre-processing, processing, and post-processing whereas the second phase includes construction of predictive models for different empirical techniques. These techniques are ANN, Ridge Regression, and Regression Analysis including the Multiple Linear Regression and Multiple Nonlinear Regression. The third phase consists of procedures followed to construct and fully built up mathematical model that would represent the chemical system. Consequently, and finally, the fourth phase includes optimization of the chemical process by mean of mathematical modelling. Specifically talking, this study displayed lists of forecasted or predicted readings generated experimentally from the lube oil recycling unit. Subsequently, these lists were utilized to empirically construct predictive models and metrically test performance of each model to explore each model's strengths and flaws. Moreover, the mathematical model developed was used to build all functional relations for all parameters taking into in to consideration all constraints existing in the lube oil recycling unit. Likewise, optimal variables were obtained for the temperature and vacuum with which we could optimize our system with higher productivity, quality and lower waste cost.

3.2 Research Methodology

The research methodology is implemented with the mean of quantitative methods and statistical tools. As illustrated in Figure (3.1), work flow diagram shows 4 phases during the research methodology starting with data collection and collection, through model construction and ending with building of mathematical model for the chemical process exist in the lube oil recycling unit.

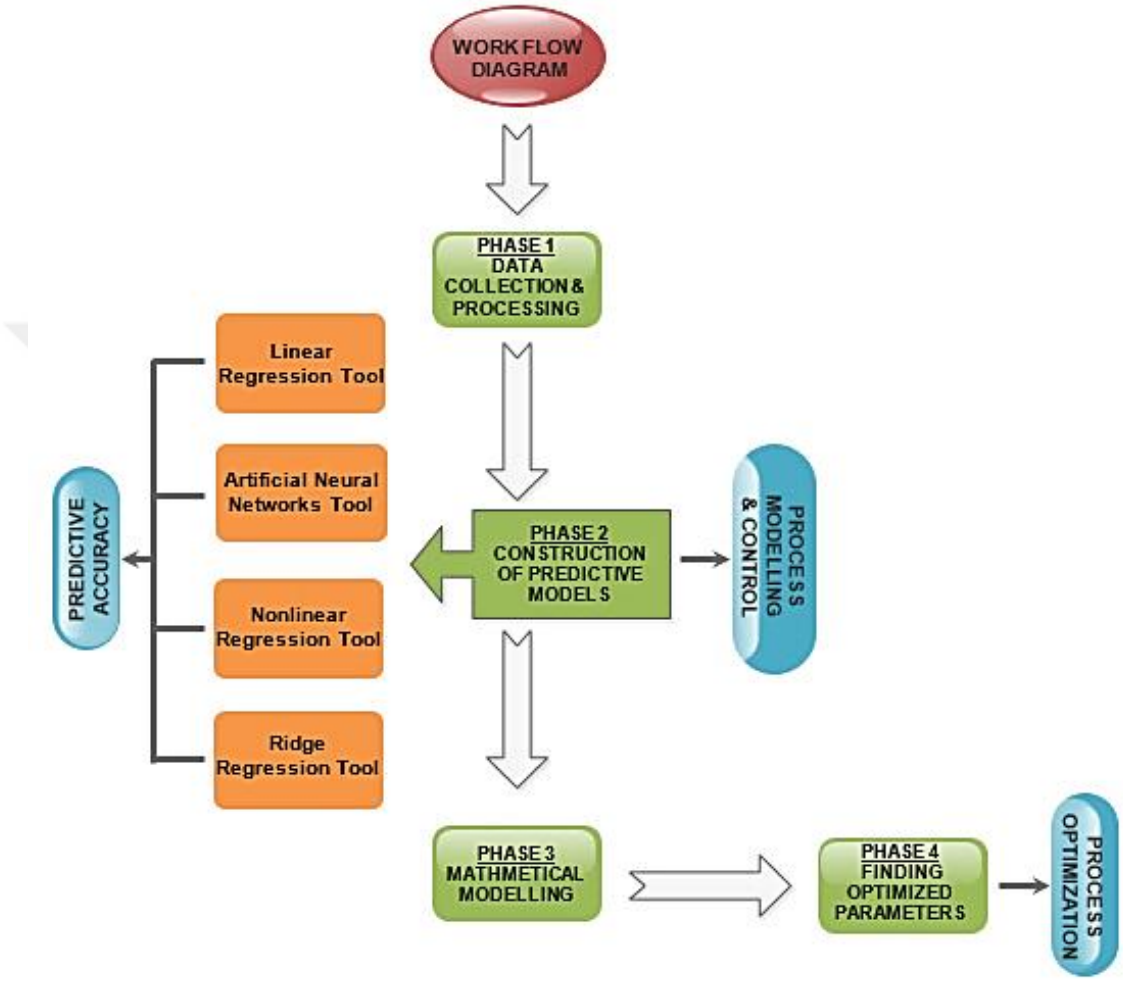


Figure 3.1 Methodological work flow diagram

3.3 Introduction to The Used Lubricant Recycling Unit

This work was conducted on a typical real-life chemical system consisting of complete arrangement of components which function integrally as a recycling unit of used oil. The unit was installed in Adana, Turkey where it was run and underwent many operating trials. Figures A.1 to A.6 in Appendix 1 shows the unit and its auxiliaries. The unit was distinctively designed and constructed based on a hybrid (combinational) system including both vertical and horizontal distillation units. The unit produces three main end products namely: light gas oil (LGO), moderate gas oil (MGO) and heavy gas oil (HGO) which are used in various power plants as a fuel. However, the unit encountered many problems including the huge waste produced along with the end products.

3.3.1 Description of The System

This system is intended to fractionate disposed used car and hydraulic oil (automotive oil) into useful products such as diesel and different weight gas oil. The system can cook around 20,000 litres of used car oil or hydraulic oil in one feed or one day. Moreover, the system is hybrid that it consists mainly of integrated vertical and horizontal distillation column and reactor ultimately which inherits the system its high economical and efficiency values through continuity (non-stop production) and less cost as well.

3.3.2 System Process Parameters

This system works on various temperature values (range: 380c° – 400c°) and vacuum values (range: 656mmHg – 730mmHg). However, we can classify them into two main groups as below:

1. Independent variables:

a) Temperature

b) Vacuum

2. Dependent variables:

a) Quantity of each distillate (heavy, moderate and light gas oil)

b) Quality of each distillate (heavy, moderate and light gas oil)

3.4 Phase (1)

3.4.1 Collecting The Data

Data collection can be defined as the process of collecting and gathering information and data. The data, then, is measured based on targeted variables in a systematic way. However, data collection involves humans and machines and both can be prone to errors. The alternative model development paradigm is based on developing relations based on process data. Input-output models are much expensive to develop. However, they only describe the relationships between the process inputs and outputs, and their utility is limited to features that are included in the available data sets (Palazoglu and Kayihan, 2007). Therefore, data in this research were collected by mean of experimental methods to provide a ground for establishing models for the chemical processes.

However, the lubricant recycling unit contains two input (independent or controllable) variables and three output (dependent or uncontrollable) variables. A set of 82 observations were taken experimentally from the chemical system from the daily record during the testing period. It is a worthy to say that the system was run under three constraints. The full description of parameters involved in the chemical system is illustrated in Figure (3.2) as below:

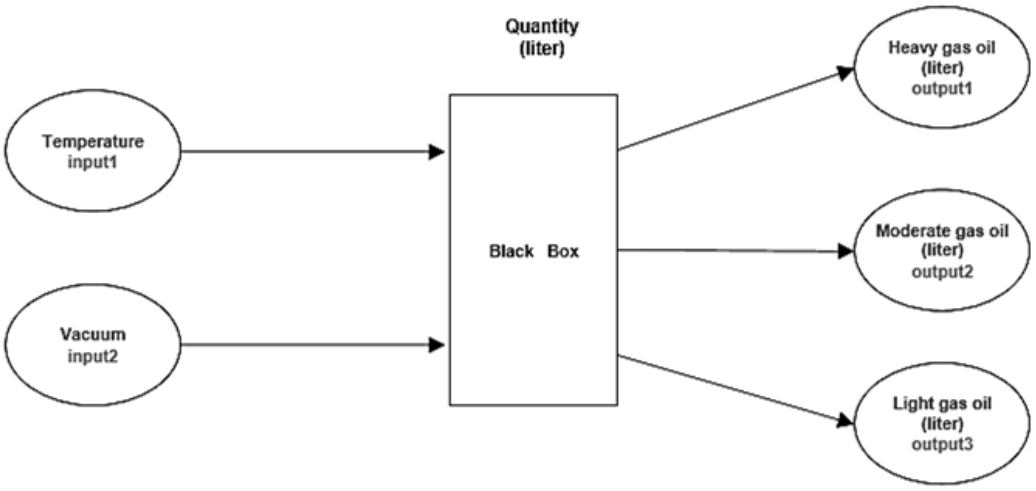


Figure 3.2 Parameters involved in the system

There are several ways to collect data such as survey, simulation or experiment. However, in our study, we collected data by conducting an observational study. In data

sampling, we took readings by direct observation method. We tried to fit our sample size to our system in regard of the nature of the system. That is to say, between 60 to 100 samplings is rational and can represent sufficiently the real system as far as chemical process is concern and our sample size was set up to 82.

3.4.2 Pre-Processing The Data

Pre-processing is a process which is the primary step in data mining and discovery and implies eliminating the inappropriate, noisy, unreliable, incomplete, redundant data, which can have a negative impact on the entire process. The reason is that the process results are sensitive to the quality of data (Lovrek and Jain, 2008).

In the chemical and real-world databases are highly affected by undesirable factors such the presence of noise, inconsistent and superfluous data and huge sizes in both dimensions. Therefore, input data must be provided in a way it serves as input for certain applications with algorithms.

3.4.2.1 Data Cleaning

Data cleaning is meant by "cleaning" targeted data by substituting or filling any missing data or values, smoothing noisy data, identifying and then omitting all outliers, and fixing inconsistencies.

In this phase of data pre-processing, we first detect and "clean out" all corrupt records in the datasets of input values e.g. missing data, outlier or noisy data, and anomalies, which were treated and removed from the dataset and ready for processing.

3.4.2.2 Data Transformation

Data transformation is to use simple mathematical formulations or learning curves to convert different measurements of selected, and clean, data into a unified numerical scale for the purpose of data analysis (Olson and Delen, 2008).

We transform our data to data proper for analysis and thus, we perform linear transformation on the original data using min-max normalization method. This method suggests that we find our maximum and minimum values from our dataset and assign them to a range between [1, -1] or [0,1].

In order to restrain all values within the range [-1,1] of all input and output data, we deployed the mathematical Equation (3.1):

$$x' = 2\frac{x - \min x}{\max x - \min x} - 1 \quad (3.1)$$

It is importance to notice that the purpose normalizing our dataset is to avoid very high results during integration process. This phenomenon occurs when input values are fed into the neuron and cause the transfer function (in case of ANN modelling) to show modest performance.

3.4.3 Post-Processing the Data

After getting the dataset normalized with proposed ANN models, the generated set of values undergo denormalization process including the network`s output.

3.4.3.1 De-Normalization the Data

As explained in the pre-processing phase, values in the datasets were regularly fed into the network and the output of neural network is de-normalized using the Equation (3.2):

$$y_D = 0.5[y_N(x_{max} - x_{min}) + (x_{max} + x_{min})] \quad (3.2)$$

Where y_N is the output of the proposed network, y_D is the de-normalized network output, x_{max} and x_{min} are the optimum values in the dataset.

3.5 Phase (2)

3.5.1 Construction of Predictive Models

To control processes in any chemical plant, it is very critical to understand the processes and their behaviour. To do so, we need to construct models that can predict future events for the system. This allows us to take preventive action and avoid unnecessary spending and waste. Predictive modelling is an ensemble of statistical algorithms coded in a statistical tool, which when applied on historical data, outputs a mathematical function (or equation). It can in-turn be utilized and formulated to predict future events based on some inputs (on which the model operates) from the future to drive a goal in business context or enable better decision making in general (Kumar, 2016).

3.5.1.1 First Tool: Multiple Linear Regression (MLR) Model

In this research, we unleashed statistical relationship as:

a) Two input variables (temperature and vacuum)
b) One output variable (light gas oil quantity

Once we are fully satisfied with the statistical relationship exists, we modelled the relationship mathematically and then used the model of prediction. Excel and SPSS programs were used to generate linear function of the chemical system.

3.5.1.2 Second Tool: Multiple Nonlinear Regression (MNLR) Model

Nonlinear regression resembles its previous type; linear regression; in that it explores the association between variables with dependent and independent nature to evaluate the involvement of the independent variables and to identify their impact on the dependent variable. However, nonlinear regression approach tries fitting a nonlinear equation by considering polynomial function of input variables. In this research, we unleashed statistical relationship as:

a) Two input variables (temperature and vacuum)
b) One output variable (light gas oil quantity

Once we were confident that this statistical relationship exists, we modelled the relationship mathematically and then used the model of prediction. MATLAB program was used to generate polynomial function of the chemical system.

3.5.1.3 Third Tool: Artificial Neural Networks (ANN) Model

The ANN is extremely a simplified and "shortened" model and emulation of the brain and essentially a function approximator that transform inputs to outputs to the best of its ability and according to Kevin (1997), artificial neural networks may be thought of as simplified models of the networks of neurons that occur naturally in the animal brain.

Here, our objective is to build a model that represent the real-world existing systems. This model should correctly mimic and predict the future results of the real system(s).

3.5.1.3.1 Network Architecture

In this study, we chose to implement the supervised network where ANN models are systematically trained to generate anticipated outputs against the inputs of the sample, preparing them in well manner to suit model in proper way.

Multilayer perceptron (MLP), which is a member of the feedforward neural networks (FFNN) family, was sought in our 1st model. That is FFNN shows more reliable and accurate results when applied to chemical processes. Finally, MATLAB software with neural network toolbox is used to build up the network models and generate the conforming MATLAB codes for further development.

3.5.1.3.2 ANN Topology

A two-layer feed-forward network can categorize any vectors in a random way by given sufficient number of network's neurons in its middle-hidden layer(s). Tangent sigmoid transfer function is the function used in the entire hidden layer while *pureline* transfer function is adapted by the output layer neurons.

In this study, we tried out four different ANN topologies so as to have the best possible ANN model. Every time we obtained new topology by altering the configuration. These topologies are: (2:12:3), (2:12:3), (2:30:3), (2:36:3). This configuration is shorthanded representation which means for instance, notation for this proposed 3-N-1 network topology is described as 2 is the total number of neural nodes in the input layer, N represents the total number of neurons located in the hidden layer, and 3 represents the number of neurons located in e Where Y is the dependent the output layer. The proposed network model is provided in Figure (3.3).

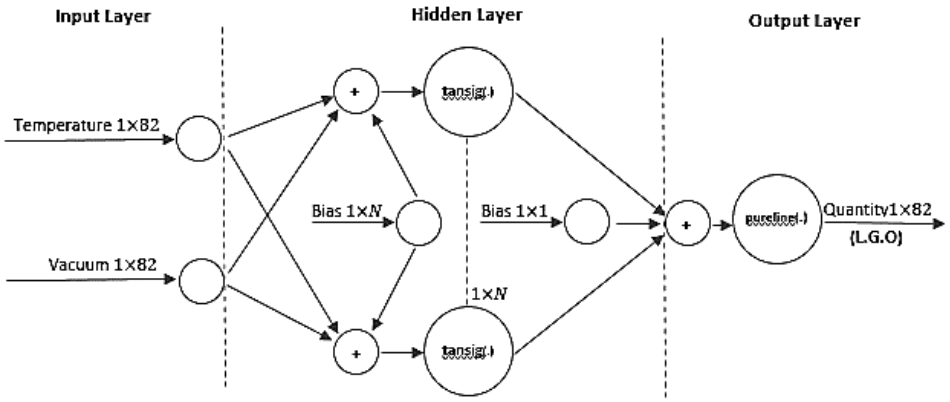


Figure 3.3 Proposed network model of the system

3.5.1.3.3 Training and Validation of the Proposed ANN Model

The purposes behind executing training and validation steps is to find and reach an optimal weight space by which mapping of input and target datasets is applicable. After the ANN was generated previously, pre-processed and normalized datasets are

arranged into the network. There are total 82 samples for each gas oil type (L.G.O, M.G.O, and H.G.O).

The datasets were sorted and divided into three data subsets, i.e., training, validation, and testing set for 10 different set data. That is to say, samples were proportionally configured as below:

1. 58 samples for training purposes (70%)
2. 12 samples for validation purposes (15%)
3. 12 samples for testing purposes (15%)

where

- Training: these are presented to the network during training, and the network is adjusted according to its error.
- Validation: validation is used to calculate and measure network generalization, and to pause training process by the time generalization stops improving.
- Testing: these holds no influence on training process and so offer an independent quantity of network performance during and after training (MATLAB, 2016).

3.5.1.3.4 Training Algorithms and Activation Functions

There are three well-known training algorithms and are successfully used in modelling various chemical process systems:

1. Levenberg-Marquardt (LM)
2. Scaled Conjugate Gradient (SCG)
3. Bayesian Regularization (BR)

These algorithms belong to the backpropagation algorithms group and are assigned for the multilayer perceptron (MLP) feedforward neural networks. They were separately investigated in our model and are compared finally. i.e., training algorithm which yield better performance in the model was selected to train the ANN model.

3.5.1.3.5 Activation Functions

In this study, two types of activation function (transfer function) where adapted.

1. Sigmoid function (tangential) – Hidden layers' neurons

We applied the sigmoid function in our hidden layer's neurons. Sigmoidal functions are function approximation and can best represent processes with non-linearity in nature and with dynamic behaviour. They are very popular in chemical process control applications. The sigmoid function consists of 2 functions, logistic and tangential. The values of logistic function range from 0 and 1 and -1 to +1 as in figure (3.4).

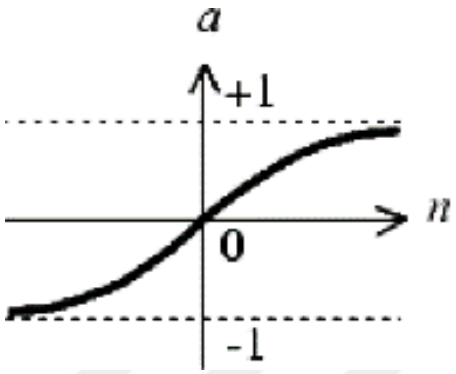


Figure 3.4 Sigmoid function

The mathematical representation of the sigmoidal function is described in Equation (3.3):

$$a = tansig(n) = \frac{2}{(1+e^{-2n})} - 1 \qquad (3.3)$$

2. Linear transfer function – Output layers' neurons

We applied the linear transfer function in our output layer's neurons. Linear transfer functions calculate a layer's output from its net input. Figure (3.5) shows this type of function.

