Oil Demand Forecasting in Malaysia in Transportation Sector using Artificial Neural Network (ANN)

by

Muhammad Afiq Bin Abd Rashid
14935

Dissertation submitted in partial fulfilment of
the requirements for the
Bachelor of Engineering (Hons)
(Mechanical)

JANUARY 2015

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF ORIGINALITY

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

Muhammad Afiq Bin Abd Rashid

CERTIFICATION OF APPROVAL

Oil Demand Forecasting in Malaysia in Transportation Sector using Artificial Neural Network (ANN)

by

Muhammad Afiq Bin Abd Rashid
14935

A project dissertation submitted to the
Mechanical Engineering Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirement for the
BACHELOR OF ENGINEERING (Hons)
(MECHANICAL)

Approved by,

____

Dr Jebaraj Sargunam Azariah

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
January 2015

ABSTRACT

Energy industry in Malaysia is one of critical sector that plays an important role in contributing the nation economic growth. The main energy source in Malaysia is from the petroleum and natural gas while the sector that consumed the most energy is transportation sector. Since both of these are the main energy source and consumer, a forecasting model is required to be developed to provide the oil demand forecast in transportation sector. This research analyses different forecasting models including time series regression technique, Auto Regressive Integrated Moving Average (ARIMA), double moving average method, double exponential smoothing method, triple exponential smoothing method and Artificial Neural Network (ANN) model (Univariate and Multivariate) to predict the future oil demand in transportation sector in Malaysia. In order to select the best forecasting model, the model validation is done using the error analysis technique such as Root Mean Square Error (RMSE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE) and Correlation Coefficient (R^2). Based on the model validation result, it is found that the Artificial Neural Network gives the least error in all of the error analysis techniques. Thus, Artificial Neural Network model is used to forecast the oil demand in transportation sector in Malaysia. The oil demand forecasted by the model in transportation sector in Malaysia for the year 2020, 2025 and 2030 are 559.44, 581.779 and 609.941 kg of oil equivalent respectively.

ACKNOWLEDGEMENT

The author would like to thank and extend his heartfelt gratitude to the persons who have made this Final Year Project as milestones for him to nurture his engineering skills and knowledge. First and foremost, the author would like to express his appreciation and praise to God for His guidance and blessing throughout this entire Final Year Project period.

Sincere gratitude goes to author's supervisor, Dr Jebaraj Sargunam Azariah, Senior Lecturer Mechanical Engineering for giving great opportunities and guidance. The knowledge, encouragement and support from Dr Jebaraj are the catalyst for the author to make sure the project is successfully completed. The co-operation is much indeed appreciated.

The author's appreciation would be incomplete without giving credit to all lectures involves in evaluating this project especially those from Mechanical Engineering Department. The author would like to express his appreciation to Dr Turnad Lenggo Ginta and Dr Rahmat Iskandar Khairul Shazi, course coordinators for Final Year Project I and II for assisting the author and the other students throughout the course period. Last but not least, special thanks to my beloved family and friends for all their support and helps throughout the semester.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION AND NOMENCLATURES

The following are frequently used abbreviations in this document and they shall be deemed to have the following meaning

ANN	Artificial Neural Network
AR	Auto Regressive
ARIMA	Auto Regressive Integrated Moving Average
CIA	Central Intelligence Agency
EIA	Energy Information Administration
FYP	Final Year Project
GDP	Gross Domestic Product
GP	Genetic Programming
LEAP	Long Range Energy Alternatives Planning
LRT	Light Rail Transit
LS-SVM	Least Square Support Vector Machine
MA	Moving Average
MAPE	Mean Absolute Percentage Error
MLR	Multiple Linear Regression
MPE	Mean Percentage Error
MR	Multiple Regression
R^2	Correlation Coefficient
RMSE	Root Mean Square Error
SPSS	Statistical Package for Social Science

CHAPTER 1

INTRODUCTION

1.1. Background of Study

Malaysia is the second-largest oil and natural gas producer in the Southeast Asia and the world second largest exporter of liquefied natural gas according to Energy Information Administration report. The oil reserves in Malaysia are the fourth highest in Asia Pacific after China, India and Vietnam. Most of the oil produced comes from the offshore fields. In terms of world largest oil production, Malaysia ranks at the 29^{th} in the world with the total production of 642 700 barrel per day. In the year 2013, Malaysia total imports of crude oil are about 160 500 barrel per day which ranks at the 37^{th} in the world for the highest world oil importer.

Although Malaysia are able to produce large number of crude oil and supply sufficient energy for their own demands without relying too much on the imports, these non renewable resources are depleting and major plans need to be considered to before energy consumption outpace the domestic production. In 2013, the total oil consumption in Malaysia are about 623 000 barrels per day. Malaysia total oil consumption increased from 465 024 barrel per day in 2000 to 598 340 barrel per day in 2010 which saw an increased by 77.72% within this ten years period (US Energy Information Administration, 2013).

Economic development and population growth for the past few decades have increased the demand for oil consumption by different sectors in Malaysia. One of the major sectors that are affected by this urbanisation is transportation sector. In good economic condition, people are able to buy more vehicles for their own convenience due to the high purchasing power and increase in household income. From year 2006 until year 2013, number of car registered in Malaysia has increased from 458294 cars to 655793 cars which saw an increase of 69.88% of total car registration within the six years period based on the statistics by Road Transport Department Malaysia.

Other reason that leads to increase of number of vehicles used in Malaysia is due to limited good and efficient public transport in certain areas. People have to rely on their vehicles to travel either within the short distances or long distances. Most of public transport such as Light Rail Transit (LRT), monorail, commuter and buses are provided in urban areas only while those living in rural areas have very limited access to these services.

The energy consumption in transportation sector occupies 37% of total energy consumption in 2012 while industrial sector consumes 30%, non energy 16%, residential and commercial sectors 15% and agriculture 2%, as shown in figure 1 (National Energy Balance, 2013). Among these sectors, transportation have seen highest increasing trend during the study period. Between the year 1980 and 2012, the energy consumption by transportation sector have increased by 15 043 ktoe between 2012 and 1980 follow by industrial sector 11 049 ktoe, non-energy 7203 ktoe, residential and commercial 6239 ktoe and agriculture 1052 ktoe as shown in figure 2. These data shown that the transportation has becoming major consumer of energy in Malaysia and this trend is expected to continue due to the current economic growth in Malaysia.

1.2. Problem Statement

Malaysia's energy production and development are the most important sector for the entire economy as it makes up almost 20% of the total gross domestic product. Starting in 2010, the tax rate and investment incentives have been introduced by Malaysia to promote more oil and natural gas exploration and development in the country's deepwater and marginal fields. These efforts aim to sustain the energy production in Malaysia and meet the future energy needs.

Developing and designing good energy demand in various sectors play an important role in both developed and developing countries for investors and policy makers. Underestimation of future energy demands may cause serious and dreadful consequence to the social life and economic growth of a country while overestimation would lead to inefficient and wasteful energy consumption.

1.3. Objectives

The purposes of this study are as following:

i. To develop forecasting models for oil consumption in transportation sector.

ii. To find the optimum forecasting model based on the error analysis techniques.

iii. To forecast the oil demand in transportation sector in Malaysia for years 2020, 2025 and 2030.