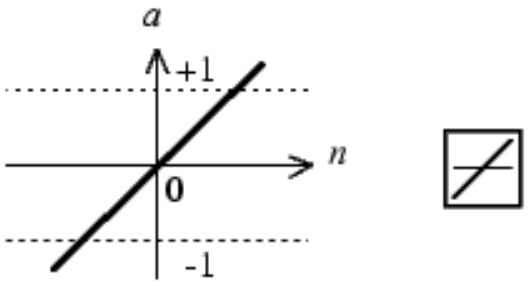


Figure 3.5 Linear transfer function

where *a = purelin (n) is the* transfer function is adapted by the output layer neurons aslo. The overall model design of network neuron including all layers is illustrated in Figure (3.5) as below:

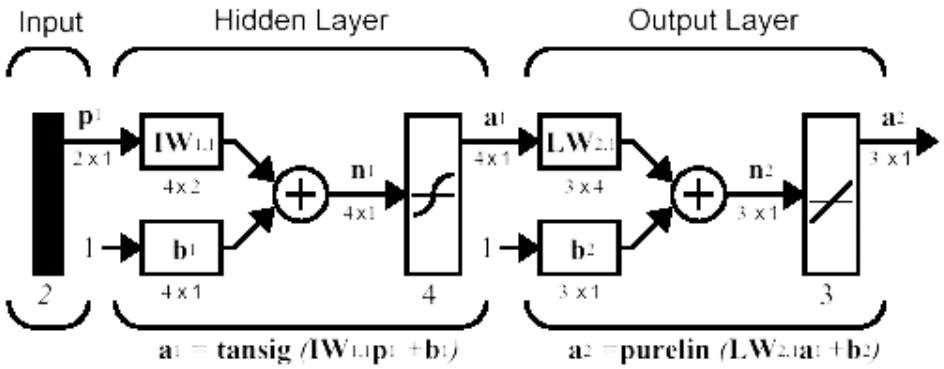


Figure 3.6 Overall network design

3.5.1.4 Fourth Tool: Ridge Regression (RR) Model

Ridge Regression, as was discussed in the previous chapter, belongs to the Regularization algorithms and it is a technique used when the data suffers from multicollinearity. To build ridge regression model, the ridge regression model equation need to be described in the form of matrix as in Equation (3.4).

$$Y = X\mathrm{B} + e \tag{3.4}$$

Where Y is dependent value while X variable holds for the independent value, and B is a constant value represents the regression coefficients that represents the error or residuals.

3.5.1.4.1 Standardization

Standardization of variables is the initial step in ridge regression analysis. Both dependent and independent variables' means are subtracted and divided by their standard deviations. Nevertheless, standardization can be performed statistically when deploying data into the NCSS software. Thus, in this study we created our model by using NCSS software.

3.5.2 Performance Metrics

We generated with the constructed models lists of forecasted or predicted readings and they were compared with their relevant actual values. Finally, we applied performance metrics on all four models' results using the following performance metrics:

1. MAD; The mean absolute deviation is the sum of absolute differences between the actual value and the forecast divided by the number of observations.

where Equation (3.5) represents the MAD metric:

$$\text{MAD} = \frac{\sum_{t=1}^{n} |A_t - F_t|}{n} \tag{3.5}$$

2. MSE; The mean square error is perhaps the most generally used error performance metric. All it does is penalizing any larger errors so as to squaring larger numbers imposes a greater impact than squaring smaller one. Lower values are better. Zero means no error.

where Equation (3.6) represents the MSE metric:

$$\text{MSE} = \frac{\sum_{t=1}^{n} (A_t - F_t)^2}{n} \tag{3.6}$$

3. RMSE; The root mean square error is the square root of the MSE.

where Equation (3.7) represents the RMSE metric:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n} (A_t - F_t)^2}{n}} \tag{3.7}$$

4. MAPE; The mean absolute percentage error is the average of absolute errors divided by actual observation values.

Where Equation (3.8) represents the MAPE metric:

$$\text{MAPE} = \frac{\sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|}{n} \times 100 \tag{3.8}$$

5. Regression R Values (R) - measures the correlation between outputs and targets. An R value of 1 means a close relationship, 0 a random relationship.

Where Equation (3.9) represents the R-value metric:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \times \sqrt{n(\sum y^2) - (\sum y)^2}} \quad (3.9)$$

3.6 Phase (3)

3.6.1 Mathematical Modelling

Mathematical modelling was beautifully explained by Meerschaert (2013) as the bridge that joins mathematics with the rest of the world. Modelling can avoid or reduce the need for costly, undesirable, or impossible experiments with the real world.

Curve fitting is the process of constructing a curve, or mathematical modelling, that has the best fit to a series of data points (Arlinghaus, 1994). Our objective here is to obtain all mathematical functions that represent or highlight the relationships among our parameters which will certainly help us understand better these relationships and assist us in constructing our mathematical models. To do so, we utilized the surface-fitting technique found in the MATLAB 2016.

There are two parameters categories in modelling inter-relationships with mathematical functions. Each function includes two separate independent variables and one dependent variable as below:

a) Quantity parameters

1. Temperature/vacuum - light gas oil (LGO)

2. Temperature/vacuum -moderate gas oil (MGO)

3. Temperature/vacuum -heavy gas oil (HGO)

b) Quality parameters

1. Temperature/vacuum - light gas oil (LGO)

2. Temperature/vacuum -moderate gas oil (MGO)

3. Temperature/vacuum -heavy gas oil (HGO)

3.7 Phase (4)

3.7.1 Finding Optimized Parameters

Optimization of models are designed to determine the values of the control variables which lead to the optimal outcome, given constraints of the problem (Meerschaert, 2013). In this phase, we found the optimized parameters through solving the mathematical model just generated in the third phase. To do so, we applied the *five-step* method on our model. These steps are as follow:

Step 1: Ask the question- Question must be asked in the form of mathematical terms for all assumed mathematical quantities like temperature (T) in degree Celsius (c^o), quantity in litres (L) and so on for the remaining identified variables involved in our system. In addition, assumptions and objective function must be states as well as stated below:

a) Variables:
   - T = x = temperature (degree Celsius, c^o)
   - V = y = vacuum (mmHg)
   - LGO($\boldsymbol{qty}$) = light gas oil/quantity (liter, l)
   - MGO($\boldsymbol{qty}$) = moderate gas oil/quantity (liter, l)
   - HGO($\boldsymbol{qty}$) = heavy gas oil/quantity (liter, l)
   - HGO($\boldsymbol{quly}$) = heavy gas oil/quality (liter/m3)
   - MGO($\boldsymbol{quly}$) = moderate gas oil/quality (liter/m3)
   - LGO($\boldsymbol{quly}$) = light gas oil/quality (liter/m3)

b) Assumption:
   - HGO($\boldsymbol{quly}$) = 855 liter/m3
   - 853.5 ≥ MGO($\boldsymbol{quly}$) ≥ 851.5 liter/m3
   - ≥ LGO($\boldsymbol{quly}$) ≥ 843 liter/m3
   - ≤ x ≤ 400 c^o
   - 655 ≤ y ≤ 730 mmHg

c) Objective:
   - Maximize P(total)

Step 2: Select the modelling approach and conditions- Problems can be stated in a standard form for which effective general solution procedure exists. We specified our

modelling approach as two variables optimization problem. i.e., temperature and vacuum are our two independent variables in the system.

Step 3: Formulate the Model-We need to take the question exhibited in step 1 and bring it back or reformulate to its original standard form selected in the step 2, so that we can apply the standard general solution procedure. In our case we reformulate the objective function Equation (3.10) as:

$$p = \mathrm{LGO}(qty) + \mathrm{MGO}(qty) + \mathrm{HGO}(qty) \tag{3.10}$$

where p is the total production of all final products

Step 4: Solve the model- In this step, we built and mathematical model by mathematical programming using constrained nonlinear optimizer found in MATLAB version 2016a, optimization tool. We utilized our mathematical functions obtained previously in step 1inluding the objective function as well as the constraints.

Step 5: Answer the question- Finally, we obtained the optimal parameters by which we answer our question to the maximizing of the system's final products under specified constraints. The targeted controllable variables (temperature and vacuum) are given certain range where they retain optimal condition in the system.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This chapter entails the outcomes obtained after deploying methodical procedures on our work. The chapter displays the results and analysis of the tools and techniques implemented. However, this chapter is divided, as with the previous chapter, into four sections. Each section introduces one phase of our work with the results and diagrams of course, i.e., section (1) exhibits results of data after pre-processing while section (2) displays all detail and results pertaining the four empirical tools implemented in our work with their performance comparison along with all diagrams and tables. Section (3) shows the mathematical model designed for our system associated with all graphs and mathematical functions. Finally, section (4) provides information on the optimality of the system's process parameters and conditions.

4.2 Data Analysis and Pre-Processing

There were 82 observations taken from the lube oil recycling unit's daily historical record on various timing. Working temperature and vacuum constitutes the inputs variables of the system while quantity and quality (for each light, moderate, and heavy gas oil) constitute the output variables of the system. The data set is tabulated in table (A.1) and (A.2) listed in appendix I.

We performed statistical analysis on our dataset using SPSS software in order to check their applicability ahead of processing. Tables (4.1) and (4.3) show statistical data analysis. As was observed in the table that we have obtained almost all attributes with slightly positively- skewed which indicates that data are skewed right-tail. Moreover, we obtained one single attribute with negative skewness. This means that our datasets are "somewhat" asymmetrical distributed but cannot be handle as a normal distribution.

Standard normal distribution has a kurtosis of zero and since our data showing near-to-zero values, our data distribution is nearly normally distributed. We can observe that all attributes retain small negative values of kurtosis except one attribute (Moderate Gas Oil, litre with 4.314).

Table 4.1 Statistical data of quantity parameters

		Temp (C°)	Vacuum (mmHg)	Light Gas Oil (litre)	Moderate Gas Oil (litre)	Heavy Gas Oil (litre)
N	Valid	82	82	82	82	82
	Missing	0	0	0	0	0
Mean		389.67	711.26	5844.68	4956.76	3932.59
Std. Error of Mean		.598	.747	25.711	25.662	23.642
Median		388.80	710.00	5788.50	4890.50	3982.50
Mode		383[a]	710	5665[a]	4728[a]	3730
Std. Deviation		5.411	6.764	232.826	232.380	214.084
Variance		29.278	45.748	54207.874	54000.55	45832.09
Skewness		.255	.267	1.035	1.663	-.099
Std. Error of Skewness		.266	.266	.266	.266	.266
Kurtosis		-.934	-.534	1.323	4.314	-1.475
Std. Error of Kurtosis		.526	.526	.526	.526	.526
Range		20	29	1243	1249	710
Minimum		380	700	5503	4673	3580
Maximum		400	729	6746	5922	4290

Table 4.2 Skewness of temperature and vacuum

		Temperature (C°)	Vacuum (mmHg)
N	Valid	82	82
	Missing	0	0
Skewness		.255	.267
Std. Error of Skewness		.266	.266
Kurtosis		-.934	-.534
Std. Error of Kurtosis		.526	.526

Table 4.3 Statistical data of quality parameters

		Temp (C°)	Vacuum (mmHg)	Light Ga Oil (litre/M3)	Moderate Gas Oil (litre/M3)	Heavy Gas Oil (litre/M3)
N	Valid	82	82	82	82	82
	Missing	0	0	0	0	0
Mean		389.67	711.26	844.226	852.323	854.671
Std. Error of Mean		.598	.747	.2356	.1132	.1779
Median		388.80	710.00	844.000	852.500	854.500
Mode		383[a]	710	843.0	852.5	854.0[a]
Std. Deviation		5.411	6.764	2.1331	1.0255	1.6106
Variance		29.278	45.748	4.550	1.052	2.594
Skewness		.255	.267	.341	.247	.081
Std. Error of Skewness		.266	.266	.266	.266	.266
Kurtosis		-.934	-.534	-.414	-.575	-.820
Std. Error of Kurtosis		.526	.526	.526	.526	.526
Range		20	29	9.5	4.5	6.5
Minimum		380	700	840.0	850.5	851.5
Maximum		400	729	849.5	855.0	858.0

However, positive kurtosis indicates a "heavy-tailed" distribution and negative kurtosis indicates a "light tailed" distribution.

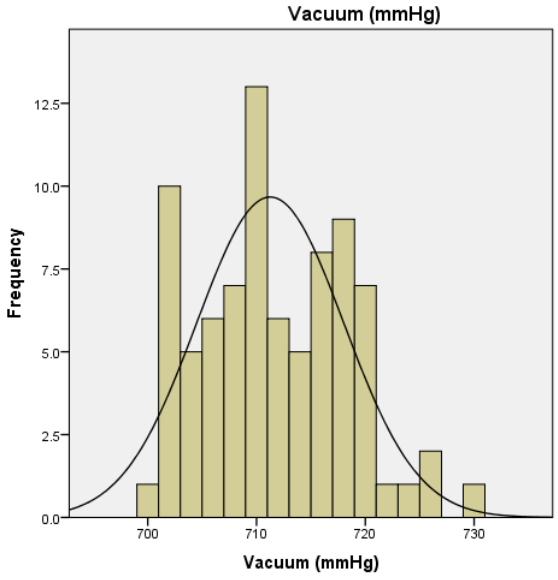


Figure 4.1 Asymmetrical distribution of parameters

Table (4.2) shows the skewness and the kurtosis of the inputs (temperature and vacuum) as they indicate that data distribution is slightly skewed to the right tail as shown in Figure (4.1).

4.2.1 Data Cleaning

In the pre-processing of the data, we first detect and "clean out" all corrupt or inaccurate records in the datasets of both input and output e.g. missing data, outlier or noisy data, and anomalies, which were treated and removed from the dataset and now we are able to execute data sampling process.

4.2.2 Data Transformation

We transform our data to data proper for analysis and thus, we perform linear transformation of the original data using min-max normalization method. This method suggests that we find our maximum and minimum values from our dataset and assign them to a range between [1, -1].

In order to bring all values into the range [-1,1] of all input and output data, we deployed the mathematical Equation (4.1) as below:

$$x' = 2\frac{x - \min(x)}{\max(x) - \min(x)} - 1 \qquad (4.1)$$

The reason to apply normalization process on the input values is that when data are fed into the neuron, it may need to integrate processes which may render very high results. This will cause the activation function, which is the tangent sigmoid function in this case, perform badly when attempting to tackle small changes in input data and thus, it may lose its sensitivity. As a result, we obtain the newly normalized data which is ready for processing.

4.3 Construction of Predictive Models

In order to control processes in any chemical plant, it is very critical to understand the processes and their behaviour. Predictive modelling is an ensemble of statistical algorithms coded in a statistical tool, which when applied on historical data, outputs a mathematical function (or equation). It can in-turn be used to predict outcomes based on some inputs (on which the model operates) from the future to drive a goal in business context or enable better decision making in general.

4.3.1 First Tool: Multiple Linear Regression (MLR) Model

In this research, we aimed at discovering the statistical relationship between variables involved in the controlling of the system's processes. Therefore, we selected one single pattern including 2 independent input variables (temperature and vacuum) and one output variable (light gas oil quantity) as shown in Figure (4.2), and once we were confident that this statistical relationship exists, we modelled the relationship mathematically and then used the model as a predictive model. Excel and SPSS programs were used to generate multiple regression model as a linear function of two variables in first-order class as shown in Equation (4.2):

$$f(x, y) = -8015 + 36.45x - 0.485y \tag{4.2}$$

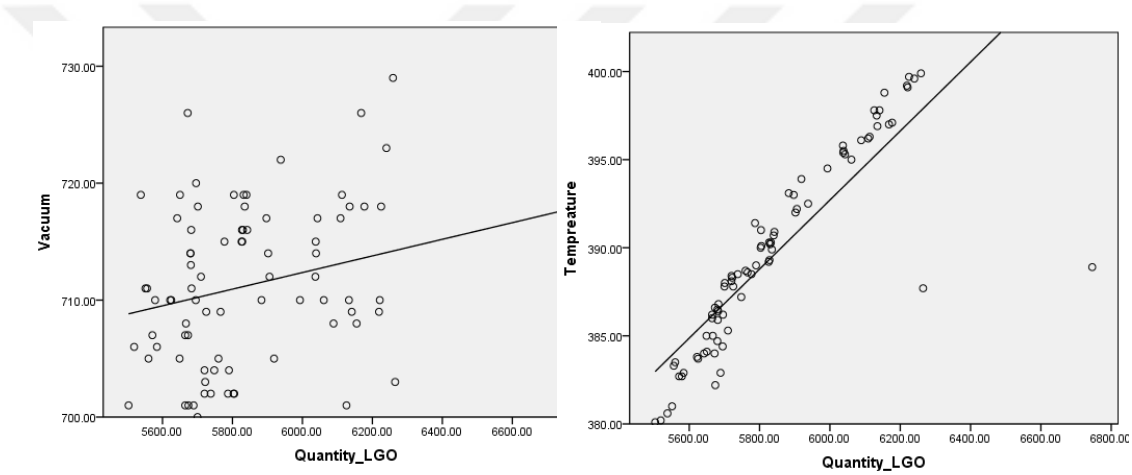


Figure 4.2 MLR model of vacuum and temperature versus quality

As shown in Table (4.4), actual and predicted values were compared and the performance metric is applied on the linear model as illustrated in Figure (4.3).

Table 4.4 Performance metrics of multiple linear regression model

Multiple Linear Regression Model Prediction Performance				
n	**MAD**	**MSE**	**RMSE**	**MAPE**
82	329	16186	402	**5.62**

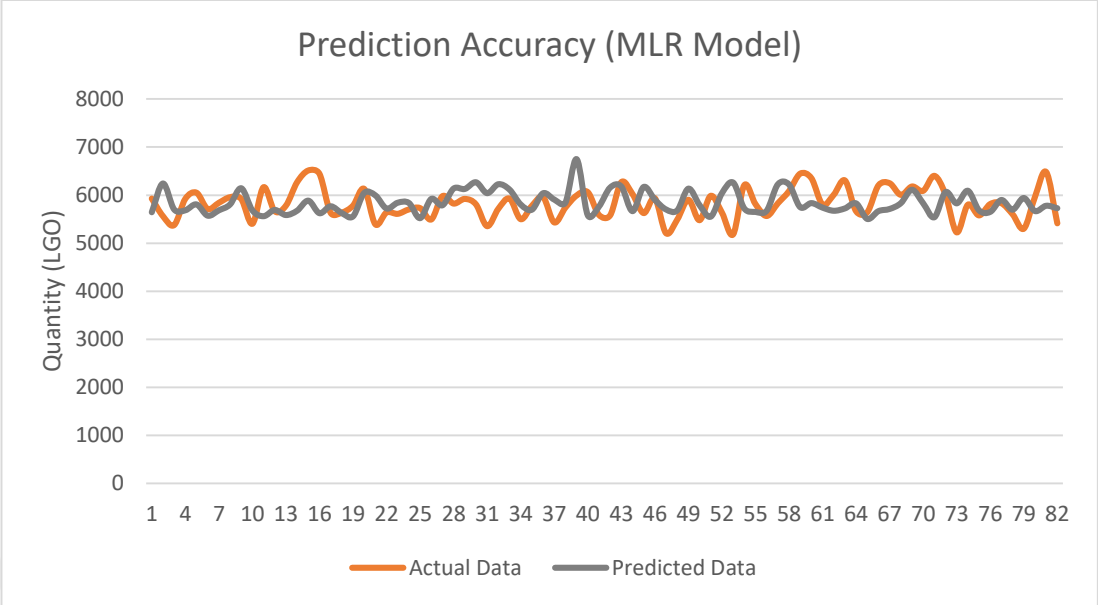


Figure 4.3 Prediction accuracy of multiple linear regression model

4.3.2 Second Tool: Multiple Nonlinear Regression (MNLR) Model

Nonlinear regression models resemble their previous type. However, they attempt to adapt better representation of variables inter-relation by using higher degree of functions, like polynomial function. In this research, we unleashed statistical relationship between 2 input variables (temperature and vacuum) and one output variable (light gas oil quantity) and formulated this relation with second- order polynomial function. MATLAB program was used to generate this function directly. The multi regression model with second-order and two variables can be written as in Equation (4.3):

$$f(x,y) = 83363 - 558.3 \times x + 65 \times y + 0.7727 \times x^2 - 0.01111 \times x \times \mathrm{y} - 0.04106 \times y^2 \quad (4.3)$$

As shown in Table (4.5), actual and predicted values were compared and the performance metric is applied on the linear model as illustrated in Figure (4.4).

Table 4.5 Performance metrics of multiple nonlinear regression model

Multiple Nonlinear Regression Model Prediction Performance				
n	**MAD**	**MSE**	**RMSE**	**MAPE**
82	304	10870	329	**5.18**

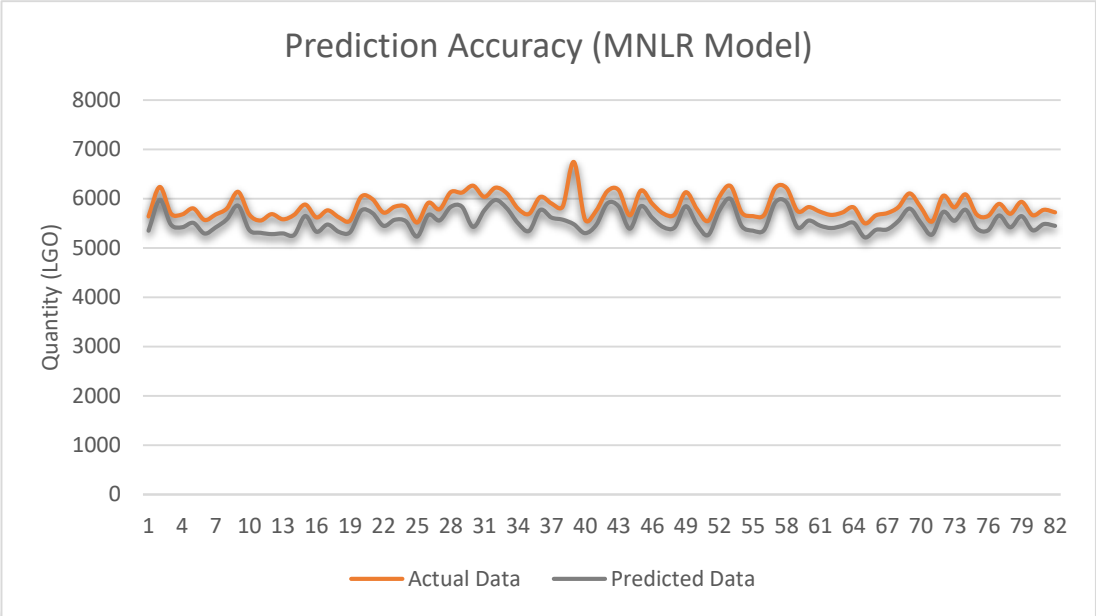


Figure 4.4 Prediction accuracy of multiple nonlinear regression model

4.3.3 Third Tool: Artificial Neural Network (ANN) Model

In the ANN modelling our objective is to build a model that represent the real-world existing systems. This model should correctly mimic and predict the future results of our real system.

The advantage of the ANN is to learn the relationship between inputs and outputs by non-statistical approach (black box). These ANN-based methodologies don't require any predefined mathematical models. They are a universal nonlinear function approximator.

4.3.3.1 ANN Topology

In this study, we tried out four different ANN topologies in order to gain the best possible ANN model configuration. Every time we obtained new topology by altering the number of neurons in the hidden layer(s). These topologies are:

a) (2:12:1)
b) (2:20:1)
c) (2:30:1)
d) (2:36:1)

These four different arrangements were created as shown in Figure (4.5).

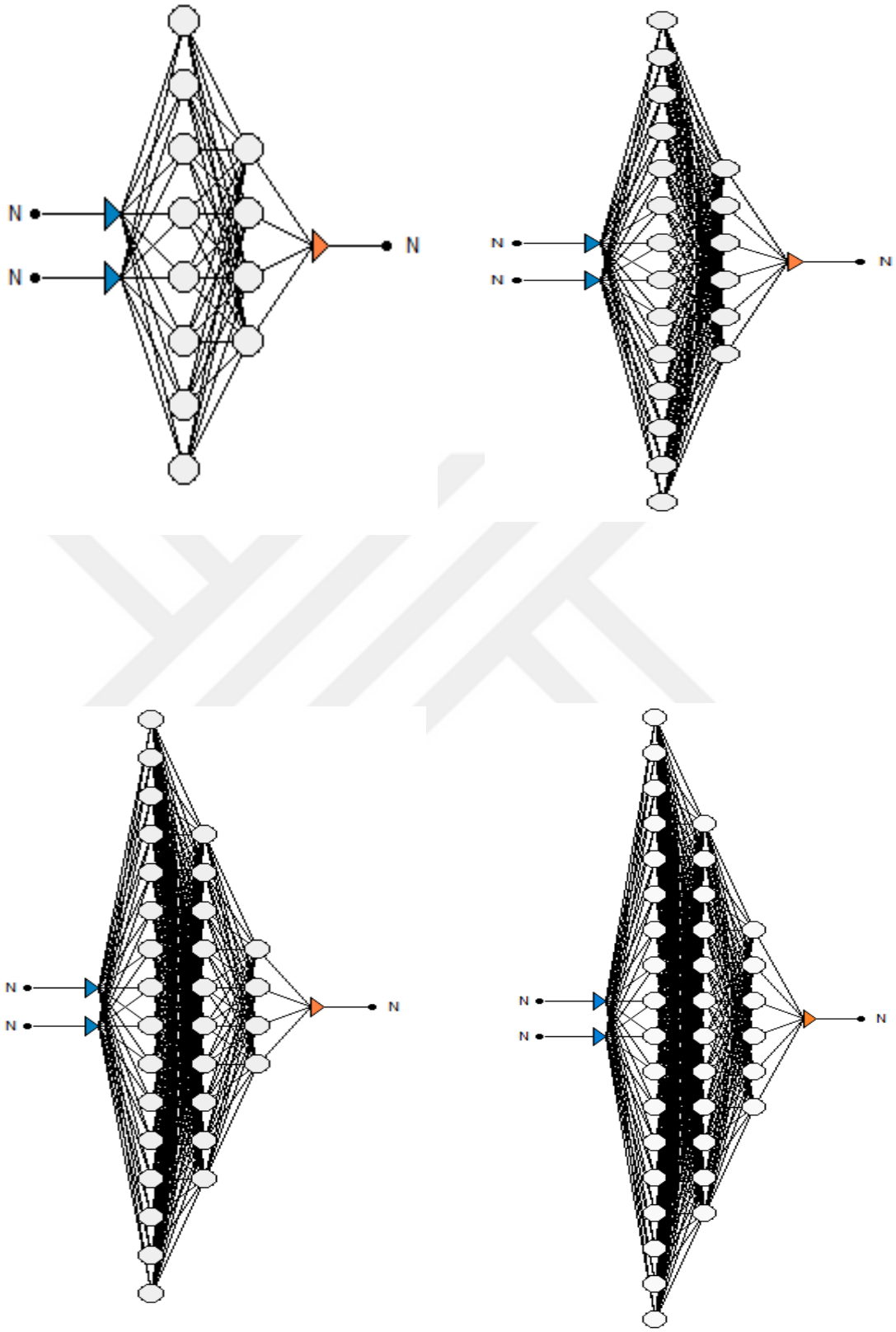


Figure 4.5 Four different ANN topologies: 2-12-1, 2-20-1, 2-30-1 and 2-36-1

4.3.3.2 Training Algorithms and Activation Functions

Initially, we had 10 different samples which were randomly picked. These samples contain 82 observations for each.

The dataset of each sample was randomly divided into three different portions for all 10 samples as below:

- **58** samples for training purposes (70%)
- **12** samples for validation purposes (15%)
- **12** samples for testing purposes (15%)

Finally, we tried out our 10 random samples on the four proposed ANN topologies with the three-well-known training algorithm as below:

- Levenberg-Marquardt (LM)
- Scaled Conjugate Gradient (SCG)
- Bayesian Regularization (BR)

Table (4.6) shows an example of one sample that has been randomly constructed 10 times and trained 10 times with 58, 12, and 12 sampling assignment for training, validation and testing respectively using the Scaled Conjugate Gradient (SCG) as training algorithm on the 2-12-1 ANN topology.

Table 4.6 10-time training of ANN model

	Train	**Train # 1**	**Train # 2**	**Train # 3**	**Train # 4**	**Train # 5**	**Train # 6**	**Train # 7**	**Train # 8**	**Train # 9**	**Train # 10**	**Test (mean)**
MSE	Train	0.029	0.050	0.033	0.030	0.046	0.033	0.029	0.031	0.052	0.030	0.0367
	Valid	0.020	0.020	0.017	0.013	0.016	0.018	0.015	0.016	0.019	0.016	0.0172
	Test	0.129	0.149	0.139	0.131	0.137	0.148	0.137	0.148	0.150	0.138	0.1410
R-Value	Train	0.910	0.892	0.910	0.895	0.908	0.920	0.910	0.909	0.898	0.903	0.9061
	Valid	0.949	0.967	0.959	0.962	0.949	0.953	0.968	0.959	0.949	0.959	0.9579
	Test	0.788	0.790	0.783	0.802	0.795	0.780	0.792	0.787	0.801	0.790	0.7911
Epoch Iteratio		37	39	43	39	44	38	35	37	45	42	40
Time (second)		4	10	9	4	6	5	6	4	7	3	6
R-Value (All)		0.891	0.885	0.9	0.879	0.872	0.881	0.884	0.881	0.884	0.88	0.884

4.3.3.3 ANN Training

(1) 2-12-1 topology:

Table 4.7 (2-12-1) network performance versus training algorithms

	Levenberg-Marquardt (LM)			**Bayesian Regularization (BR)**			**Scaled Conjugate Gradient (SCG)**		
	Train	**Valid**	**Test**	**Train**	**Valid**	**Test**	**Train**	**Valid**	**Test**
MSE	0.04694	0.05914	**0.03558**	0.04577	**0.00000**	0.01339	0.03673	**0.01721**	0.01410
R-Value	**0.91326**	0.88011	0.90796	0.91466	0.00000	**0.96193**	0.90618	**0.95793**	0.79112
Epoch Iteration	9			652			40		
Time (second)	≈1			44			6		
R-Value (All)	0.909			0.919			0.884		

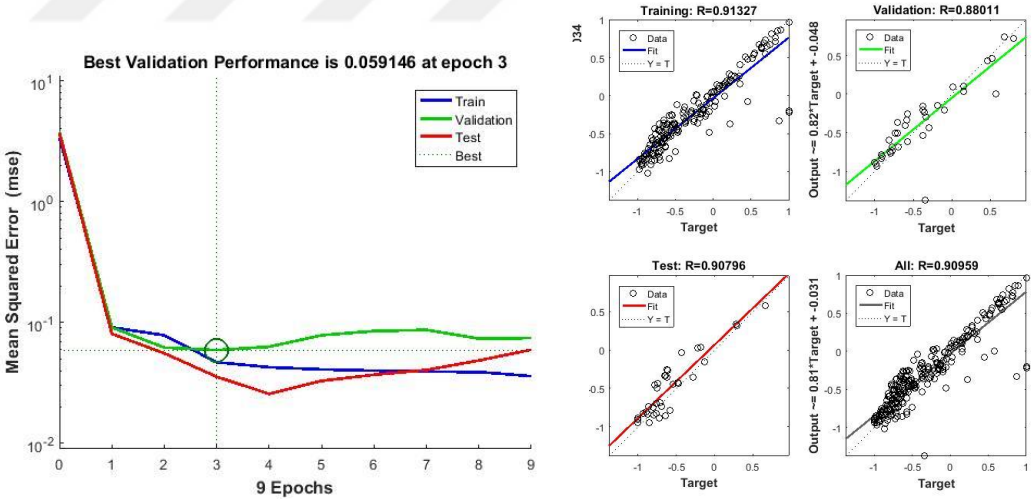


Figure 4.6 (2-12-1) best validation performance and R-value

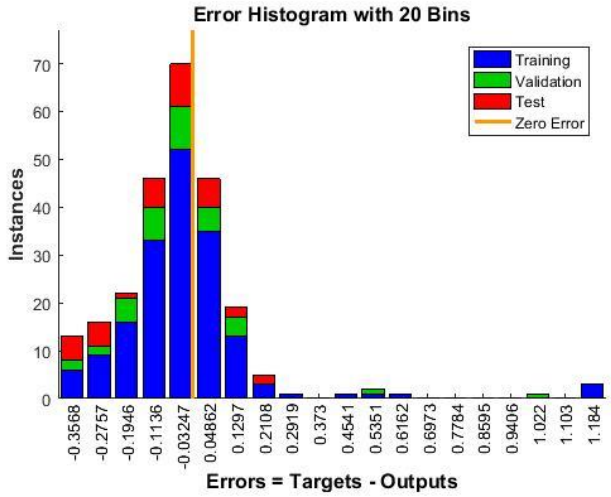


Figure 4.7 (2-12-1) error histogram

(2) 2-20-1 topology:

Table 4.8 (2-20-1) network performance versus training algorithms

	Levenberg-Marquardt (LM)			**Bayesian Regularization (BR)**			**Scaled Conjugate Gradient (SCG)**		
	Train	**Valid**	**Test**	**Train**	**Valid**	**Test**	**Train**	**Valid**	**Test**
MSE	0.07955	0.06673	**0.06449**	0.04627	0.00000	**0.01186**	0.03990	0.06977	**0.01629**
R-Value	**0.94839**	0.85658	0.88099	0.91125	0.00000	**0.98336**	0.92403	0.90415	**0.96409**
Epoch Iteration	12			727			30		
Time (second)	≈1			≈48			≈2		
R-Value (All)	0.924			0.919			0.920		

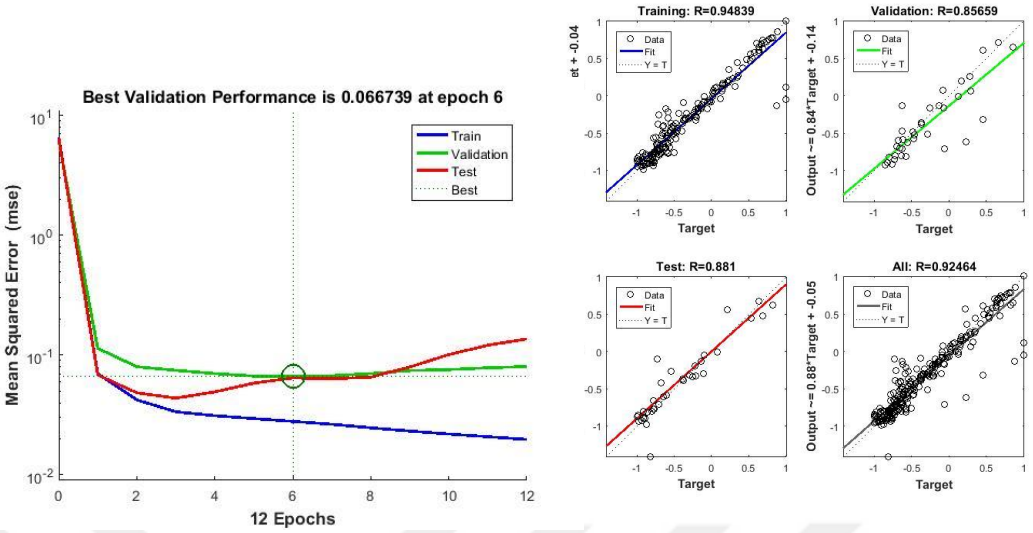


Figure 4.8 (2-20-1) best validation performance and R-value

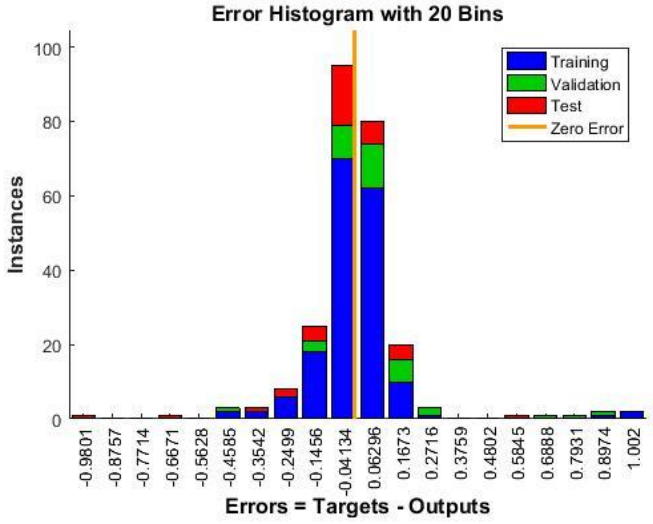


Figure 4.9 (2-20-1) error histogram

(3) 2-30-1 topology:

Table 4.9 (2-30-1) network performance versus training algorithms

	Levenberg-Marquardt (LM)			**Bayesian Regularization (BR)**			**Scaled Conjugate Gradient (SCG)**		
	Train	**Valid**	**Test**	**Train**	**Valid**	**Test**	**Train**	**Valid**	**Test**
MSE	**0.02501**	0.03376	0.07602	0.03868	**0.00000**	0.04918	0.05505	**0.01779**	0.06974
R-VALUE	**0.95221**	0.94047	0.83587	**0.92583**	0.00000	0.87852	0.89722	**0.97179**	0.82821
Epoch Iteration	9			1000 (limit)			32		
Time (second)	≈1			≈67			≈3		
R-Value (All)	0.935			0.921			0.899		