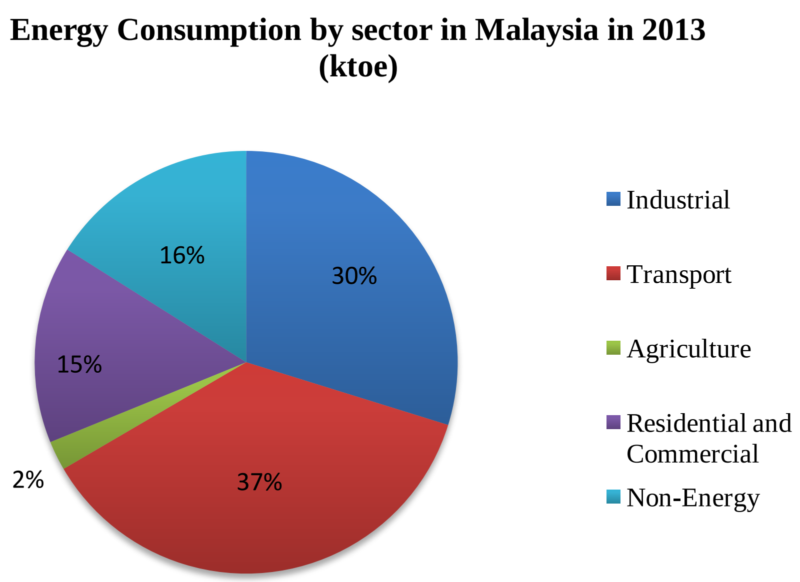

Figure 1: Energy consumption by sector in Malaysia in 2013

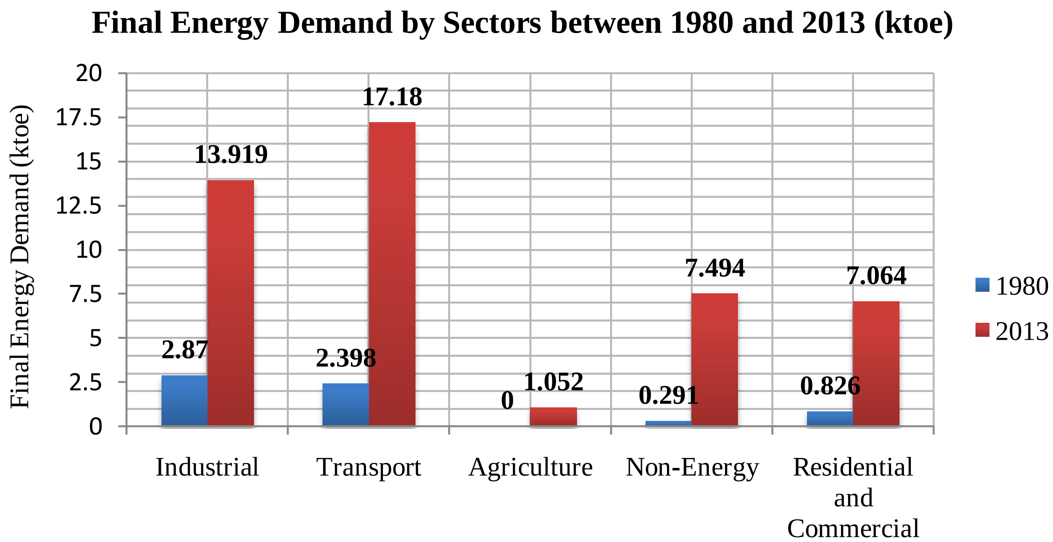

Figure 2: Final Energy Demand by Sectors between 1980 and 2013

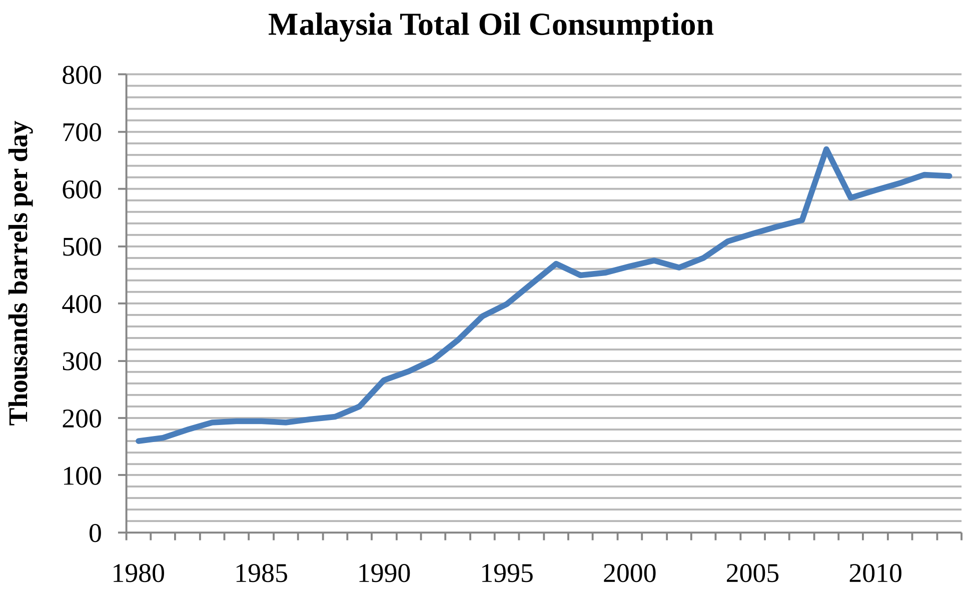

Figure 3: Malaysia Total Oil Consumption

CHAPTER 2

LITERATURE REVIEW

Artificial neural network is a structured mathematical model that consists of a huge number of simple processing units. These unites are interconnected to each other directly by synapses. The most vital characteristics of this energy model is its ability to generate, form and discover new knowledge automatically without any help similar to how the human are functioning [1]. The human brain works by transmitting signal from one neuron to another neuron. These neurons are able to collect and process the incoming electromagnetic signal, if the processed signal is not strong enough, it will not be further transmitted in the human brain. This allows the brains to continually process the data and enable the humans to think, recognize, imagine and act based on the transmitted signal information. The artificial neural network works on the same principle but is less complicated than the human brain [2].

The past research papers have developed various energy models such as energy planning models, energy supply-demand models, energy models and forecasting based on different technique including Artificial Neural Network and are presented here. A study on comparing regression analysis, artificial neural networks (ANNs), with the new technique called least squares support vector machines (LS-SVMs) in predicting the consumption of electricity energy in Turkey by [3] has proven that LS-SVM is more accurate. [4] conduct a study on annual electricity consumption in high energy consumption industrial sectors by means of artificial neural network (ANN) method. Based on the outcomes of the study, this method is proven to be more precise compared to conventional regression model. The developing trends of China's energy production and consumption in 2015 and 2020 are forecasted by [5] by applying grey forecasting model and novel Markov approach.[6]

forecasted the aggregated electricity demand of a group of signed up domestic consumers by means of a mathematical model. Structural time series model is applied to estimate Turkish residential electricity demand in 2020 [7]. [8] adopted panel cointegration approach in analyzing the demand for electricity and provides out-of-sample forecasting at the sectoral level. The forecasting of electricity demand for the state of Maharashtra, India in 2030 has been done by [9] via the method of Long Range Energy Alternatives Planning (LEAP). [10] developed Multiple regression (MR) model and genetic programming (GP) model in forecasting daily electricity consumption of an administration building located at the Southwark campus of London South Bank University in London.