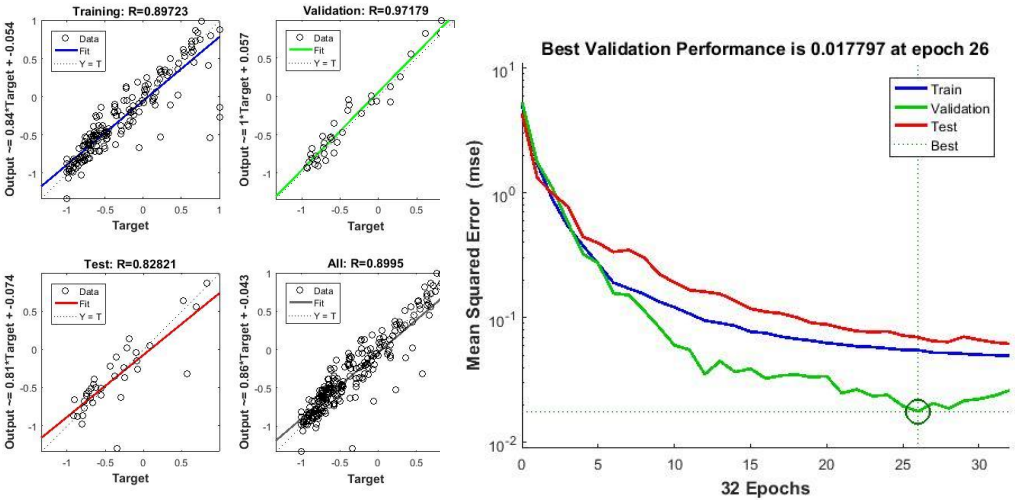


Figure 4.10 (2-30-1) best validation performance and R-value

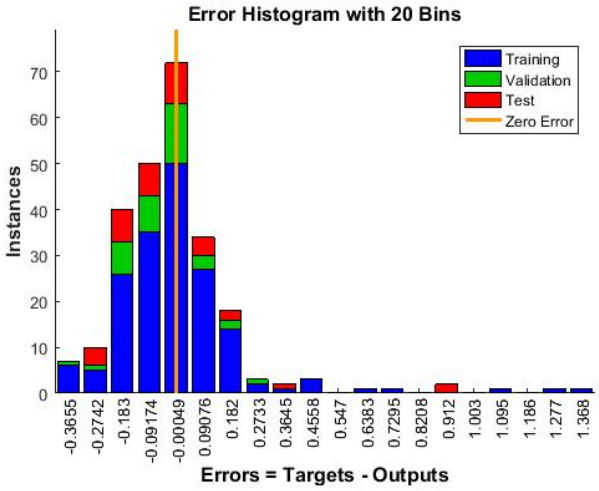


Figure 4.11 (2-30-1) error histogram

(4) 2-36-1 topology:

Table 4.10 (2-36-1) network performance versus training algorithms

	Levenberg-Marquardt (LM)			**Bayesian Regularization (BR)**			**Scaled Conjugate Gradient (SCG)**		
	Train	**Valid**	**Test**	**Train**	**Valid**	**Test**	**Train**	**Valid**	**Test**
MSE	**0.02633**	0.02734	0.08286	0.03902	**0.00000**	0.05323	0.03989	0.03385	**0.11178**
R-Value	**0.95086**	0.93255	0.84340	**0.92442**	0.00000	0.88013	0.92744	**0.93486**	0.74926
Epoch Iteration	8			1000			31		
Time (second)	≈1			≈67			≈2		
R-Value (All)	0.933			0.919			0.903		

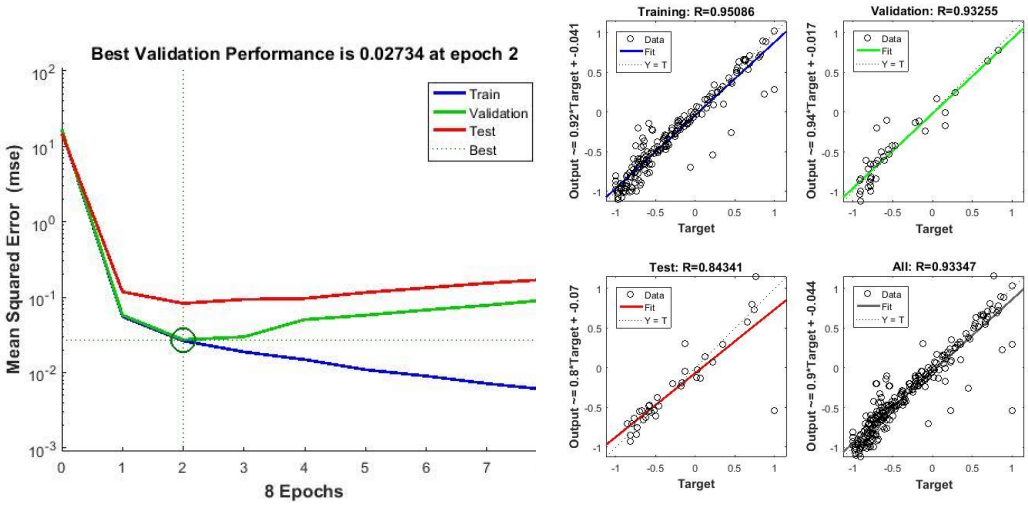


Figure 4.12 (2-36-1) best validation performance and R-value

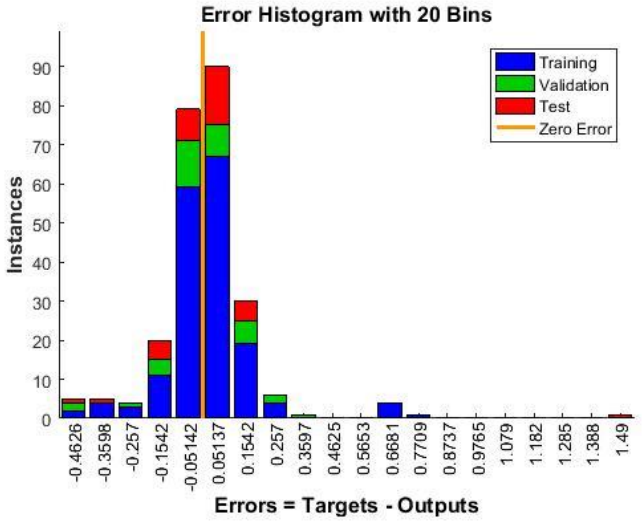


Figure 4.13 (2-36-1) error histogram

4.3.3.4 ANN Result Analysis

As shown in tables (4.7), (4.8), (4.9) and (4.10) from the previous experiments and are combined altogether in Table (4.11) for comparison purposes. In general, validation or testing sets have higher prediction accuracy (or almost slightly below) as was compared to the set of training phase. This means no absolute or clear overfitting on this data (training data) is experienced and hence, the generalization point is reached. Yet the SCG algorithm showed high degree of generalization and hence no overfitting at all. The validation set performance always indicates and reflects the accurateness of forecasting ability in all training algorithms. However, SCG algorithm showed the best

value of mean square error (MSE) in respect to the validation with 0.01779. Interestingly, these MSE validation values are very similar and were achieved in the neural network structure whose topology falls within 30 neurons category. However, BR algorithm exhibited best performance regarding the testing set with 0.01186. Furthermore, in respect to the topology of the neural network's neurons, we noticed that 2-30-3 configuration showed the best correlation among the target data and the model's output data with 0.935, 0.921 and 0.889 pertaining the LM, BR and SCG algorithms respectively. As a result, it can be claimed that LM with 2-30-3 network configuration showed better correlation followed by SCG and BR algorithms.

Eventually, with respect to time, it is obviously clear that BR algorithm took significant time before reaching its best validation performance associated with big number of epochs. On the other hand, LM algorithm showed very little time to

Table 4.11 Summary of training algorithms with respect to various topologies

Performance metrics	**R-value (All)**			**MSE**								
Algorithm	Levenberg-Marquardt (LM)	Bayesian Regularization	Scaled Conjugate Gradient (SCG)	**Levenberg-Marquardt (LM)**			**Bayesian Regularization (BR)**			**Scaled Conjugate Gradient (SCG)**		
Topology				**Train**	**Valid**	**Test**	**Train**	**Valid**	**Test**	**Train**	**Valid**	**Test**
2-12-1	0.90	0.91	0.884	0.046	0.059	0.035	0.045	0.000	0.013	0.0367	**0.0172**	0.0172
2-20-1	0.92	0.91	0.920	0.079	0.066	0.064	0.046	0.0000	**0.011**	0.039	0.069	0.016
2-30-1	**0.93**	**0.92**	**0.899**	0.025	0.0337	0.076	0.038	0.0000	0.049	0.055	**0.017**	0.069
2-36-1	0.93	0.91	0.903	0.026	**0.0273**	0.082	0.039	0.000	0.053	0.039	**0.033**	0.111
Best Validation	**0.93**	**0.92**	**0.924**	**0.02734**			-			**0.01779**		
Max time	1	67	3									

accomplish its task. The same case for the SCG algorithm with few seconds extra taken as compared to the LM algorithm.

To compare each model's predictivity performance, simulation test was conducted on the three models which were trained by the following algorithms: (1) Levenberg-Marquardt (LM), (2) Bayesian Regularization (BR) and (3) Scaled Conjugate Gradient (SCG). Thus, we simulated our best generated model with some real dataset. We picked randomly 12 data samples representing the quantity of the light gas oil attribute from our data Table (4.12).

Table 4.12 Prediction accuracy of training algorithms versus actual data

Quantity observations		Scaled Conjugate Gradient (SCG)	Levenberg-Marquardt (LM)	Bayesian Regularization (BR)
N	**Actual (litre)**	**Predicted (litre)**	**Predicted (litre)**	**Predicted (litre)**
1	6240	6313	6366	6501
2	6259	6487	6449	6441
3	5790	5970	5930	5963
4	5777	5619	5634	5667
5	5725	5834	6028	5531
6	5696	5837	5582	5915
7	5672	5503	5832	5596
8	5584	5809	5513	5877
9	5503	5633	5727	5709
10	5680	5861	5601	5838
11	6089	6098	6175	6018
12	5556	5653	5688	5742

There are two groups of data i.e. actual and predicted values. We simulated our data in our models and obtained our predicted values. Hence, predicted values were compared to the actual values obtained from the real system.

As we notice from Table (4.13) and Figures (4.14), (4.15) and (4.16), SCG algorithm showed best performance and is able to predict values more accurately as compared to other algorithms. SCG secured MAPE value equal to 2.42 followed by LM and BR algorithms with values 2.55 and 3.07 respectively. Figure (4.17) shows the overall comparison among the three training algorithms.

Table 4.13 Performance metrics for various training algorithms

	Levenberg-Marquardt (LM)	**Scaled Conjugate Gradient (SCG)**	**Bayesian Regularization (BR)**
n	12	12	12
MAD	147	142	178
MSE	25768	23787	35654
RMSE	161	154	189
MAPE	2.55	**2.42**	3.07

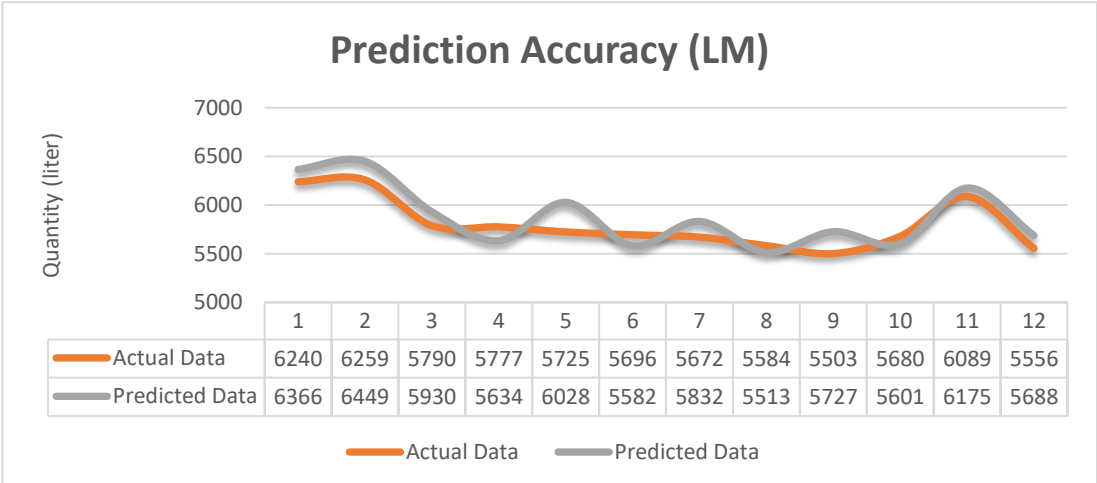
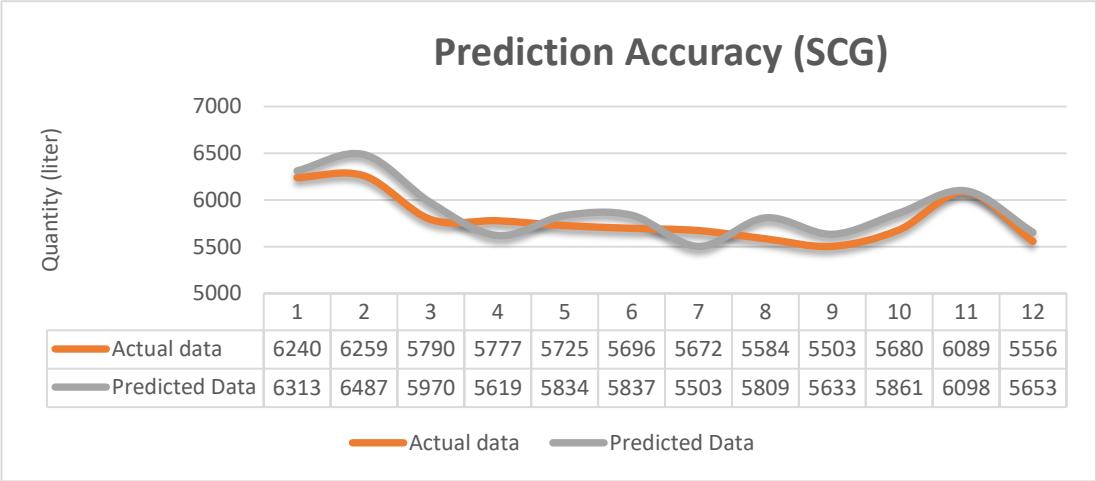
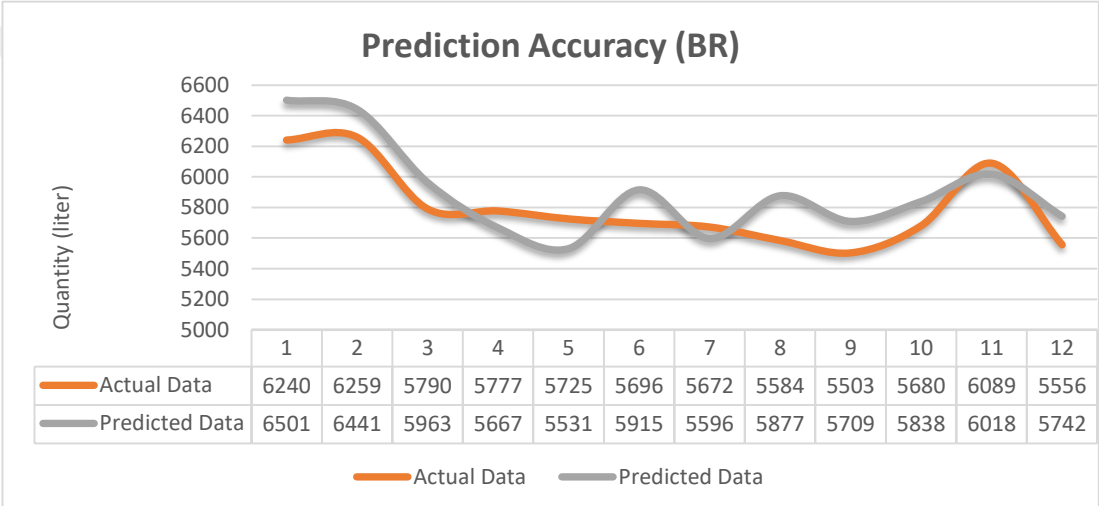


Figure 4.14 Prediction accuracy of Levenberg-Marquardt training algorithm



	1	2	3	4	5	6	7	8	9	10	11	12
Actual data	6240	6259	5790	5777	5725	5696	5672	5584	5503	5680	6089	5556
Predicted Data	6313	6487	5970	5619	5834	5837	5503	5809	5633	5861	6098	5653

Figure 4.15 Prediction accuracy of Scaled Conjugate training algorithm



	1	2	3	4	5	6	7	8	9	10	11	12
Actual Data	6240	6259	5790	5777	5725	5696	5672	5584	5503	5680	6089	5556
Predicted Data	6501	6441	5963	5667	5531	5915	5596	5877	5709	5838	6018	5742

Figure 4.16 Prediction accuracy of Bayesian Regularization training algorithm

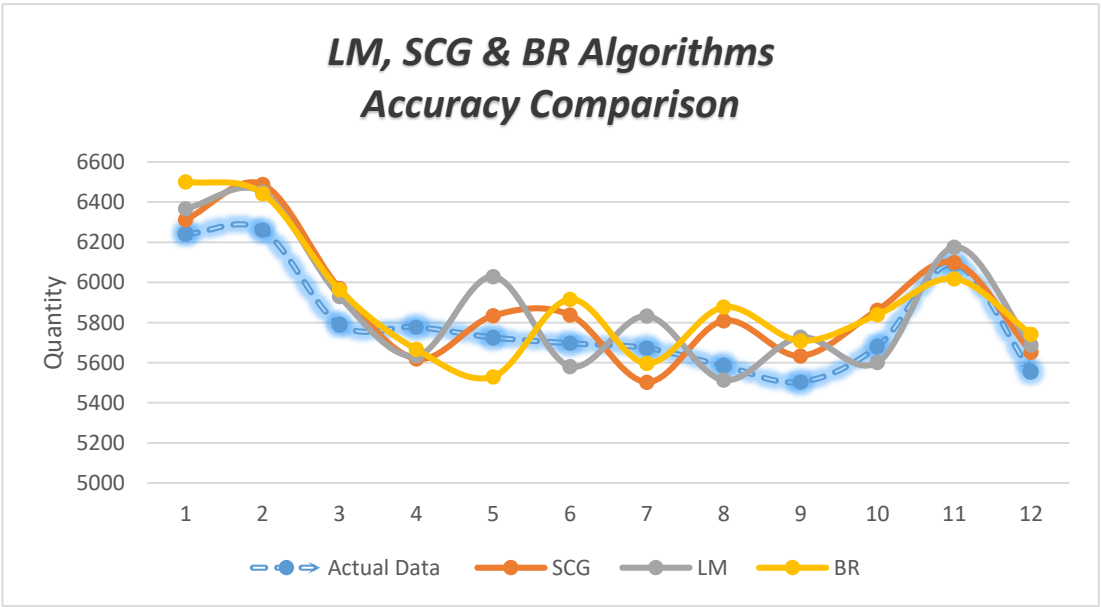


Figure 4.17 Prediction accuracy of LM, BR and SCG training algorithms

4.3.4 Fourth Tool: Ridge Regression (RR) Model

Ridge Regression belongs to the Regularization algorithms and it is a technique used when the data suffers from multicollinearity where independent variables are highly correlated. However, it is the first step toward standardizing obtained variables of both dependent and independent variables. Nevertheless, standardization can be performed statistically when deploying data into the NCSS interface page. Table (4.14) shows the RR model performance metrics while Figure (4.18) shows the prediction accuracy of RR model.