2.1 Structure of Artificial Neural Network model

The typical structure of an Artificial Neural Network model for univariate or multivariate has three different layers as shown in figure 4. The first layer is the input layer consist input node depending on the number of variables used in the model. The second layer is the hidden layer and the number of nodes in hidden layer is depends on the user. The user can modified the number of nodes in order to find the least error for the model. Based on the past research paper, the suitable number of nodes for hidden layer is between five to ten nodes only. Building complex model will not ensure better forecasting result. The third layer is the output layer where the values of each node in hidden layer are summed and multiplied with their associated weights. For each of the input and hidden layer, they have their own bias value to stabilize the weightage of nodes in each layer.

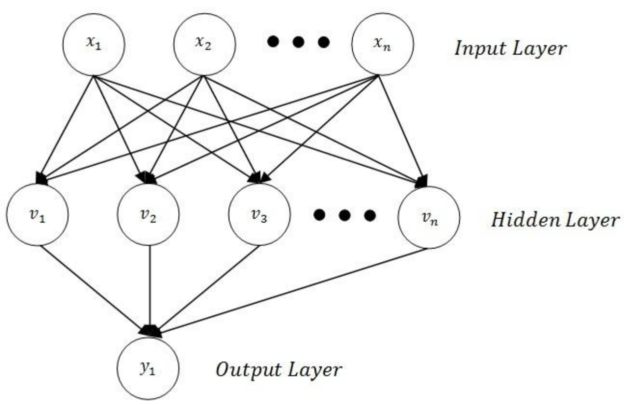

Figure 4: Structure of feed forward multilayer model

2.2 Advantages and disadvantages of Artificial Neural Network

The advantages of using artificial neural network model in predicting future energy consumption is it can handle nonlinearities among variables as the expected trends of energy consumption data is nonlinear. The ability of artificial neural network to develop a complex nonlinear relationship without needing any assumption from the nature of relationship makes this model suitable to use in different application and produce an accurate result. Although there have been lot of satisfactory result using artificial neural network as energy forecasting model, there are challenges faced by the researchers in developing the model for their own research studies. The researchers must be able to determine the suitable architecture of the model which are the number of layers, the number of nodes in each layer and numbers of arcs which connect the nodes to each other [1]. Another disadvantage of ANN is its characteristic of black box thinking. This approach does not enable the researchers to study the relationship between the input parameters and output parameters [2]. Another drawback of this method is the researchers need to use trial and error method in developing the best model due to its non-systematic technique to determine the network structure [11].

In order to get the most accurate result and prediction, the network parameter of an artificial neural network model should be adjusted by developing different types of model consisting of different variable. The ability of artificial neural network to model any nonlinear relationship enables it to forecast more accurate result compare to other energy models. [13] The report explains that the input parameters or independent variables should cover all possible variables that influence the output variable of interest.

Artificial neural network approach has been used by [2] to predict the future energy demand of Thailand by using gross domestic product, population and the numbers of registered vehicles as input parameters. Energy model using artificial neural networks haves various independent variables; [11] designed an ANN model using gross domestic product, population and transport amount (vehicle-kilometre) to forecast Turkey's transport energy demand, [12] chose ANN as a forecasting tool to predict the energy demand in United States and compare the result with the multiple linear regression (MLR) model by using gross domestic product and price of energy carries as independent variables, [13] developed a model for residential energy demand in United States using ANN technique and seven independent variables; resident population, gross domestic product, household size, median household income, cost of residential electricity, cost of residential natural gas and cost of residential heating oil.

Thus this study considers different types of independent variables based on the socioeconomic indicators such as gross domestic product, population, number of vehicle registration, export and imports of goods. Gross domestic product (Unit: 10^{10} US dollar) from 1980 to 2012, export and import of goods and services (Unit: US dollar) and population (Unit: million person) were obtained from World Bank Data (2014). Number of vehicle registration (Unit: million vehicles this includes automobile, bus, and truck but excluded two-wheel motorcycle) is obtained from Road Transport Department Malaysia (2014).

CHAPTER 3

METHODOLOGY

In this study, 10 different forecasting models were developed based on the past oil consumption data in transportation sector in Malaysia from the year 1980 until 2013. The models are time series regression methods including linear model, exponential model, power model and quadratic model, double moving average method, double exponential smoothing method, triple exponential smoothing method, Auto Regressive Integrated Moving Average (ARIMA) model, Artificial Neural Network (ANN) model (Univariate and Multivariate). The methodologies of each of the model are shown in figure 5 and discussed as follows.

3.1 Time Series Regression Methods

Time series regression method is an analysis used to find a best fit line from the historical data by minimizing the squared error between the forecast data and past data using the Microsoft excel software. Different types of models were used in this study to forecast the new data and steps are explained in the following section.

3.1.1 Linear Model

The linear model finds the best fitted line using a linear equation by plotting the historical data as a single variable. The equation for linear model is in the form of,

$$y = at + b$$

where y is the forecast data, t is the time period and a and b are the fixed constants.

3.1.2 Exponential Model

The exponential model forecast data by predicting the data to follow trend of increasing or decreasing at a constant growth rate. The model follows the exponential curve in form of,