Table 4.14 Performance metrics of Ridge Regression model

Ridge Regression Model Prediction Performance				
n	**MAD**	**MSE**	**RMSE**	**MAPE**
82	192	75874	275	**3,27**

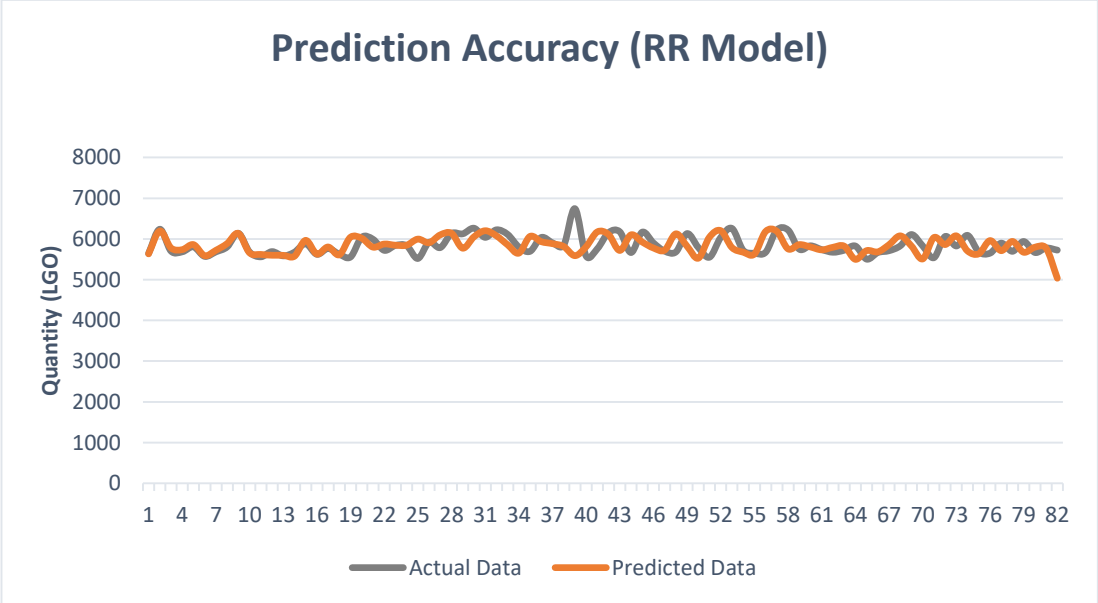


Figure 4.18 Prediction accuracy of ridge regression model

4.4 Construction of Mathematical Model

As explained in the previous chapter, prior to the establishment of a mathematical model for our system, we aimed at obtaining all possible mathematical functions that can typically represent the processes in the system taking into consideration all constraints available in the system. Therefore, it is necessary to highlight the relationships among the parameters which will certainly help us understand better these relationships and assist us in constructing the mathematical models. To do so, we utilized the surface-fitting technique found in the MATLAB software.

The mathematical functions are listed as below:

4.4.1 Mathematical Functions of The Quantity Parameters

1. Temperature & vacuum = Light Gas Oil (LGO)

$$\text{LGO}(liter) = f(x,y) = 83363 - 558.3 \times x + 65 \times y + 0.7727 \times x^2 - 0.01111 \times x \times y - 0.04106 \times y^2 \quad (4.4)$$

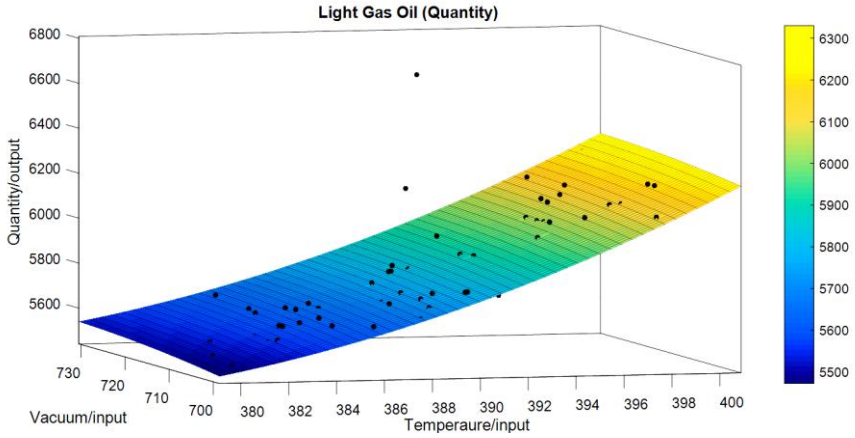


Figure 4.19 Surface-fitting of between input variables (temperature and vacuum) and output variable (light gas oil) in term of its quantity

2. Temperature & vacuum = Moderate Gas Oil (MGO)

$$MGO(liter) = f(x,y) = 96245 - 606.1 \times x + 55.45 \times y + 0.7509 \times x^2 - 0.07234 \times x \times y - 0.05733 \times y^2 \quad (4.5)$$

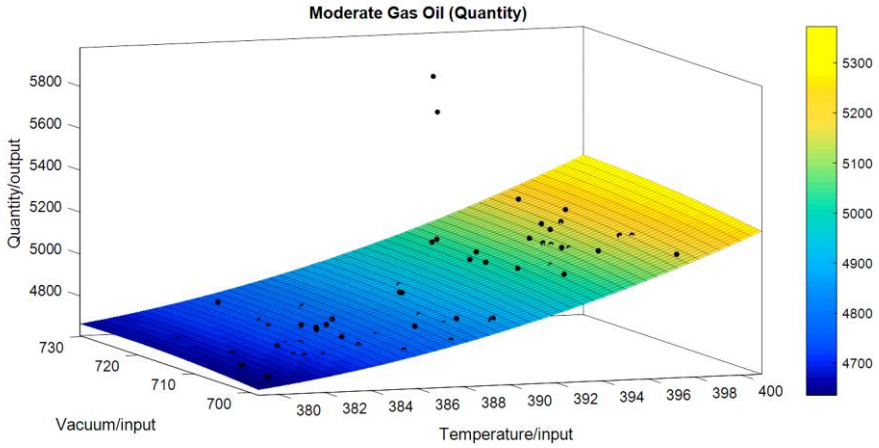


Figure 4.20 Surface-fitting of between input variables (temperature and vacuum) and output variable (moderate gas oil) in term of its quantity

3. Temperature & vacuum = Moderate Gas Oil (HGO)

$$HGO(liter) = f(x,y) = -70258 + 659.7 \times x - 175.2 \times y - 1.036 \times x^2 + 0.2614 \times x \times y - 0.05199 \times y^2 \quad (4.6)$$

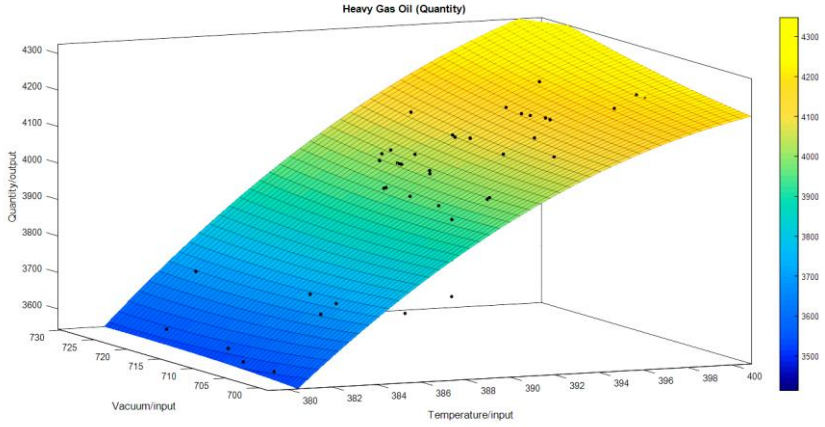


Figure 4.21 Surface-fitting of between input variables (temperature and vacuum) and output variable (heavy gas oil) in term of its quantity

4.4.2 Mathematical Functions of the Quality Parameters

4. Temperature & vacuum = Light Gas Oil (LGO)

$$LGO(liter/m^3) = f(x,y) = 1471 - 0.6577 \times x - 1.63 \times y - 0.0005 \times x^2 + 0.00199 \times x \times y + 0.000624 \times y^2 \quad (4.7)$$

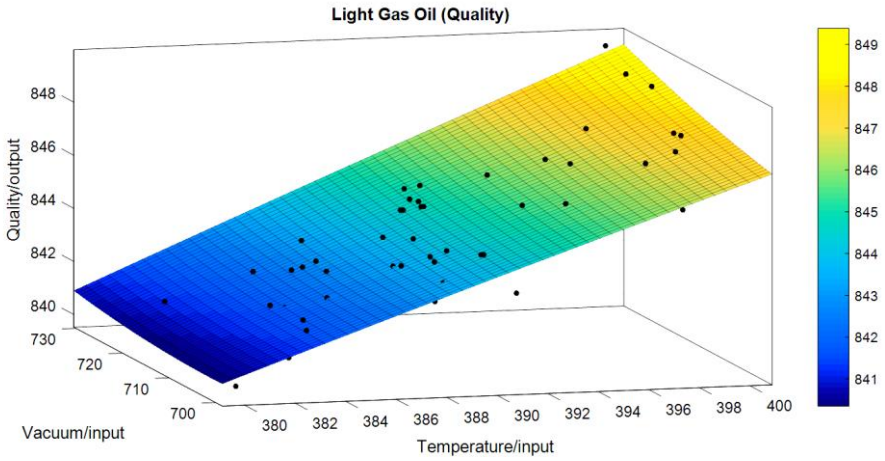


Figure 4.22 Surface-fitting of between input variables (temperature and vacuum)

5. Temperature & vacuum = Moderate Gas Oil (MGO)

$$MGO(liter/m^3) = f(x,y) = 1494 - 0.7056 \times x - 1.521 \times y - 0.00057 \times x^2 + 0.00185 \times x \times y + 0.00057 \times y^2 \quad (4.8)$$

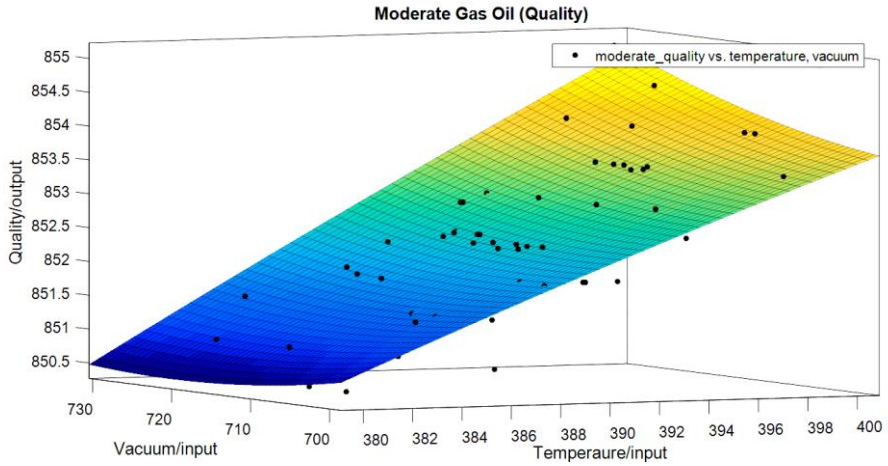


Figure 4.23 Surface-fitting of between input variables (temperature and vacuum)

6. Temperature & vacuum = Heavy Gas Oil (HGO)

$$HGO(liter/m^3) = f(x,y) = 125.4 + 2.785 \times x + 0.349 \times y - 0.00253 \times x^2 - 0.0008 \times x \times y - 0.0000027 \times y^2 \quad (4.9)$$

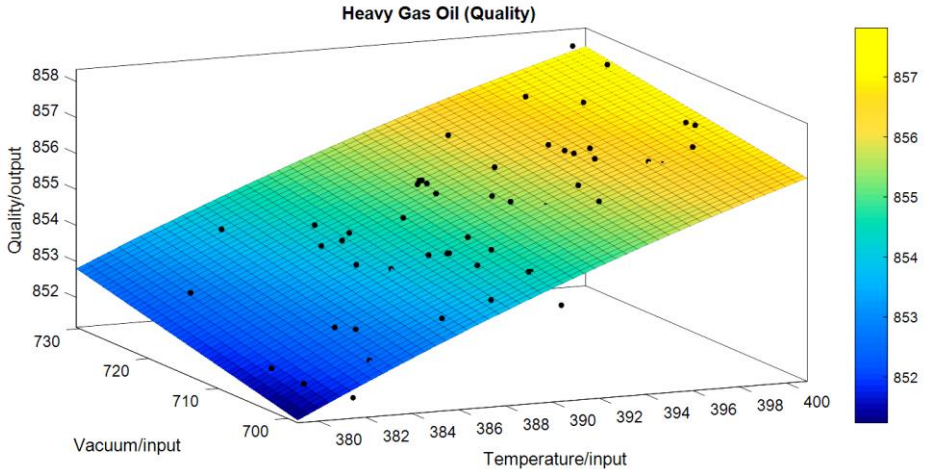


Figure 4.24 Surface-fitting of between input variables (temperature and vacuum)

4.5 Finding Optimized Parameters

In this phase, we found the optimized parameters through solving the mathematical model just generated in the third phase. To do so, we applied the *five-step method* on our model. These steps are as follow:

Step 1: Determine all variables, assumption, constraints and objective

a) Variables:

- T = x = temperature (degree Celsius, c^o)
- V = y = vacuum (mmHg)
- LGO(qty) = light gas oil/quantity (liter, l)
- MGO(qty) = moderate gas oil/quantity (liter, l)
- HGO(qty) = heavy gas oil/quantity (liter, l)
- HGO($quly$) = heavy gas oil/quality (liter/m3)
- MGO($quly$) = moderate gas oil/quality (liter/m3)
- LGO($quly$) = light gas oil/quality (liter/m3)

b) Assumption & constraints:

- Heavy gas oil, HGO($quly$) = 855 liter/m3 (4.10)
- 853.5 ≥ Moderate gas oil, MGO($quly$) ≥ 851.5 liter/m3 (4.11)
- 846 ≥ Light gas oil, LGO($quly$) ≥ 843 liter/m3 (4.12)
- 380 ≤ x ≤ 400 c^o

- $655 \leq y \leq 730$ mmHg

c) Objective:

- Maximize P(total)

Where, P is the total production and our objective is to bring these end products up to 70%. In other Word:

$$max(x, y) = \textit{Heavy gas oil}\,(\text{HGO}) + \textit{Moderate gas oil}\,(\text{MGO}) + \textit{Light gas oil}\,(\text{LGO}) \geq 70\% \quad (4.13)$$

Step 2: Select the modelling approach

We specified our modelling approach as two variables optimization problem. Where temperature and vacuum are our two independent variables in the system and quantity of light, moderate and heavy gas oil respectively are our three dependent variables.

Step 3: Formulate the model

We used the questions obtained in step (1) and reformulate them in a standard form selected in the step 2, so that we can apply the standard general solution procedure. In our case we reformulate the objective function as in Equation (4.14):

$$\text{P} = \text{LGO(qty)} + \text{MGO(qty)} + \text{HGO(qty)} \quad (4.14)$$

where P is the total production of all final products.

Step 4: Solve the model

We built our mathematical model by mathematically programming our model using constrained nonlinear optimizer found in MATLAB software, optimization tool. We utilized our mathematical functions obtained previously in step 1including the objective function and the constraints. The objective function can be noted as:

$$Max\ (\text{Production}(total)) = \text{LGO}(quantity) + \text{MGO}(quantity) + \text{HGO}(quantity) \quad (4.15)$$

Substituting equations (4.4), (4.5) and (4.6) in equation (4.15), we obtain this formula:

$$max(x, y) = 109350 - 504.7x - 54.75y + 0.4876x^2 + 0.32263xy - 0.0464y^2 \quad (4.16)$$

The constraints can be written down as:

$$843 \leq 1471 - 0.6577x - 1.63y - 0.00052x^2 + 0.00199xy + 0.00623y^2 \leq 846 \quad (4.17)$$

$$851.5 \leq 1494 - 0.7056x - 1.521y - 0.0005719x^2 + 0.001854xy + 0.00057y^2 \leq 853.5 \quad (4.18)$$

$$125.4 + 2.785x + 0.349y - 0.00253x^2 - 0.00079xy - 0.00000027y^2 = 855 \quad (4.19)$$

$$380 \leq x \leq 400 \quad (4.20)$$

$$655 \leq y \leq 730 \quad (4.21)$$

Hence,

By rewriting and combining Equations (4.16), (4.17), (4.18), (4.19), (4.20) and (4.21), we obtain the mathematical programming which can be written in standard form as:

Objective function

$$max(x, y) = 109350 - 504.7x - 54.75y + 0.4876x^2 + 0.32263xy - 0.0464y^2$$

Constraints

1. $843 \leq 1471 - 0.6577x - 1.63y - 0.00052x^2 + 0.00199xy + 0.00623y^2 \leq 846$
2. $851.5 \leq 1494 - 0.7056x - 1.521y - 0.0005719x^2 + 0.001854xy + 0.00057y^2 \leq 853.5$
3. $125.4 + 2.785x + 0.349y - 0.00253x^2 - 0.00079xy - 0.00000027y^2 = 855$
4. $380 \leq x \leq 400$
5. $655 \leq y \leq 730$

Finally, we transformed the model into MATLAB Codes and format as below:

```
function f= test(x)
f=109350 - 504.7*x(1) - 54.75*x(2) + 0.4876*x(1)^2 + 0.32263*x(1)*x(2) -
0.0464*x(2)^2;
f=-f;
function [c,ceq]=testconst(x)
c1= 1471 - 0.6577*x(1) - 1.630*x(2) - 0.0005237*x(1)^2 + 0.001993*x(1)*x(2)
+ 0.0006239*x(2)^2 - 846;
c2= -628 + 0.6577*x(1) + 1.630*x(2) + 0.0005237*x(1)^2 - 0.001993*x(1)*x(2)
- 0.0006239*x(2)^2;
c3= 1494 - 0.7056*x(1) - 1.521*x(2) - 0.0005719*x(1)^2 + 0.001854*x(1)*x(2)
+ 0.0005677*x(2)^2 - 853.5;
c4=- 642.5 + 0.7056*x(1) + 1.521*x(2) + 0.0005719*x(1)^2 - 0.001854*x(1)*x(2)
- 0.0005677*x(2)^2;
ceq1= 125.4 + 2.785*x(1) + 0.349*x(2) - 0.002533*x(1)^2 - 0.000791*x(1)*x(2)
- 0.0000002671*x(2)^2 - 855;
c5=x(1) - 400;
c6=380 - x(1);
c7=x(2) - 730;
c8=655 - x(2);
c=[c1;c2;c3;c4;c5;c6;c7;c8];ceq=(ceq1);
end
```

Step 5: Answer the question and find the optimal values

In the final step, we obtained the optimized controllable parameters through solving the mathematical model with the MATLAP application. Below are the optimum controllable values, i.e. Temperature and vacuum optimal values are:

a) Temperature = 398.4C^{o}
b) Vacuum = 656 mmHg

Thus, the optimized total quantity of the system can be obtained by substituting optimal values of temperature and vacuum in the Equation (4.14), maximum total production quantity = 14107 litres.

$$\text{LGO} + \text{MGO} + \text{HGO} = 5648 + 4570 + 3889 = 14107 \text{ litres} \qquad (4.22)$$

This theoretical value represents 70.54% of the total feeding lube oil into the system. As a result, our objective was met with respect to the process optimization under given constraints with only 1% error (since the average actual production obtained after applying the optimal parameters on the system: 14097-14093-14096-14096.5-14091 litres whose average is 14094.7 litres). However, it is important to mention that we obtained max production equal to 11400 litres under satisfied constraints prior to applying of the optimal parameters on our unit and before starting constructing our dataset observation from the daily record. Moreover, the individual products were also computed by substituting temperature and vacuum optimal values in the Equations (4.4), (4.5) and (4.6). we obtain the following:

1. LGO($\boldsymbol{quantity}$) = 5648 liters (28.24%)
2. MGO($\boldsymbol{quantity}$) = 4570 liters (22.85%)
3. HGO($\boldsymbol{quantity}$) = 3889 liters (19.45%)

To ensure our constraints were met, we substituted optimum and vacuum values in equations (4.7), (4.8) and (4.9), we obtain the following:

4. LGO($\boldsymbol{quality}$) = 845.93 (constraint satisfied)
5. MGO($\boldsymbol{quality}$) = 852.89 (constraint satisfied)
6. HGO($\boldsymbol{quality}$) = 854.99 ≈ 855 (constraint satisfied)

It is, however, important to say that setting range of values for the temperature and vacuum is feasible for the system on condition it won't violate our objective rules. As stated, our objective was to secure more than 70% of the end products under certain constraints. Therefore, by satisfying Equations (4.4), (4.5), (4.6), (4.7), (4.8) and (4.9),

we can get range of data as

a) For temperature: (397.7 C^o - 398.8C^o)
b) For vacuum: (654 mmHg -657 mmHg)

Noting that these values were limited within 5% change in the total production quantity which allow to remain within the objective's value bounds.

4.6 Discussion

The objective behind this study was to explore some of the various empirical tools and techniques deployed widely and particularly in the field of chemical processes. Empirical approach was adopted in this work for the reason that empirical methods is generally developed to use when the actual process is too complex and the underlying phenomena are not well understood or when the numerical solution of the fundamental model is quite difficult or when the empirical model provides satisfactory predictions of the process characteristics. Experimental plant data are utilized to model a relationship between the process input and process output as an empirical model using a mathematical framework.

The work introduces four different empirical techniques and their applications in modelling of processes exist in the chemical process plants. We discuss here the key advantages and disadvantages of each technique and set a ground for scientific comparison among these techniques in term of prediction performance.

4.6.1 Multiple Linear Regression (MLR)

As was defined in the latter chapters, multiple linear regression attempts to highlight hidden relationships between explanatory variables (two or more variables) with a response variable by relating a linear equation to the established observed data. Although it is a linear regression, but there are some advantages as well disadvantages listed below:

A. **Advantages**

1. MLR is fairly simple and applicable in most of the regression modelling.
2. MLR is a statistical model that, when relationships are almost linear, shows optimal results.
3. MLR is easy to deploy with compare to other regression techniques.
4. MLR is easy to understand and used to perform simulation.

B. Disadvantages

1. MLR is often inappropriately used to model non-linear relationships.
2. MLR is limited to linear relations only.
3. MLR only tracks at the mean of the dependent variable.
4. MLR is sensitive to outliers.
5. MLR models tend to over-fit at the time the number of parameters is higher than the number of when number of readings or observations.

4.6.2 Multiple Nonlinear Regression (MNLR)

Again, as it was explained in the previous chapters, multiple nonlinear regression is used to identify the existing relationship between all independent variables and one single dependent variable by appropriating a nonlinear equation to observed data.

A. **Advantages**

1. MNLR is used and more appropriate when relation between variables is nonlinear.
2. MNLR proves the capability of examining the qualified effect of one or more predictor variables to the standard value.
3. MNLR shows the capacity to identify outliers, anomalies, or other flaws.

B. Disadvantages

1. MNLR shows less predictive accuracy as compared to other machine learning tools.

2. MNLR models are inherently more difficult to than multiple linear models.

3. Multiple nonlinear regression models need more time to be solved.

4. MNLR models generally need a larger number of parameters and their mathematical formulas are complex, i.e., second-order polynomial equations or even higher order.

4.6.3 MLR Vs. MNLR Models in Chemical Processes

Most chemical processes are complex in nature as they exhibit non-linear and non-equilibrium behaviour. This phenomenon describes the unsteadiness and the dynamic trend of every chemical reaction in any process. Therefore, models which integrate nonlinearity in their behaviour are much convenient to represent these processes. Furthermore, MNLR models are more accurate and robust with less outlier values as compared to the MLR models.

4.6.4 Artificial Neural Networks (ANN)

As was discussed in the late chapters, artificial neural network (ANN) is a computational model that mimics the function of biological neural networks and retains its structure. ANN is believed to have numerous advantages over many machine learning tools. Some of the main advantages and limitations are listed below.

A. **Advantages**

1. ANN is relatively easy to use and can perform nonlinear statistical modelling for prediction purposes.

2. ANN models are quite appropriate for modelling biochemical and chemical processes.

3. ANN can learn from the observing data set.

4. ANN is very cost-effective that saves time and money

5. ANN is non-parametric model while majority of statistical tools and methods are parametric or law-based model that need complex background of statistics.

6. ANN can prove leaning and generalizing ability.

7. ANN has the capability to detect complex relationship associated with those in the nonlinear relationships.

8. ANN proved itself as a new replaceable technique to the conventional statistical techniques and multiple regression and.

9. ANN models are more precise and accurate as compared to many other classical models.

B. Disadvantages

1. ANN includes “Black Box” which originates more computational time and burden, proneness to overfitting, and forms an empirical nature of model development.

2. ANN cannot handle with models that are higher in complexity sometimes.

3. ANN models are not probabilistic, e.g., a continuous value as an output is difficult to translate into probability.

4.6.5 Ridge Regression (RR)

Ridge Regression is considered as one of the fundamental regularization technique used when the data suffers from multicollinearity where independent variables are highly correlated. RR is being limitedly used at a time whenever explanatory variable of a model undergo nonlinearity.

A. **Advantages**

1. RR generally works well when with independent variables are highly correlated.

2. RR resembles other regression techniques in that model coefficients are determined.

3. RR is powerful techniques generally used for generating robust models even in the presence of a big number of features.

4. RR can reduce model complexity by process called coefficient shrinkage.

5. RR is used to prevent overfitting.

6. RR models can outperform other regression models without regularization.

7. RR models do well in particular in the absence of a subset of true coefficients that are small or even zero.

B. Disadvantages

1. RR is not very useful in case of exorbitantly high number of features, say in millions.

2. RR method shrinks coefficients toward zero. This introduces bias, but can greatly reduce the variance, resulting in a better mean-squared error.

3. RR is very limited in application.

4. RR does not perform as good as when all of the true coefficients are moderately large.

4.6.6 Empirical Models' Prediction Performance

The main goal of this work was to spotlight on the performance of various empirical models generated for the purposes of controlling the chemical processes. Prediction ability and accuracy is the key character of any empirical model. Similarly, each empirical model with different technique differs from other empirical models by its prediction performance and accuracy. In this work, four different empirical models with different techniques were generated in order to forecast chemical process future values, i.e., production and quality. Subsequently, and after taking performance measurement on their prediction performance with several metrics, it is obvious that as indicated in Table (4.15), artificial neural network (ANN) tool shows better performance than any other tools tried in this work with mean absolute percentage error equal to 2.42 followed by ridge regression (RR), multiple nonlinear regression

(MNLR), and multiple linear regression (MLR) tools with 3.27, 5.18 and 5.62 respectively as shown in Figure (4.25).

Table 4.15 Performance metrics of MLR, MNLR, ANN and RR algorithms

Metrics \ Model	Multiple Linear Regression (MLR)	Multiple Nonlinear Regression (MNLR)	Artificial Neural Networks (ANN)	Ridge Regression (RR)
MAD	329	304	142	192
MSE	161861	108705	23787	75874
RMSE	402	329	154	275
MAPE	5.62	5.18	2.42	3.27

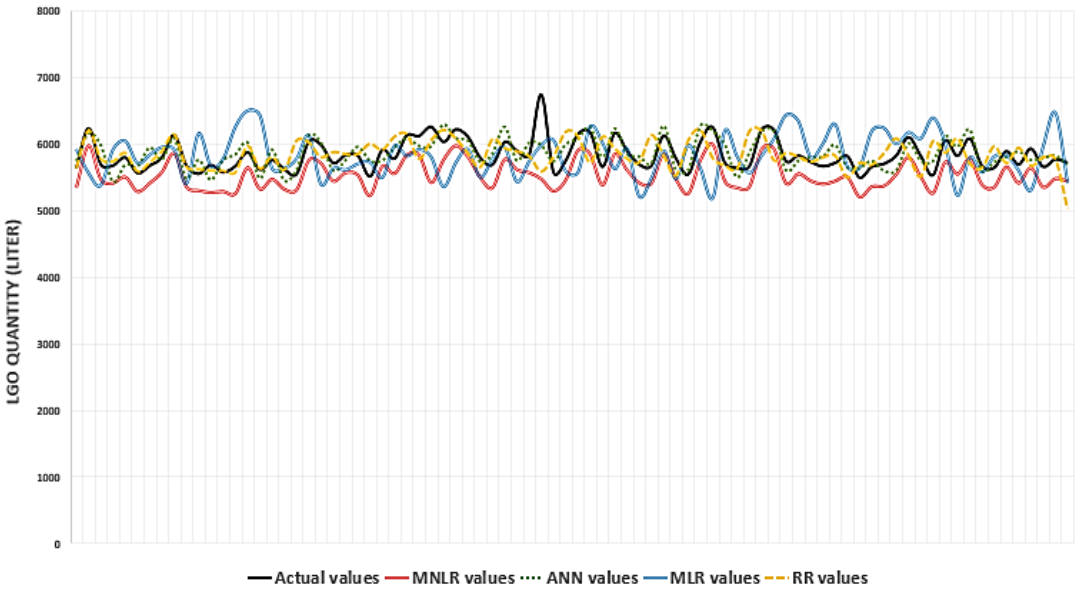


Figure 4.25 Accuracy performance of MLR, MNLR, ANN and RR training (difference of mean of methods)

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 Introduction

This chapter gives a concise conclusion on the work done with some practical evidences showing the importance of the study. Moreover, the chapter provides suggestions to the work done with future recommendations as well.

5.2 Conclusion

Process modelling has become widely a popular area of interest particularly in the field of chemical process control and optimization. Application of empirical model on the chemical processes can help in optimizing the chemical system which in return, can cut considerable expenses and reduce time and effort. There are millions of dollars being save yearly from the crude oil refineries across the world because their processes are controlled and brought to optimum. Empirical modelling is one of the modelling technique and was used for decades. However, there are newly emerging tools like neural networks which proved better modelling performance as compared to its peers in the same area.

In this work, four different empirical modelling tolls were deployed on our chemical system namely “Used lube oil recycling unit”. The Neural Network model, which is one of the four models generated, was used to control and optimize the system’s processes. Prior to the work, we had weekly equal to 4300 litres loss of crude oil in form of used lube oil which represents around 3% of the final products because of incorrect working parameter values. Consequently, this renders our system to economically be inefficient with approximately 9,460TL loss every week (around 37,840TL in a month). This loss was removed entirely after processes were controlled through obtaining the optimal values of the system’s controllable parameters. Furthermore, we could enhance production rate form 57% to above 70% which enabled us to increase profit to about 23% at a time we could satisfy our system’s

quality constraints for the products namely: Light Gas Oil (LGO), Moderate Gas Oil (MGO) and Heavy Gas Oil (HGO) with standard concentration constraints as below:

- Heavy gas oil, HGO($quly$) = 855 liter/m3
- 853.5 ≥ Moderate gas oil, MGO($quly$) ≥ 851.5 liter/m3
- 846 ≥ Light gas oil, LGO($quly$) ≥ 843 liter/m3

Another objective of this work was to compare between four well-known empirical modelling techniques in term of their generated model's predictive performance, namely: (1) Multiple Linear Regression, (2) Multiple Nonlinear Regression, (3) Artificial Neural Network and (4) Ridge Regression. As a result, the artificial neural network (ANN) model showed better performance than the other models with mean absolute percentage error equal to 2.42 followed by ridge regression (RR), multiple nonlinear regression (MNLR), and multiple linear regression (MLR) tools with 3.27, 5.18 and 5.62 respectively.

Eventually, empirical models as compared to fundamental models or the first-law models have the following positive-edges and properties listed below:

1. Empirical models are simple and lead to limited knowledge.
2. Empirical models can rapidly develop even when there is limited physical understanding of the system.
3. Empirical model's outputs are generally readily understood and interpreted.
4. Empirical models are cost-effective and faster to handle as compared to the fundamental models.
5. Empirical models offer explicit, direct control over the structure of the resulting model.

On contrary, empirical models carry some significant disadvantages as:

1. Empirical models sometimes can fail to capture the true behaviour over a large region of operating space.

2. Empirical models normally require an expendable large data set, from both historical record and during process as well.

3. Empirical models inherently do not offer a diagnostic or indicative methodology to identify and recognize the root cause of problems.

4. Empirical models generally lack the direct physical interpretation that characterizes fundamental models.

5. Empirical models are based exclusively on finite collections of observed data. Therefore, prediction outside the range of these datasets may be unreliable.

6. Empirical models generally require accurate extrapolations in case of optimization and control.

5.3 Recommendation

In general, the modelling approaches fall in two categories, one is the parametric or fundamental modelling and the other is the empirical modelling. However, in the recent years, lot of efforts were put into enhancing or finding new ways and methods for building more efficient and accurate predictive models which can be used in the modelling of chemical processes. Therefore, it was finally possible to combine both unique models into one model with enhanced properties. This model, however, is called a “hybrid” or “mixed” model which can overcome the drawbacks of each previously said approach while combining the advantages of each approach as well. Hence, it is highly recommended to direct all effort and research work toward exploring more effective modelling approaches and techniques. Reader may consult the work presented by Psichogios and Ungar (1992) who pioneered usage of the hybrid technique in the chemical process to identify the biochemical reactor or (Dahm and Oliveira, 1997) who combined both the fundamental and empirical models to emulate the processes in the polymerization reactors. He developed the mass balance equations for the reactant within the fundamental approach, while developed the unknown rates of the reactions within the empirical approach.

REFERENCES

Ahmed, DF, Shakoor, AM. (2016). Neural network control for a batch distillation column, *Tikrit Journal of Engineering Sciences*, **23(1)**, 10-19.

Aiken, LS, West, SG. (1991). Multiple regression: testing and interpreting interactions. Newbury Park, London, New Delhi: Sage Publications.

Allgöwer, F, Zheng, A. (2000). Nonlinear model predictive control. 1st edition. Basel: Springer Basel AG.

Ansari, RM, Tade, MO. (2000). Nonlinear model-based process control: Applications in petroleum refining. London: Springer.

Archdeacon, TJ. (1994). Correlation and regression analysis: A Historian's guide. USA: University of Wisconsin Press.

Baughman, DR, Liu, YA. (1995). Neural networks in bioprocessing and chemical engineering. California: Academic Press, Inc.

Bari, A, Chaouchi, M, Jung, T. (2017). Predictive analytics for dummies. 2nd edition. Canada, New Jersey: John Wiley & Sons, Inc.

Balchen, JG, Gilles, ED, Waller, KV, Rawlings, JB. (1993). Dynamics and control of chemical reactors, distillation columns and batch processes. 1st edition. UK, New York, Seoul, Tokyo: Pergamon Press.

Berber, R. (1995). Methods of model based process control. **293**: 1st edition. Dordrecht: Springer-Science, Business Media, B. V.

Bequette, BW. (2003). Process control: Modelling, design, and simulation. New Jersey: Pearson Education, Inc.

Berk, RA. (2004). Regression analysis: A constructive critique. USA: Sage Publications.

Ben´ıtez, JM, Castro, JL, Requena, I. (1997). Are artificial neural networks black boxes? *IEEE,* **8 (5)**, 1156-1164.

Bequette, BW. (1991). Nonlinear control of chemical process: A review, *American Chemical Society,* **30**, 1391-1413.

Birkes, D, Dodge, Y. (1993). Alternative methods of regression. New York, Chichester, Brisbane, Toronto, Singapore: Wiley.

Biyanto, TR, Widjiantoro, BL, Jabal, A, Budiati, T. (2010). Artificial neural network based modelling and controlling of distillation column system, *International Journal of Engineering, Science and Technology*, **2(6)**, 177-188.

Bishop, CM. (1995). Neural networks for pattern recognition. USA: Oxford University Press Inc.

Bonvin, D. (1994). Advanced control of chemical processes. UK, USA, Japan: Pergamon.

Carr, RW. (2007). Comprehensive chemical kinetics: Modelling of chemical reactions. **42:** 1st edition. Amsterdam: Elsevier.

Cameron, IT, Hangos, KM. (2001). Process modelling and model analysis. **4**. USA, UK: Academic Press.

Chang, AF, Pashikanti, K, Liu, YA. (2012). Refinery engineering: Integrated process modelling and optimization. Germany: Wiely-VCH Verlag GmbH & Co. KGaA.

Chatterjee, S, Hadi, AS. (1988). Sensitivity analysis in linear regression. Canada: Wiley.

Chatterjee, S, Hadi, AS. (2012). Regression analysis by example. 5th edition. New Jersey, Canada: Wiley.

Chetouani, Y. (2007). Using artificial neural networks for the modelling of a distillation column, *International Journal of Computer Science & Applications*, **4(3)**, 119-133.

Chu, BT, Chen, SS. (1992). Intelligent modelling, diagnosis and control of manufacturing processes. **4:** 1st edition. Singapore: World Scientific Publishing Co. Pte. Ltd.

Cinar, A, Palazoglu, A, Kayihan, F. (2007). Chemical process performance evaluation. Boca Raton: CRC Press.

Deng, Z, Choi, KS, Jiang, Y. (2014). Generalized hidden-mapping ridge regression, knowledge-leveraged inductive transfer learning for neural networks, fuzzy systems and kernel methods, *IEEE*, **44(12)**, 2585-2599.

Dobre, TG, Maracano, JS. (2007). Chemical engineering: Modelling, simulation and similitude. Germany: Wiely-VCH Verlag GmbH & Co. KGaA.

Elnashaie, SSEH, Garhyan, P. (2003). Conservation equations and modelling of chemical and biochemical processes. New York: Marcel Dekker, Inc.

Estopier, MO, Jobson, M, Smith, R. (2013). Operational optimization of crude oil distillation systems using artificial neural networks, *computer and chemical engineering*, **59**, 178-185.

Fan, J, Gijbels, I. (2003). Local polynomial modelling and its applications. Boca Raton, London, New York, Washington, D.C.: Chapman & Hall.

Fazlali, A, Hosseini, S, Yasini, B, Moghadassi, A. (2009). Optimization of operating conditions of distillation columns: an energy saving option in refinery industry, *Songklanakarin Journal of Science and Technology*, **31(6)**, 661-664.

Fernandes, FAN, Lona, LMF. (2005). Neural network applications in polymerization processes, *Brazilian Journal of Chemical Engineering*, **22(3)**, 401-418.