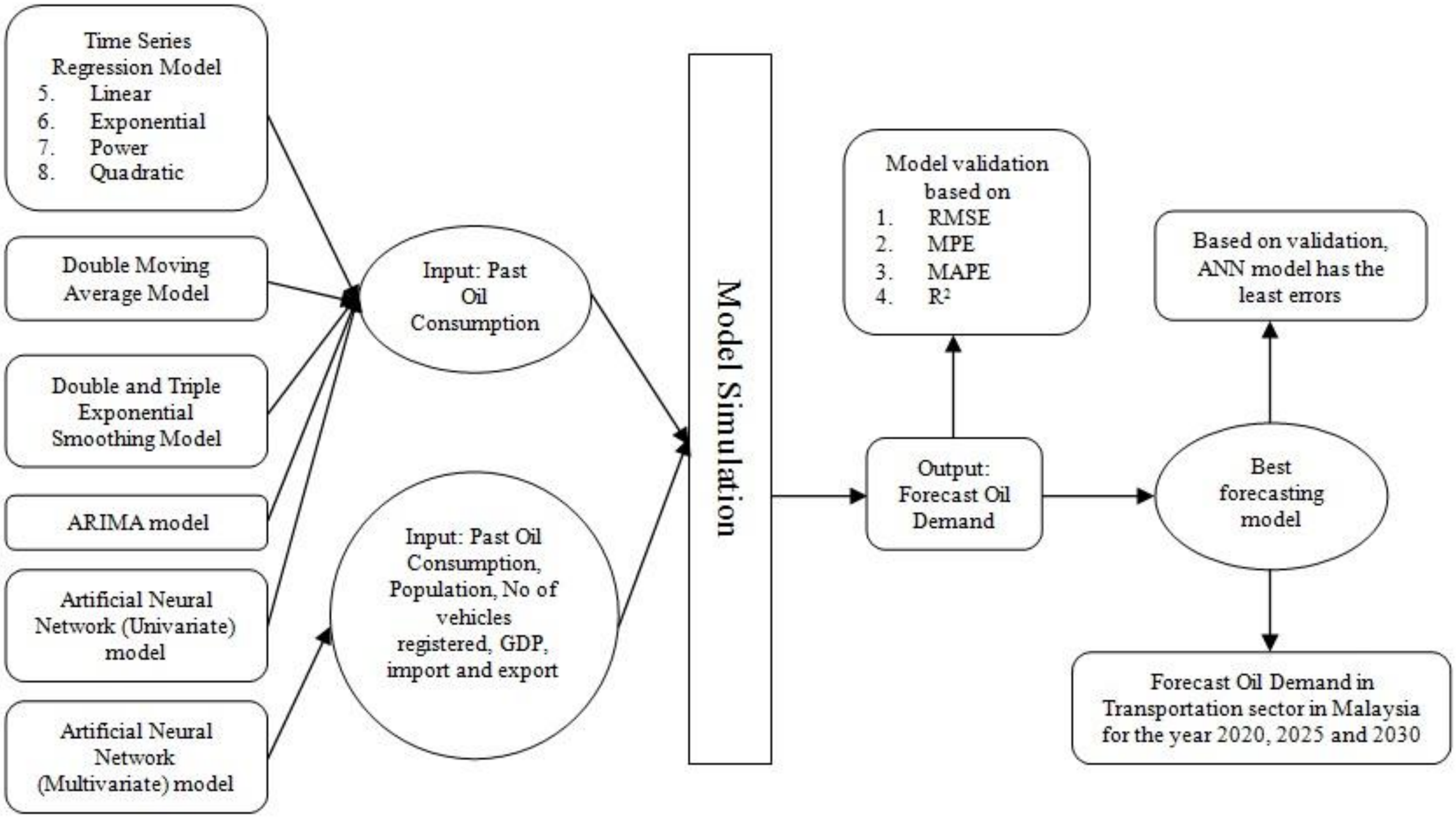

Figure 5: Schematic representation of formulation of oil demand forecasting models

$$y = e^{(a+bt)}$$

where y is the forecast data, t is the time period and a and b are the fixed constants

3.1.3 Power Model

Power model can be in the form of direct power model, inverse power model or quadratic model. In this study, the model used is direct power model since the data is increasing gradually over time and the equation for the model is as follow,

$$y = at^b$$

where y is the forecast data, t is the time period and a and b are the fixed constants.

3.1.4 Quadratic Model

Quadratic model is able to describe the path or trend of the data if the data follow a specific trend without any seasonal variation. The equation of the model is as follow,

$$y = at^2 + bt + c$$

where y is the forecast data, t is the time period and a, b and c are the fixed constants.

3.2 Double Moving Average Method

Moving average method is used for smoothing the time series to find the series trend and to measure the seasonal fluctuation of the data. The method is done by moving the arithmetic mean values of the data through the time series and applied twice in order to get the smoother time series. In this study, 3 year data are included in each moving average term.

$$y_{t+2} = \frac{y_t + y_{t+1} + y_{t+2}}{3}$$

The data are then plotted to find the best fit linear line and equation in the form of:

$$y = at + b$$

where y is the new forecast data using double moving average method and t are the time period starting from year 0. a and b are average change in Y and y-intercept based on the best fit line.

3.3 Double Exponential Smoothing Method

In the previous method, moving average assigned weight for each previous as equally. In contrast with exponential smoothing, the weight for previous data is decreasing gradually as the time series progress in order to forecast the new data. The formulate for exponential smoothing is as follow,

$$y_{t+1} = \alpha X_t + (1 - \alpha) y_t$$

where y_t is the forecast data, X_t is the actual data, t is the time period and α is the weighting factor from 0 to 1. The same step is applied twice to get new data. The new data is then plotted to find the best fit line based on linear model.

3.4 Triple Exponential Smoothing Method

Triple exponential smoothing method follow the same procedure as double exponential smoothing method but the step is repeated three times and the forecast equation follow the quadratic model.

3.5 Auto Regressive Integrate Moving Average (ARIMA) Model

The Auto Regressive Integrate Moving Average (ARIMA) model has two main structures to forecast any set of data. The first one is non stationary operator, the long term prediction is determined based on the different and the constant between the past data. The second structure is stationary operators which consist of AR and MA used to determine the short prediction value. The model predicts future value by combining the linear past value and series of error. The model was developed using SPPS Statistic software and model validation was done in Microsoft excel.

3.6 Artificial Neural Network (ANN) Model (Univariate and Multivariate)

In this study, the training process was done in MATLAB software and follows the feed-forward multilayer neural network with back propagation technique. For univariate model, there is only one input node which is the past oil consumption in transportation sector while for the multivariate model, there are six input nodes which are the past oil consumption in transportation sector, GDP, population, number of vehicles registered, imports and exports.

In error back propagation technique, there are different passes which are forward pass and backward pass [13]. The past data is divided into two parts, for training and validation purpose. During the training process, the weight for each node remains constant in forward pass. The data for each variable is applied to the input nodes and its effects are propagated through the network from input layer to hidden layer and finally to output layer. The weightage for each node is then adjusted based on the error correction rule in the backward pass to minimize the error in output layer. The training is done repeatedly to find the least error using the new weightage value for each node.

From the input to hidden layer, the activation transfer function used in the model is log sigmoid function and the formula is shown as follow,

$$v_j = \frac{1}{1 + \exp(-x_j)} \;, \; j = 1,2,3 \ldots n$$

Where v_j is value of node in the hidden network, x_j is total weighted sum of all inputs node plus the bias of neuron j. In the forward pass, the function signal at the output of neuron in each layer is computed as:

$$x_j(n) = \sum_{i=0}^{m} w1_{ij}(n) x_i(n)$$

where m is the total number of inputs excluding the bias inside the input layer and $w1_{ij}(n)$ is the synaptic weight connecting input nodes and hidden node and $x_i(n)$ is the input signal of input node. The index i refers to the input to the ith input terminal to the network. The activation transfer function from hidden layer to output layer is pure linear and the formula is shown as follow,

$$y_1 = \sum_{i=0}^{k} w2_{ij}(n) v_i(n)$$

where k is the total number of inputs excluding the bias inside the input layer and $w2_{ij}(n)$ is the synaptic weight connecting hidden node and output node and $v_i(n)$ is the hidden signal of hidden node. The index i refers to the input to the ith input terminal to the network. The error between the observed data and network output in back propagation technique is calculated as follow:

The back propagation technique during training process calculates the error between actual data and output of network and the error signal are as follows:

$$\delta_y = (d - y)y(1 - y)$$

where δ_y is the error in output layer between actual data and output of artificial neural network model in input layer, y is the actual data, and d is the output of artificial neural network model.

$$\delta_v = v(1 - v)\sum_y \delta_y w_{vy}$$

where δ_v is the error in hidden layer, v is the value in hidden node, y represent the number of output nodes and w_{vy} is weight of the connection from hidden layer to output layer.

$$\delta_x = x(1 - x)\sum_v \delta_v w_{xv}$$

where δ_x is the error in input layer, x is the value in input node, v represent the number of hidden nodes and w_{xv} is weight of the connection from input layer to hidden layer. Weight between hidden layer to output layer and input layer to hidden layer are adjusted as follow:

$$\Delta w_{vy} = \alpha\delta_y$$

$$w_{vy} = w_{vy} + \Delta w_{vy}$$

$$\Delta w_{xv} = \alpha\delta_y x$$

$$w_{xv} = w_{xv} + \Delta w_{xv}$$

where α is the learning rate, Δw_{vy} is the adjusted amount for weight connecting hidden layer to output layer and Δw_{xv} is the adjusted amount for weight connecting input layer to hidden layer. The training process stops when all w_{xv} in the previous iteration are below threshold.

3.7 Gantt Chart and Project Milestone for FYP I

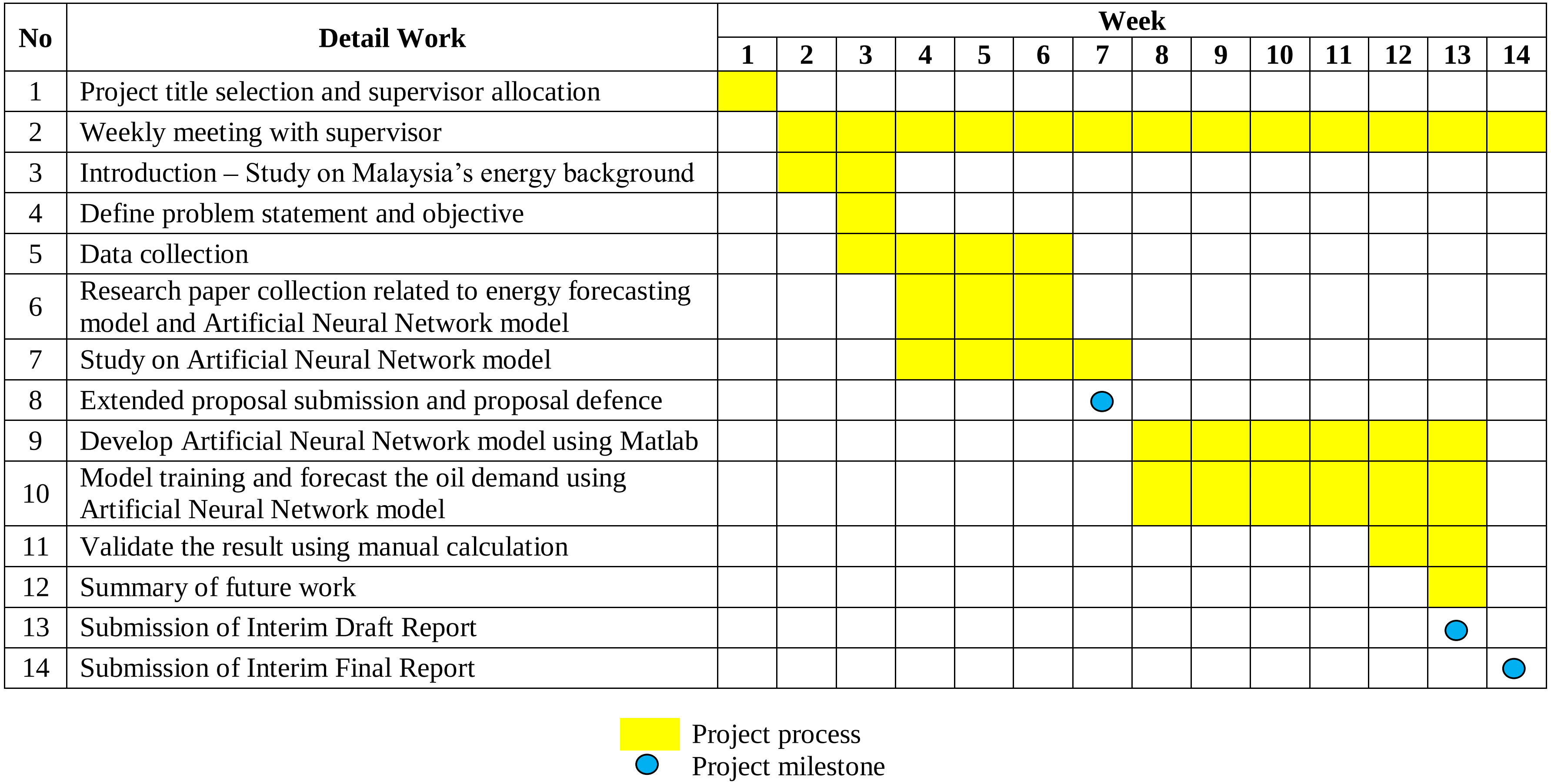

Figure 6: Gantt chart and project milestone for FYP I

3.8 Gantt Chart and Project Milestone for FYP II

No	Detail Work	Week													
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	Study on regression model, moving average model, exponential smoothing model and ARIMA model	■	■	■	■	■									
2	Develop time series regression model		■												
3	Develop double moving average model		■	■											
4	Develop double and triple exponential smoothing model		■	■											
5	Develop ARIMA model using SPSS software			■	■										
6	Model validation			■	■	■									
7	Forecast oil demand using best model					■	■								
8	Progress Report submission							●							
9	Report compilation								■	■	■	■	■	■	
10	Pre-SEDEX										●				
11	Draft Final Report submission											●			
12	Dissertation submission											●			
13	Technical Paper submission												●		
14	Viva													●	
15	Project Dissertation submission														●

■ Project process
● Project milestone

Figure 7: Gantt chart and project milestone for FYP II

CHAPTER 4

RESULT AND DISCUSSION

4.1 Artificial Neural Network Univariate Model

For the first forecasting model, the artificial neural network was developed using a single independent variable as input value which is the oil consumption by the transportation sector in Malaysia from 1980 to 2013 and target value is similar to the input value. The model and training process for each layer of the network was developed and done inside the MATLAB toolbox.

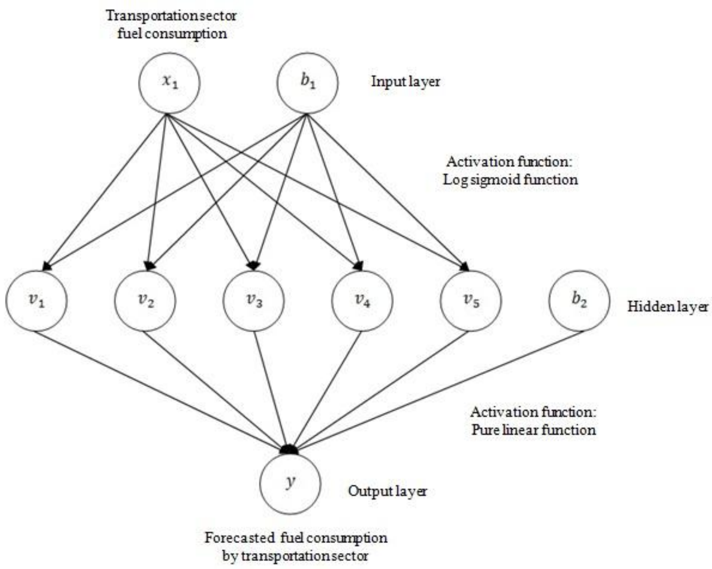

Figure 8: Artificial Neural Network structure for univariate model

The structure of the neural model is shown in figure 8; there are two different layers inside the network. The first hidden layer has five neurons and the activation function is log sigmoid function while the output layer is the summation of all values of neurons in hidden layer with the second bias values in one neuron using the pure linear function as the activation function.