Freund, RJ, Wilson, WJ, Sa, P. (2006). Regression analysis IM. 2nd edition. USA, UK: Elsevier, Inc.

Freud, RJ, Littell, RC. (2000). SAS system for regression. 3rd edition. USA: SAS Institute Inc.

Gary, JH, Handwerk, GE. (2001). Petroleum refining. 4th edition. New York, Basel: Maercel Dekker, Inc.

Genç, ÖS. (2006). Modelling and control of bioprocesses by using artificial neural networks and hybrid model, *İzmir Institute of Technology*, 3-32.

Golberg, MA, Cho, HA. (2010). Introduction to regression analysis. UK: WIT Press.

Gupta, GK. (2014). Introduction to data mining with case studies. 3rd edition. Delhi: Asoke K Ghosh, PHI Learning Private Ltd.

Gurney, K. (1997). An introduction to neural networks. London: UCL Press Ltd.

He, X, Xu, S. (2009). Process neural networks: Theory and applications. China: Zhejiang University Press.

Hernández, JGS, Petriciolet, AB. (2016). Process intensification in chemical engineering: Design optimization and control. Switzerland: Springer.

Himmelblau, DM. (2000). Applications of artificial neural networks in chemical engineering, *Korean J. Chem. Eng.* **17(4)**, 373-392.

Hilt, DO, Seegrist, DW. (1977). Ridge, a computer program for calculating ridge regression estimates, *Northeastern Forest Experiment Station*, **13**, 1-7.

Hui, LQ, Hui, LW, Aziz, N, Ahmad, Z. (2011). Nonlinear process modelling of "shell" heavy oil fractionator using neural network, *Journal of Applied Sciences*, DOI.

Huet, S, Bouvier, A, Poursat, MA, Jolivet, E. (1996). Statistical tools for nonlinear regression: A practical guide with S-PLUS examples. 1st edition. New York: Springer.

Ivanov, AV. (1997). Asymptotic Theory of nonlinear regression. Dordrecht: Springer

Jana, AK. (2011). Chemical process modelling and computer simulation. 2nd edition. New Delhi: Asoke K Ghosh, PHI Learning Private Ltd.

Jakobsen, HA. (2014). Chemical reactor modelling: Multiphase reactive flows. 2nd edition. Switzerland: Springer.

Jain, MK, Iyengar, SRK, Jain, RK. (1995). Numerical methods for scientific and computer computation. Kluwer Academic Publishers.

Joly, M, Moro, LFL, Pinto, JM. (2001). planning and scheduling for petroleum refineries using mathematical programming, *Brazilian Journal of Chemical Engineering*, **19(2)**, 207-228.

Johansen, A, Foss, BA. (1997). Operating regime based process modelling and identification, *Computers & Chemical Engineering*, **21(2)**, 159-176.

Karayiannis, NB, Venetsanopoulos, AN. (1993). Artificial neural networks: Learning algorithms, performance evaluation, and application. Boston, Dordrecht, London: Kluwer Academic Publishers.

Karuppiah, R, Furman, KC, Grossmann, IE. (2008). Global optimization for scheduling refinery crude oil operations, *Computers and Chemical Engineering*, **32**, 2745-2766.

Keshani, S, Daud, WR, Nourouzi, MM. (2015). Spray drying: An overview on wall deposition, process and modelling, *Journal of Food Engineering*, **146**, 152-162.

Keith, TZ. (2015). Multiple regression and beyond: An introduction to multiple regression and structural equation modelling. 2nd edition. New York: Routledge.

Kleinbaum, DG, Kupper, LL, Nizam, A, Rosenberg, ES. (2014). Applied regression analysis and other multivariable methods. 5th edition. USA: Cengage Learning.

Klimberg, R, McCullough, BD. (2013). Fundamentals of predictive analytics with JMP. USA: SAS Institute, Inc.

Kouvaritakis, B, Cannon, M. (2001). Non-linear predictive control: Theory and practice. 1st edition. London: Institution of Engineering and Technology.

Kotu, V, Deshpande, B. (2015). Predictive analytics and data mining: Concepts and practice with rapidminer. USA: Elsevier, Inc.

Konishi, S. (2014). Introduction to multivariate analysis: Linear and nonlinear modelling. USA: CRC Press.

Kumar, E. (2008). Artificial intelligence. New Delhi: I.K. International Publishing House Pvt. Ltd.

Leffler, WL. (2008). Petroleum refining in nontechnical language. 4th edition. USA: PennWell Corporation.

Liaua, LCK, Yang, TCK, Tsaib, MT. (2004). Expert system of a crude oil distillation unit for process optimization using neural networks, *Expert Systems with Applications*, **26**, 247-255.

Malhotra, GK. (2011). Chemical process simplification improving productivity and sustainability. New Jersey: John Wiley & Sons, Inc.

Marsland, S. (2009). Machine learning: An algorithmic perspective. USA: CRC Press.

Maeder, M, Neuhold, YM. (2007). Practical data analysis in chemistry. **26**. Amsterdam: Elsevier.

Mehrotra, K, Mohan, CK, Ranka, SR. (2000). Elements of artificial neural networks. USA: MIT.

Meerschaert, MM. (2013). Mathematical modelling. 4th edition. USA, Amsterdam, UK: Elsevier.

Montgomery, DC. (2017). Introduction to linear regression analysis. 5th edition. Content Technologies. Inc.

Montgomery, DC, Peck, EA, Vining, GG. (2012). Introduction to linear regression analysis. 5th edition. New Jersey, Canada: Wiley.

Morshedi, AM. (1986). Universal dynamic matrix control. *Chemical Process Control III*, 547-577. New York: Elsevier.

Mujtaba, IM, Hussain, MA. (2001). Application of neural networks and other learning technologies in process engineering. London: Imperial College Press.

Natarajan, BK. (1991). Machine learning: A theoretical approach. USA: Morgan Kaufmann Publishers, Inc.

Nelles, O. (2001). Nonlinear system identification: From classical approaches to neural networks and fuzzy models. Germany: Springer.

Okkan, U. (2011). Application of levenberg-marquardt optimization algorithm based multilayer neural networks for hydrological time series modelling, *An International Journal of Optimization and Control: Theories & Applications*, **1(1)**, 53-63.

Olson, DL, Wu, D. (2017). Predictive data mining models. Singapore: Springer.

Olafadehan, OA, Adeniyi, VO, Popoola, LT, Salami, L. (2013). Mathematical modelling and simulation of multicomponent distillation column for acetone-chloroform-methanol system, *Advanced Chemical Engineering Research*, **2(4)**, 113-123.

Pawar, M. (2004). Data collecting methods and experiences: A guide for social researchers. USA, UK, India: New Dawn Press, Inc.

Pearson, RK. (1999). Discrete-time dynamic models. New York: Oxford University Press, Inc.

Phillips, PP, Stawarski, CA. (2008). Data collection: Planning for and collecting all types of data. **2**. USA: Pfeiffer

Popoola, LT, Babagana, G, Susu, AA. (2013). A Review of an expert system design for crude oil distillation column using the neural networks model and process optimization and control using genetic algorithm framework, *Advances in Chemical Engineering and Science*, **3**, 164-170.

Priddy, KL, Keller, PE. (2005). Artificial neural networks: An introduction. **TT68**. USA: Spie Press.

Quantrille, TE, Liu, YA. (1991). Artificial intelligence in chemical engineering. UK: Academic Press, Inc.

Ray, W.H. (1982). New approaches to the dynamics of nonlinear systems with implications for process and control system design, *Chemical Process Control*, 246-267.

Ray, W.H. (1983). Multivariable process control, *Comput. Chem. Eng*, **47**, 367-394.

Raschka, S. (2016). Python machine learning. UK: Packt Publishing Ltd.

Rigatos, GG. (2010). Intelligent industrial systems: Modelling, automation and adaptive behaviour. USA: Information Science Reference.

Ramchandran, B. (1994). Neural network model-based control of distillation columns, *Texas Tech University*, PhD dissertation.

Rasmuson, A, Andersson, B, Olsson, L, Andersson, R. (2014). Mathematical modelling in chemical engineering. 1st edition. New York: Cambridge University Press.

Rhinehart, RR. (2016). Nonlinear regression modelling for engineering applications: Modelling, model validation, and enabling design of experiments. 1st edition. UK: Wiley.

Ribas, GP, Leiras, A, Hamacher, S. (2012). Operational planning of oil refineries under uncertainty Special issue: Applied Stochastic, *IMA Journal of Management Mathematics*, **23**, 397-412.

Rojas, R. (1996). Neural networks: A systematic introduction. Berlin: Springer.

Romagnoli, JA, Palazoglu, A. (2006). Introduction to process control. Boca Raton, London, New York, Singapore: CRC Press.

Roffel, B, Betlem, B. (2006). Process dynamics and control: Modelling for control and prediction. England: Wiley.

Sadighi1, S, Mohaddecy1, RS, Ameri, YA. (2015). Artificial neural network modelling and optimization of hall-heroult process for aluminum production, *International Journal of Technology*, **3**, 480-491.

Santos, CA, Spim, JA, Ierardi, MCF, Garcia, A. (2002). The use of artificial intelligence technique for the optimization of process parameters used in the continuous casting of steel, *Applied Mathematical Modelling*, **26**, 1077-1092.

Seber, GAF, Wild, SCJ. (2003). Nonlinear regression. New Jersey: Wiley.

Sen, AK, Srivastava, MS. (1990). Regression analysis: Theory, methods and applications. 1st edition. New York: Springer.

Shore, H. (2005). Response modelling methodology: Empirical modelling for engineering and science. **8.** Singapore: World Scientific Publishing Co. Ltd.

Shalini, J, Kumar, JS, Swathanthra, PA. (2012). Steady state modelling of chemical systems by using neural networks, *International Journal of Engineering Research & Technology (IJERT)*, **1(6)**, 1-6.

Shinnar, R. (1986). Impact of model uncertainties and nonlinearities on modern controller design present status and future goal. *In Chemical Process Control*, 53-93. New York: Elsevier.

Shah, Nita, Mishra, Poonam. (2013). Oil production optimization: a mathematical model, *Journal Petrol Explore Production Technology*, **3**, 37- 42.

Simpson, R, Sastry, SK. (2013). Chemical and bioprocess engineering: Fundamental concepts for first-year students. New York, Heidelberg, Dordrecht, London: Springer.

Silva, A, Brito, JD, Gaspar, PL. (2016). Methodologies for service life prediction of buildings: With a focus on façade claddings. Switzerland: Springer International Publishing.

Siegel, E. (2016). Predictive analytics: The power to predict who will click, buy, lie, or die. New Jersey, Canada: Wiley.

Sotto, ER. (2015). Clean recovery of antioxidant compounds from plant foods, by-products and algae assisted by ultrasounds processing. Modelling approaches to optimize processing conditions, *Trends in Food Science & Technology*, **42(2)**, 134-149.

Srivastava, VK, Fahim, A. (2007), An optimization method for solving mixed discrete-continuous programming problems, *Computers and Mathematics with Applications*, **53**, 1481-1491.

Strickland, JS. (2014). Predictive analytics using R. Colorado: Lulu, Inc.

Stiesch, G. (2003). Modelling engine spray and combustion processes. New York: Springer.

Upreti, SR. (2017). Process modelling and simulation for chemical engineers: Theory and practice. 1st edition. UK: Wiley.

Weisberg, S. (2014). Applied linear regression. 4th edition. Canada: Wiley.

WEF. Water environment Federation. (2014). Wastewater treatment process modelling. 2nd edition. 14. p. Virgina.

Whitby, B. (2009). Artificial intelligence (world issue today). New York: Rosen Publishing.

Wiley, J. (2005 Int. J. Robust nonlinear control, *International Journal of Robust and Nonlinear Control*, **15**, 287-289.

Wold, S. (1992). Nonlinear partial least squares modelling II. Spline inner relation, *Chemometric and Intelligent Labouratory Systems*, **14(1, 2, 3)**, 1-448.

Yan, X. (2009), Linear regression analysis: Theory and computing. Singapore: World Scientific Publishing Co. Pte. Ltd.

Yusof, KM, Idris, A, Siong, LJ, Mun, WH, Morad, NA. (2003). Artificial neural network modelling of steady state chemical engineering systems, *Invited paper for Malaysia-Japan Seminar on Artificial Intelligence Applications in Industry*.

APPENDIX 1

Figure A.1 The lube oil recycling unit

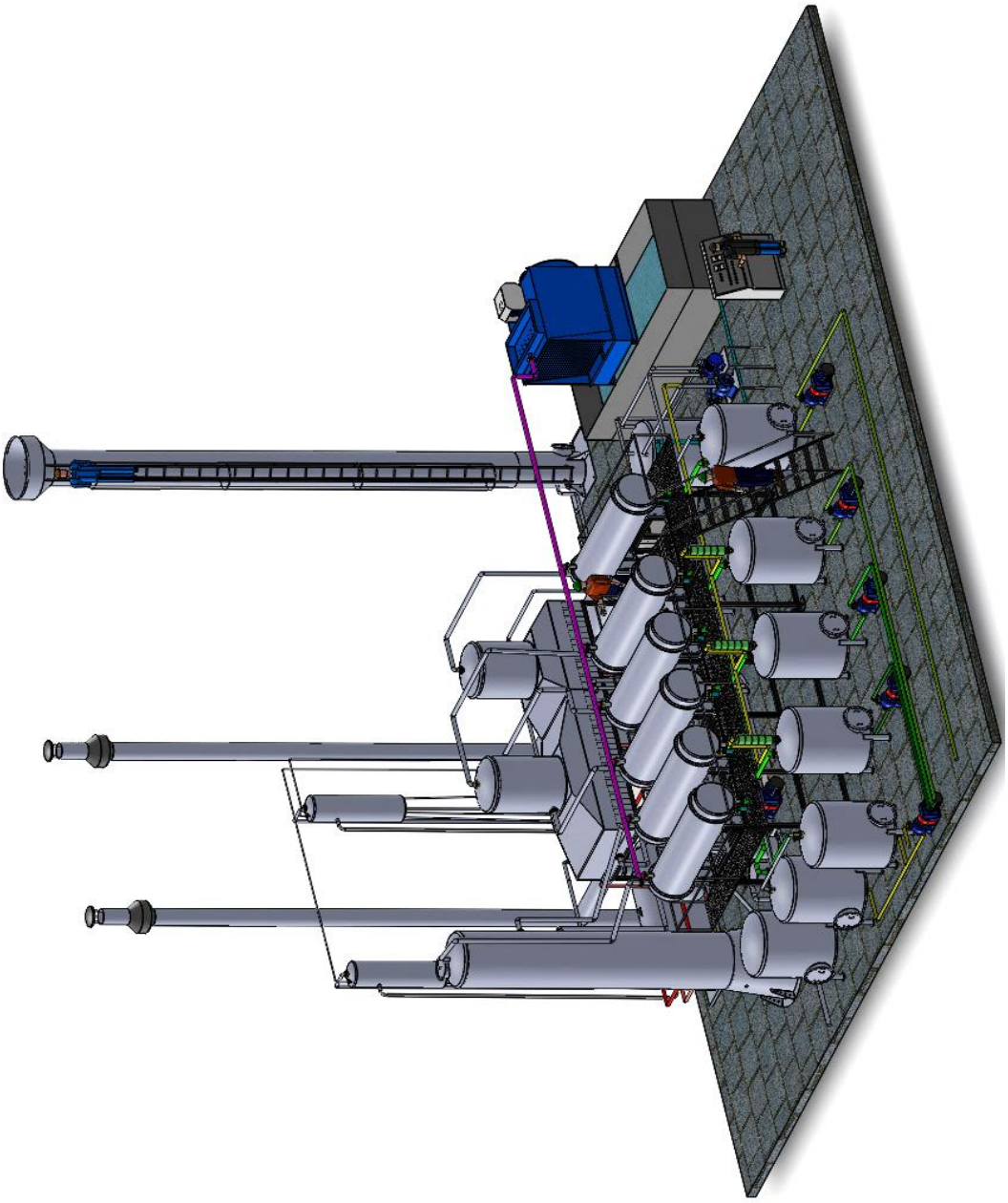


Figure A.2 3D design of the lube oil recycling unit (isometric view:1)

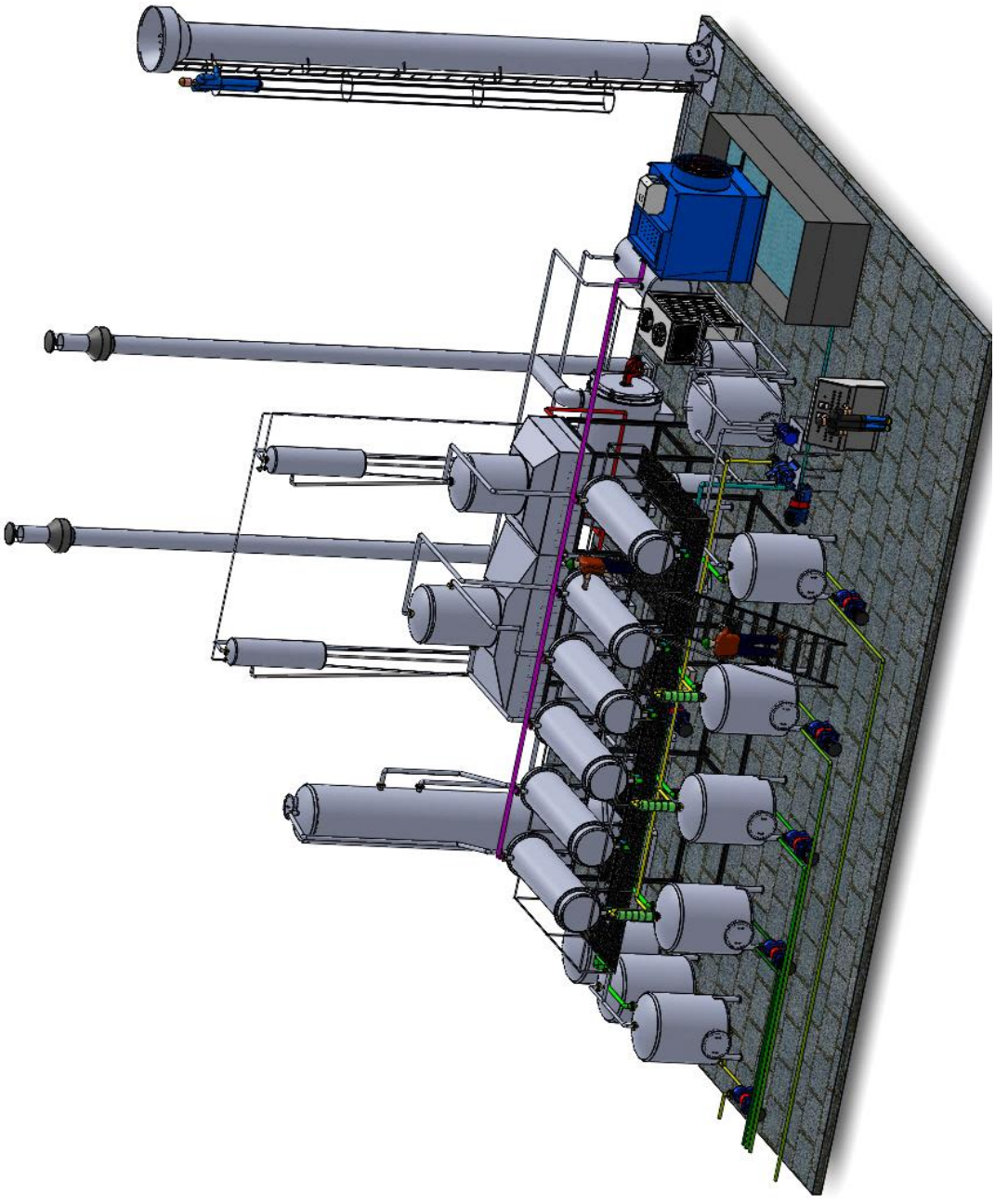


Figure A.3 3D design of the lube oil recycling unit (isometric view:2)