The formula uses to calculate the value in hidden neurons and output neuron is shown as follow:

$for\ n = number\ of\ input\ data, \qquad i = 1, 2, \ldots.n.$

$$f(x) = logsig(x) = f(x) = \frac{1}{1 + e^{-(x)}}$$

Hidden neurons:

$v_1 = logsig\left(x_i w_{1,1} + b_{1,1}\right)$ $\qquad$ $v_2 = logsig\left(x_i w_{1,1} + b_{1,2}\right)$

$v_3 = logsig\left(x_i w_{1,3} + b_{1,3}\right)$ $\qquad$ $v_4 = logsig\left(x_1 w_{1,4} + b_{1,4}\right)$

$v_5 = logsig\left(x_i w_{1,5} + b_{1,5}\right)$

Output neuron:

$$y_i = v_1 w_{2,1} + v_2 w_{2,2} + v_3 w_{2,3} + v_4 w_{2,4} + v_5 w_{2,5} + b_{2,1}$$

For the year 1981, the fuel consumption by transportation sector in Malaysia is 156.698549156201 kg of oil per capita. In order to train this set of data in the neural network, normalization of these data should be done first. The method to normalize the data in 1981 is shown below.

The minimum value of data for this input variable is = 149.4895920763 kg of oil

The maximum value of data for this input variable is = 541.975949124 kg of oil

$$x' = \frac{x - minimum\ (x)}{maximum\ (x) - minimum\ (x)}$$

$$x' = \frac{156.698549156201 - 149.4895920763}{541.975949124 - 149.4895920763}$$

$$x' = 0.0183674080651541$$

The normalized value of input data in other years is shown in the table 5.

The next step is to train the data inside the MATLAB neural network toolbox in order to obtain the suitable weight and bias values (Table 1) for each neuron in each layer. The Levenberg-Marquardt back propagation technique is used in this training process. A set of code (Appendix A) is developed to simulate and calculate the output of this network based on the weight and bias values obtained from the training process. The sample calculation for data in year 1981 is shown as follow:

Table 1:
Weight and bias value for univariable neural network model

Weight	Input layer	Hidden layer
$w_{1,1}$	14.0645	0.078222
$w_{1,2}$	−8.8472	−0.072212
$w_{1,3}$	9.0747	0.027384
$w_{1,4}$	−0.8527	−4.6931
$w_{1,5}$	12.0945	0.013581

Bias	Input layer	Hidden layer
$b_{1,1}$	12.3106	1.9396
$b_{1,2}$	5.1445	
$b_{1,3}$	−1.884	
$b_{1,4}$	−0.3969	
$b_{1,5}$	11.5339	

Calculate the value of each neuron in hidden layer:

$$v_1 = logsig\left(x_i w_{1,1} + b_{1,1}\right)$$

$$v_1 = logsig\ [(0.0183674080651541)(14.0645) + 12.3106]$$

$$v_1 = 5.831101642302948 \times 10^{-6}$$

$$v_2 = logsig\left(x_i w_{1,1} + b_{1,2}\right)$$

$$v_2 = logsig\ [(0.0183674080651541)(-8.8472) + 5.1445]$$

$$v_2 = 0.99318662508329$$

$$v_3 = logsig\left(x_i w_{1,3} + b_{1,3}\right)$$

$$v_3 = logsig\ [(0.0183674080651541)(9.0747) + (-1.384)]$$

$$v_3 = 0.228407444451109$$

$$v_4 = logsig\left(x_1 w_{1,4} + b_{1,4}\right)$$

$$v_4 = logsig\ [(0.0183674080651541)(-0.8257) + (-0.39691]$$

$$v_4 = 0.39829597253455$$

$$v_5 = logsig\left(x_i w_{1,5} + b_{1,5}\right)$$

$$v_5 = logsig\ [(0.0183674080651541)(12.0945) + 11.5339]$$

$$v_5 = 0.999990053086533$$

Calculate the result for output layer

$$y_i = v_1 w_{2,1} + v_2 w_{2,2} + v_3 w_{2,3} + v_4 w_{2,4} + v_5 w_{2,5} + b_{2,1}$$

$$\begin{aligned} y_i = {} & (0.831101642302948 \times 10^{-6})\,(0.078222) \\ & + 0.99318662508329\,(0.072212) \\ & + 0.228407444451109\,(0.027384) \\ & + 0.39829597253455\,(-4.6931) \\ & + 0.999990053086533\,(0.013581) + 1.9396 \end{aligned}$$

$$\boldsymbol{y_i = 0.018456510532472}$$

The other result for this Neural Network is being tabulated in the table 5.

4.2 Artificial Neural Network Multivariate Model

For the second forecasting model, the artificial neural network was developed using multi variable input consist of the fuel consumption by transportation sector, population, gross domestic product, number of vehicles registered, import, and export of goods and services in Malaysia from 1980 to 2013 and target value is the fuel consumption by transportation sector. The model and training process for each layer of the network was developed and done inside the MATLAB toolbox.

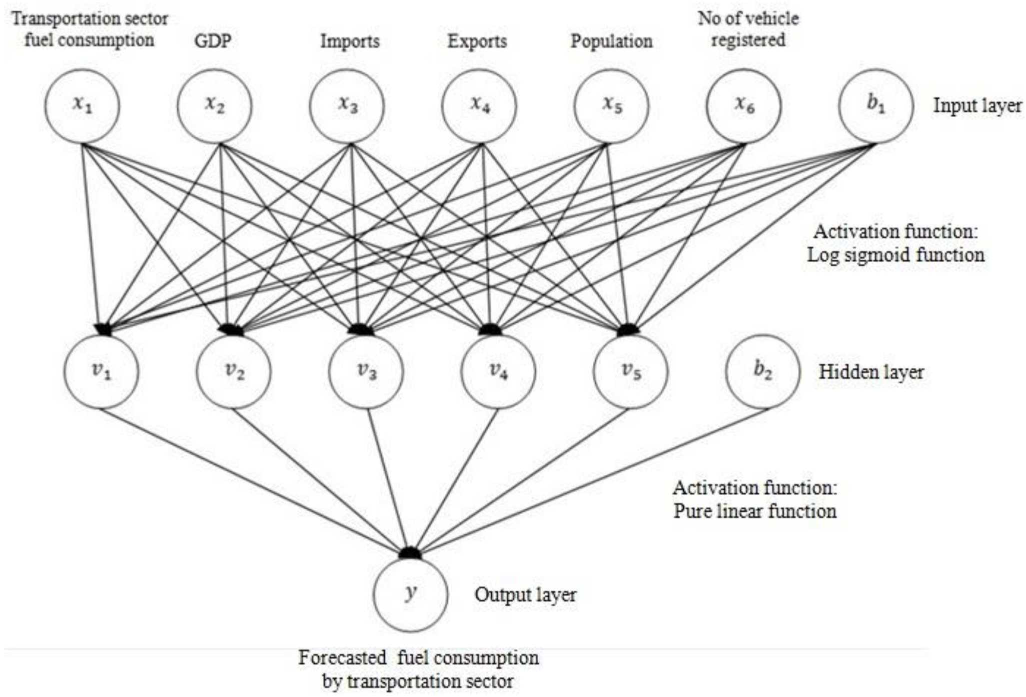

Figure 9: Artificial Neural Network structure for multivariate model

The structure of the neural model is shown in figure 9; there are two different layers inside network. The first hidden layer has five neurons and the activation function is log sigmoid function while the output layer is the summation of all values in hidden layer with the second bias values in one neuron using the pure linear function as the activation function. The formula uses to calculate the value in hidden neurons and output neuron is shown as follow:

$$for\ n = number\ of\ input\ data\ for\ each\ variable, \qquad i = 1, 2, \dots . n.$$

$$f(x) = logsig(x) = f(x) = \frac{1}{1 + e^{-(x)}}$$

Hidden neurons:

$$v_1 = logsig\left(x_{1,3}w_{1,1} + x_{2,3}w_{2,1} + x_{3,3}w_{3,1} + x_{4,3}w_{4,1} + x_{5,3}w_{5,1} + x_{6,3}w_{6,1} + b_{1,1}\right)$$

$$v_2 = logsig\left(x_{1,3}w_{1,2} + x_{2,3}w_{2,2} + x_{3,3}w_{3,2} + x_{4,3}w_{4,2} + x_{5,3}w_{5,2} + x_{6,3}w_{6,2} + b_{1,2}\right)$$

$$v_n = logsig\left(x_{1,n}w_{1,n} + x_{2,n}w_{2,n} + x_{3,n}w_{3,n} + x_{4,n}w_{4,n} + x_{5,n}w_{5,n} + x_{6,n}w_{6,n} + b_{1,n}\right)$$

Output neuron:

$$y_i = v_1w_{2,1} + v_2w_{2,2} + v_3w_{2,3} + v_4w_{2,4} + v_5w_{2,5} + b_{2,1}$$

Input data in year 1982 is used to show as the sample calculation. The method to normalized data for each input variable is similar to the previous univariable data. Listed below are the actual and normalized data for each input in year 1982:

Table 2:
Actual data and normalized data for input variables in multivariate Neural Network

Input variables	Actual data	Normalized data
Road sector fuel consumption	162.4083746	0.03291524
Gross domestic product	27287163523	0.008174369
Population	14543585	0.044691551
Export of goods and services	13679124040	0.002180073
Import of goods and services	16036111133	0.011582369
Number of vehicles registered	87542.8	0.017666949

The next step is to train the data inside the MATLAB neural network toolbox in order to obtain the suitable weight and bias values (Table 3 and 4) for each neuron in each layer. The Levenberg-Marquardt back propagation technique is used in this training process. A set of code (Appendix B) is developed to simulate and calculate the output of this network based on the weight and bias values obtained from the training process. The sample calculation for data in year 1982 is shown as follow:

Table 3:
Weight value for each input in hidden layer for multivariate neural network

Weight	Input layer					
	Input 1	**Input 2**	**Input 3**	**Input 4**	**Input 5**	**Input 6**
$w_{i,1}$	3.52450	-0.20794	0.23415	-1.91462	2.09491	-0.27475
$w_{i,2}$	1.61200	-1.67343	3.75518	-3.55421	-3.83340	-0.21691
$w_{i,3}$	-4.01079	-2.91686	-1.47103	-0.51327	-0.94176	-4.18804
$w_{i,4}$	0.82734	0.00055	-0.00196	0.00953	-0.01015	0.00313
$w_{i,5}$	2.25809	0.04048	0.01478	-0.07678	0.06683	-0.03925

Table 4:
Weight and bias value for multivariate neural network

Weight	Hidden layer
$w_{1,1}$	0.079821
$w_{1,2}$	-0.001427
$w_{1,3}$	0.004451
$w_{1,4}$	4.622908
$w_{1,5}$	0.590693

Bias	Input layer	Hidden layer
$b_{1,1}$	-3.80569	-2.492455
$b_{1,2}$	-4.52599	
$b_{1,3}$	-1.22905	
$b_{1,4}$	-0.33475	
$b_{1,5}$	2.91326	

Calculate the value of each neuron in hidden layer:

$$v_1 = logsig\left(x_{1,3}w_{1,1} + x_{2,3}w_{2,1} + x_{3,3}w_{3,1} + x_{4,3}w_{4,1} + x_{5,3}w_{5,1} + x_{6,3}w_{6,1} + b_{1,1}\right)$$

$$\begin{aligned} v_1 = logsig\ [&(0.03291)(3.52449) + (0.00817)(-0.207945) \\ &+ (0.04469)(0.23414) + (0.00218)(-1.91461) \\ &+ (0.01158)(2.09491) + (0.01766)(-0.27475) - 3.80568] \end{aligned}$$

$$v_1 = 0.0249485033108618$$

$$v_2 = logsig\left(x_{1,3}w_{1,2} + x_{2,3}w_{2,2} + x_{3,3}w_{3,2} + x_{4,3}w_{4,2} + x_{5,3}w_{5,2} + x_{6,3}w_{6,2} + b_{1,2}\right)$$

$$v_2 = logsig\,[(0.0329)(1.61199) + (0.00817)(-1.67342) + (0.04469)(3.75518) + (0.00218)(-3.55421) + (0.01158)(-3.83340) + (0.01766)(-0.21691) - 4.52598]$$

$$v_2 = 0.0124345437660615$$

$$v_3 = logsig(x_{1,3}w_{1,3} + x_{2,3}w_{2,3} + x_{3,3}w_{3,3} + x_{4,3}w_{4,3} + x_{5,3}w_{5,3} + x_{6,3}w_{6,3} + b_{1,3})$$

$$v_3 = logsig\,[(0.03291)(-4.01079) + (0.00817)(-2.91685) + (0.04469)(-1.47102) + (0.00218)(-0.51326) + (0.01158)(-0.941760) + (0.01766)(-4.18804) - 1.229052]$$

$$v_3 = 00.17701968141957$$

$$v_4 = logsig(x_{1,3}w_{1,4} + x_{2,3}w_{2,4} + x_{3,3}w_{3,4} + x_{4,3}w_{4,4} + x_{5,3}w_{5,4} + x_{6,3}w_{6,4} + b_{1,4})$$

$$v_4 = logsig\,[(0.03291)(0.827344) + (0.00817)(0.000550) + (0.04469)(-0.001963) + (0.00218)(0.00953441715766364) + (0.011582)(-0.010154) + (0.01766)(0.00312) + -0.334754]$$