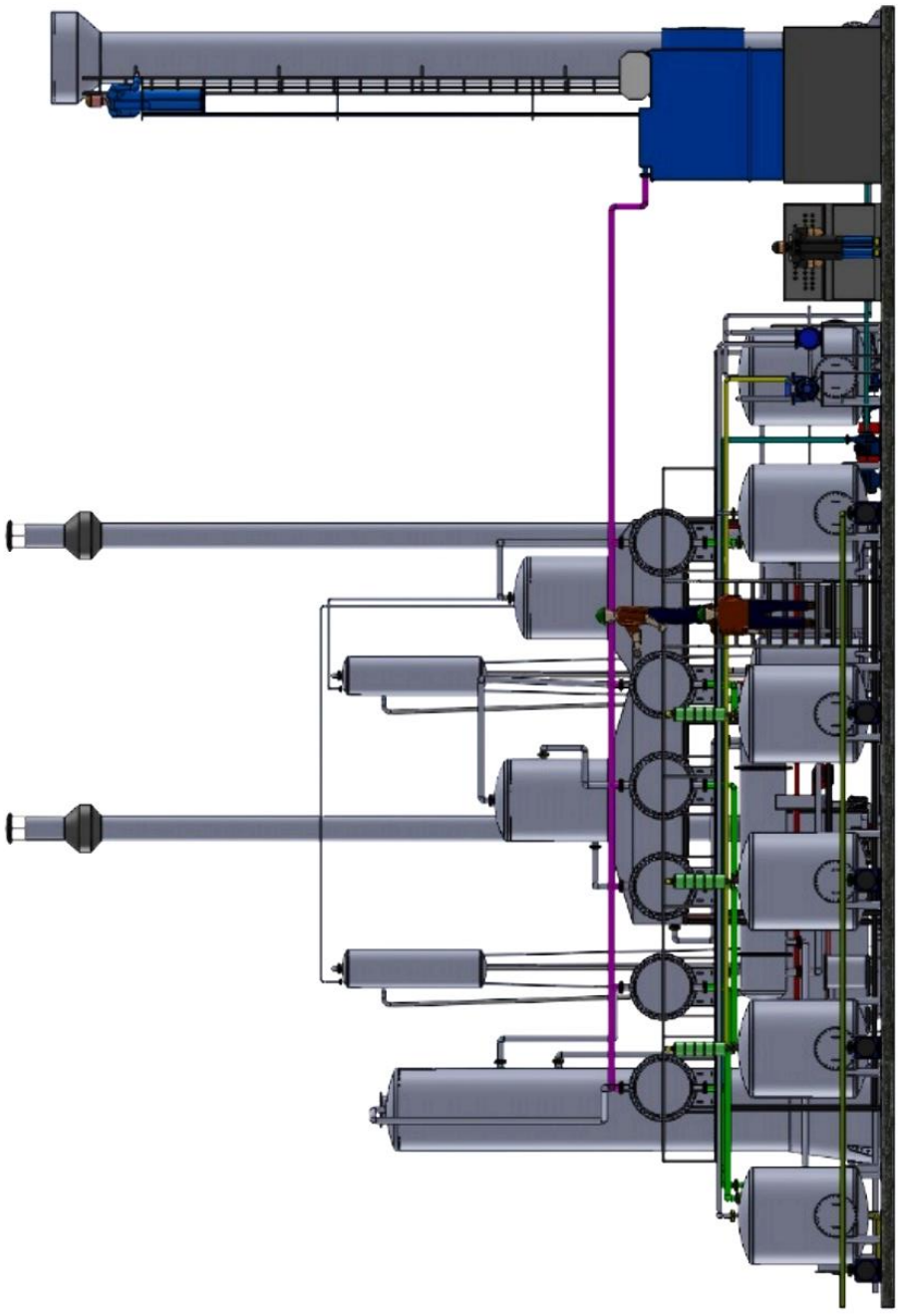


Figure A.4 3D design of the lube oil recycling unit (side view:3)

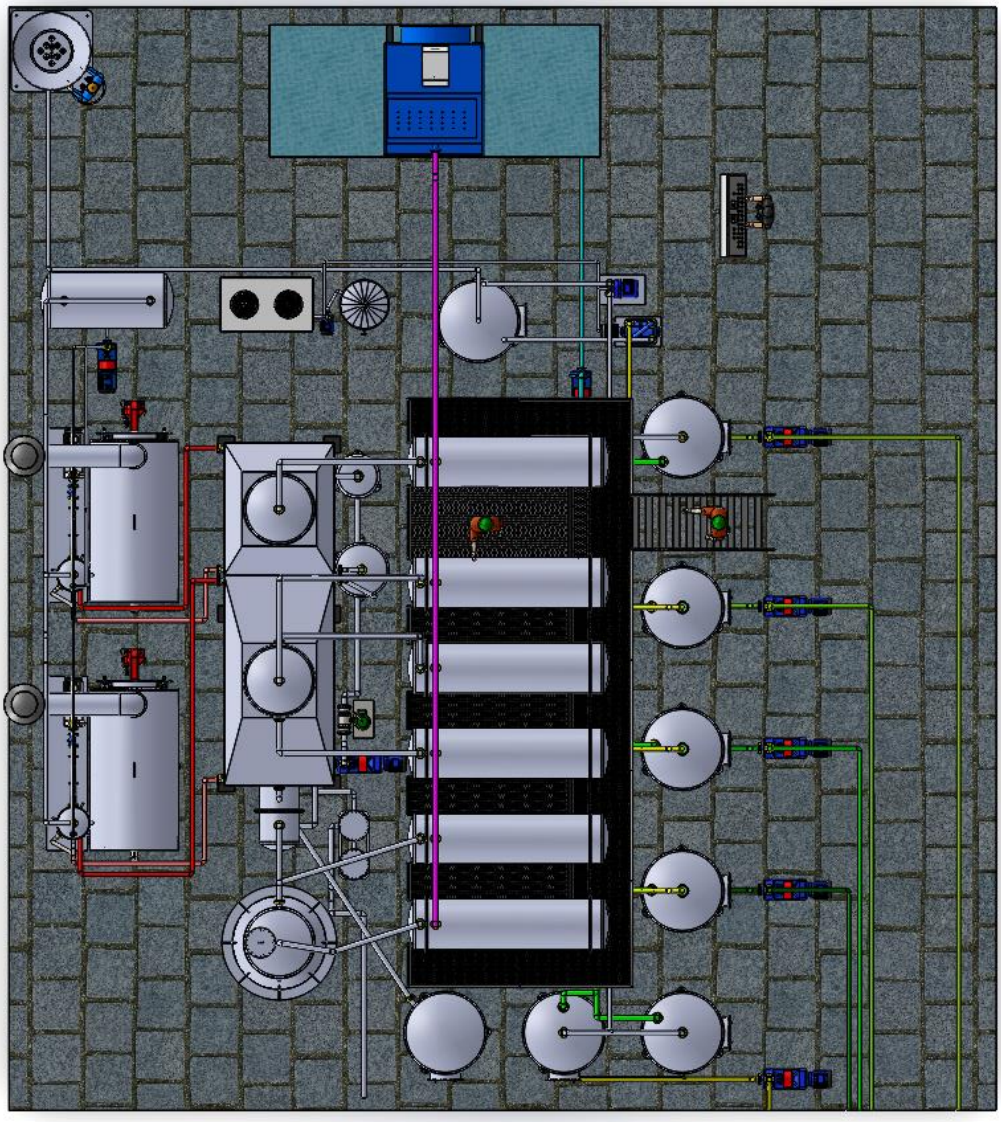


Figure A.5 3D design of the lube oil recycling unit (top view:4)

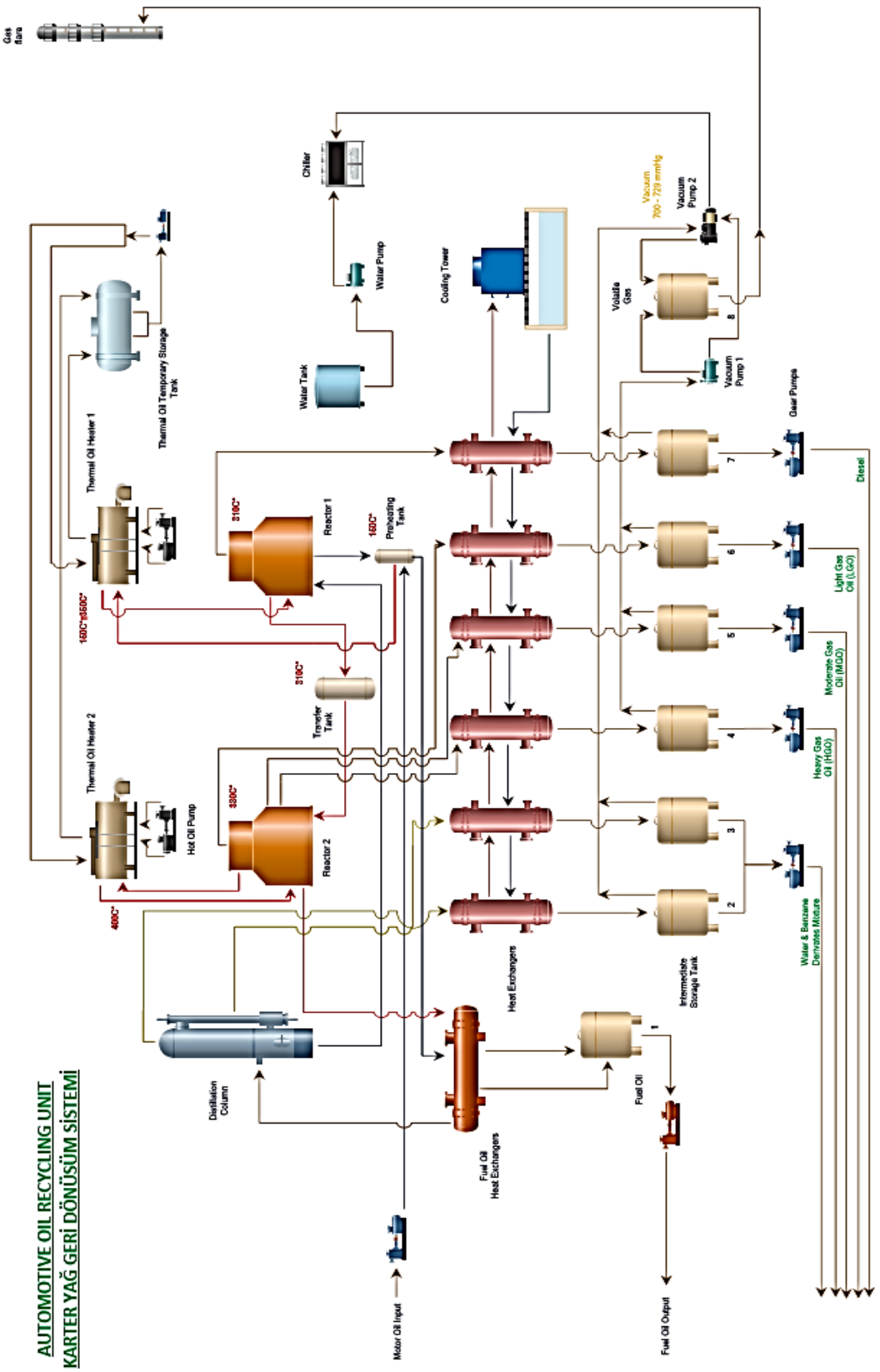


Figure A.6 Lube oil recycling unit process flow diagram

Table A.1(a) 82 observations taken from the used lubricant recycling unit from daily record (max production = 16958 litres and min production = 13756 litres)

S/NO	INPUT		OUTPUT			OUTPUT		
	INPUT DATA		DISTILLATE QUANTITY (LITER)			DISTILLATE QUALITY (CONCENTRATION)		
	TEMPERATURE C°	VACUUM mmHg	LIGHT GAS OIL	MODERATE GAS OIL	HEAVY GAS OIL	LIGHT GAS OIL2	MODERATE GAS OIL3	HEAVY GAS OIL4
1	383.7	710	5625	4739	3664	842	851	852.5
2	390.2	716	5830	4913	4037	845	852.5	855.5
3	385	708	5667	4779	3720	842.5	851.5	852.5
4	383.5	705	5560	4728	3658	842	851	853
5	389.9	718	5835	4913	4036	845.5	852.5	855
6	391.4	702	5787	4873	4011	843	852	853.5
7	395	710	6061	5140	4178	847	853.5	857
8	395.5	714	6039	5117	4162	846	853.5	856.5
9	388.9	707	6746	5842	4061	844	852.5	854
10	388.4	702	5720	4818	3854	843	852	853.5
11	390.1	702	5805	4896	4018	844.5	852	854.5
12	393	717	5897	5066	4095	846	853	856
13	396.2	717	6109	5158	4151	846	853	855.5
14	395.8	712	6037	5115	4160	846	853.5	856.5
15	397.1	718	6177	5240	4228	847.5	854	857.5
16	397.8	709	6141	5195	4200	847	853.5	856.5
17	397.5	710	6133	5186	4189	846.5	853.5	856.5
18	392.2	712	5906	5065	4121	844.5	852.5	855.5
19	388	718	5701	4795	3846	843.5	852	854
20	389.3	715	5828	4906	3980	845	853	856
21	382.7	710	5579	4728	3622	841.5	851	852.5
22	384.1	719	5650	4762	3690	842.5	851	853
23	380.6	719	5538	4690	3600	841.5	851	853
24	380.2	706	5519	4683	3580	840	850.5	852
25	381	711	5551	4693	3588	841	851	852
26	382.7	707	5571	4715	3612	841.5	851	852.5
27	385	705	5649	4766	3702	842	851.5	852.5
28	386.4	716	5682	4805	3743	843	851.5	854
29	386.2	707	5665	4789	3727	842.5	851.5	853.5
30	385.9	713	5681	4790	3738	843	851.5	854
31	384.4	710	5695	4798	3730	843.5	852	855
32	386	701	5665	4773	3720	842.5	851.5	853.5
33	386.5	714	5680	4802	3741	843	851.5	854
34	388.3	703	5722	4810	3856	843	852	853.5
35	387.7	703	6265	5262	4095	846.5	852.5	855.5
36	387.2	704	5748	4854	3800	844	852.5	855
37	383.8	710	5622	4735	3661	841.5	851	852.5
38	386.8	711	5683	4799	3739	843	851.5	854
39	388.1	704	5720	4811	3859	843	852	853.5

Table A.1(b) 82 observations taken from the used lubricant recycling unit from daily record

41	390.9	716	5842	5922	4060	842	850.5	852	
42	390.7	719	5840	4928	4057	845.5	852.5	855.5	
43	393.1	710	5883	5049	4085	845.5	853	855.5	
44	394.5	710	5993	5058	4125	845	853.5	856	
45	389.2	715	5826	4909	3978	845	853	856	
46	386.6	707	5673	4789	3730	842.5	851.5	853.5	
47	388.5	702	5738	4830	3870	843.5	852	854	
48	391	719	5804	4891	4020	843.5	852	854	
49	382.9	701	5689	4936	3757	842	851.5	852.5	
50	382.2	701	5674	4899	3730	841	851	851.5	
51	385.3	712	5710	4799	3737	843.5	852.5	855	
52	384	717	5642	4755	3681	842	851	853	
53	387.8	700	5700	4821	3765	843	852	854	
54	388.6	709	5766	4857	3990	844.5	852.5	854.5	
55	388.7	705	5760	4851	3985	844	852.5	854.5	
56	390.3	716	5827	4909	4035	845	852.5	855	
57	392	714	5902	5062	4120	844.5	852.5	855.5	
58	396.3	719	6113	5167	4160	846	853	855.5	
59	397.8	701	6126	5175	4171	846	853.5	856	
60	395.4	715	6038	5115	4162	846	853.5	856.5	
61	399.2	709	6219	5239	4222	848	854	857.5	
62	399.7	718	6225	5248	4239	849	854	855.5	
63	396.9	718	6135	5183	4179	846.5	853.5	856.5	
64	390.3	719	5832	4915	4048	845	852.5	855.5	
65	392.5	722	5938	5084	4139	845.5	853	856.5	
66	393.9	705	5919	5062	4101	846	853	856	
67	395.3	717	6043	5120	4169	846.5	853.5	856.5	
68	397	726	6168	5219	4208	847	854	857	
69	398.8	708	6155	5193	4180	847.5	853.5	857	
70	399.1	710	6221	5238	4227	848	854	857.5	
71	399.6	723	6240	5266	4251	849	854.5	858	
72	399.9	729	6259	5312	4290	849.5	855	858	
73	389	704	5790	4883	3951	844.5	852.5	855	
74	388.5	715	5777	4862	3909	844	852.5	855	
75	387.8	709	5725	4832	3781	843.5	852.5	854.5	
76	386.2	720	5696	4811	3753	843.5	852	854.5	
77	384	726	5672	4788	3710	842.5	851.5	854	
78	382.9	706	5584	4716	3629	841.5	851	853	
79	380.1	701	5503	4673	3580	840	850.5	852	
80	384.7	714	5680	4783	3712	843	852	854.5	
81	396.1	708	6089	5137	4130	845	852.5	854.5	
82	383.3	711	5556	4729	3660	842	851	853	
MAX	399.9	729	6746	5922	4290	849.5	855	858	
MIN	380.1	700	5503	4673	3580	840	850.5	851.5	

INFORMATION

Name and Surname: Al Mothana AL SHAREEF

Nationality: Syrian

Birth place and date: Islamabad – 01/04/1985

Marital status: single

Phone number: +(90) 538 075 20 22

Fax: N/L

Email: engr1st@gmail.com

EDUCATION

	Graduate school	**Year**
Master	Gaziantep University-Turkey	2017
Bachelor	University of Engineering & Technology (UET)-Pakistan	2009
High School	Saudi Arab School	2005

Work experience

	Place	**Enrollment**
2013-Present	KAYALAR ENERGY-Gaziantep	Mechanical Engineer
2011-2012	IICO-Pakistan	Program Coordinator
2007-2008	TOYOTA MOTOR-Pakistan	Maintenance Trainee

FOREIGN LANGUAGE

Arabic | English | Urdu | Pashto | Turkish

HOBBIES

Traveling | Reading | Science-oriented projects & researches.