$$v_4 = 0.423689151837986$$

$$v_5 = logsig(x_{1,3}w_{1,5} + x_{2,3}w_{2,5} + x_{3,3}w_{3,5} + x_{4,3}w_{4,5} + x_{5,3}w_{5,5} + x_{6,3}w_{6,5} + b_{1,5})$$

$$v_5 = logsig\,[(0.03291)(2.25809) + (0.008174)(0.040481) + (0.044691)(0.014779) + (0.00218)(-0.07678) + (0.01158)(0.066834) + (0.01766)(-0.03924) + 2.91326]$$

$$v_5 = 0.952051574260721$$

Calculate the result for output layer

$$y_i = v_1w_{2,1} + v_2w_{2,2} + v_3w_{2,3} + v_4w_{2,4} + v_5w_{2,5} + b_{2,1}$$

$$y_i = (0.024948)\,(0.0798207) + 0.012434\,(-0.0014267) + 0.177019\,(0.004451) + 0.423689\,(4.622908) + 0.952051\,(0.590693) + -2.4924545$$

$$\boldsymbol{y_i = 0.0313535578244641}$$

The other results for this Neural Network are shown in the table 6.

Table 5:
Output result for Univariate Neural Network model

Year	Input		Output	
	Actual Data	Normalized data	Normalized data	Data from network
1980	149.4895921	0.00000	2.7776E-05	149.5004938
1981	156.6985492	0.01837	0.018456511	156.7335207
1982	162.4083746	0.03292	0.033117793	162.4878741
1983	183.1583303	0.08578	0.086529023	183.451053
1984	184.1734846	0.08837	0.089145836	184.4781166
1985	188.7805008	0.10011	0.101023555	189.1399591
1986	195.7246704	0.11780	0.118928474	196.1673955
1987	201.5146526	0.13255	0.133852612	202.0249162
1988	214.0444803	0.16448	0.166103825	214.6830772
1989	224.712578	0.19166	0.193477397	225.4268307
1990	249.8476616	0.25570	0.257502125	250.5556632
1991	261.1442898	0.28448	0.286044532	261.7581684
1992	270.4488263	0.30819	0.309464188	270.9500639
1993	274.1935729	0.31773	0.31887252	274.6427058
1994	298.4676352	0.37958	0.379778654	298.5475323
1995	308.8002175	0.40590	0.405756367	308.7434303
1996	343.1824082	0.49350	0.492975424	342.9757205
1997	383.7963554	0.59698	0.597153344	383.8641326
1998	351.0167744	0.51346	0.51303206	350.8476763
1999	418.5001709	0.68540	0.68535037	418.4802621
2000	431.753875	0.71917	0.718900445	431.6482089
2001	458.0004248	0.78604	0.786481762	458.1729536
2002	459.4123937	0.78964	0.790191427	459.6289466
2003	479.7784743	0.84153	0.844539506	480.9598262
2004	505.0248395	0.90585	0.911871152	507.3865788
2005	491.9231809	0.87247	0.877234571	493.7921931
2006	459.9443509	0.79100	0.791591287	460.1783727
2007	481.9902752	0.84716	0.850502403	483.300182
2008	496.9902222	0.88538	0.890756002	499.0991705
2009	476.5327673	0.83326	0.835798469	477.5290885
2010	482.7443646	0.84909	0.852535793	484.0982597
2011	476.0949697	0.83214	0.834620758	477.0668528
2012	530.5012686	0.97076	0.974444684	531.9458362
2013	541.9759491	1.00000	0.999997222	541.9748589

Table 6:
Output result for Multivariate Neural Network model

Year	Input		Output	
	Actual Data	Normalized data	Normalized data	Data from network
1980	149.4895921	0.00000	-0.0013339	148.9660547
1981	156.6985492	0.01837	0.016929694	156.1342662
1982	162.4083746	0.03292	0.031353558	161.7954358
1983	183.1583303	0.08578	0.084140476	182.513581
1984	184.1734846	0.08837	0.086708418	183.5214631
1985	188.7805008	0.10011	0.098430823	188.1223471
1986	195.7246704	0.11780	0.116118716	195.0646039
1987	201.5146526	0.13255	0.130889786	200.8620474
1988	214.0444803	0.16448	0.162818902	213.3937899
1989	224.712578	0.19166	0.19002034	224.069983
1990	249.8476616	0.25570	0.254231569	249.2720145
1991	261.1442898	0.28448	0.283179387	260.6336379
1992	270.4488263	0.30819	0.307034503	269.9964455
1993	274.1935729	0.31773	0.316642427	273.7674248
1994	298.4676352	0.37958	0.378941277	298.2188734
1995	308.8002175	0.40590	0.405624525	308.6916842
1996	343.1824082	0.49350	0.493837087	343.3139112
1997	383.7963554	0.59698	0.598229774	384.2866168
1998	351.0167744	0.51346	0.513523514	351.0405652
1999	418.5001709	0.68540	0.686105818	418.7767651
2000	431.753875	0.71917	0.72030251	432.1985002
2001	458.0004248	0.78604	0.787231426	458.4671867
2002	459.4123937	0.78964	0.7908396	459.8833457
2003	479.7784743	0.84153	0.84230528	480.0829229
2004	505.0248395	0.90585	0.906327341	505.2107086
2005	491.9231809	0.87247	0.872390461	491.890946
2006	459.9443509	0.79100	0.791527267	460.1532456
2007	481.9902752	0.84716	0.847405197	482.0845708
2008	496.9902222	0.88538	0.883997916	496.4467138
2009	476.5327673	0.83326	0.832802727	476.3533004
2010	482.7443646	0.84909	0.848991119	482.7070235
2011	476.0949697	0.83214	0.832111606	476.0820449
2012	530.5012686	0.97076	0.970151491	530.2608165
2013	541.9759491	1.00000	0.999313956	541.7066864

4.3 Model Validation

In order to validate the result of each forecasting model which have been developed in the previous section, model validation is required based on the different error analysis method. Mean Percentage Error (MPE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Coefficient Factor Correlation (R^2) between the actual data and observed data were calculated for each model. The objective of this model validation is to find best model which will produce the least error for all of the error analysis method. The best model will then selected to forecast the future oil demand in transportation sector in Malaysia.

The forecast data from each model and actual data are plotted and shown in figure 10. The differences between the forecast data and actual data are then calculated to find the Mean Percentage Error (MPE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Coefficient Factor Correlation (R^2). The model validation result is shown in table 7. Based on table 7, the Artificial Neural Network (Multivariate) model shows the least error for each error analysis which are 0.440739, -0.1122, 0.153859 and 0.999995 for Mean Percentage Error (MPE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Coefficient Factor Correlation (R^2) respectively.

4.4 Oil demand Forecasting

The input variables used in this study to forecast future oil demand in transportation sector are past oil consumption, population, GDP, number of vehicles registered, imports and exports. As mentioned in the previous section, different types of forecasting methods were used to predict the future oil demand in transportation sector in Malaysia. In order to validate the forecast data in each method, the Mean Percentage Error (MPE), Root Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE) and Correlation (R^2) between the actual data and observed data were calculated. The error results for each model are tabulated in Table 7.

Based on Table 7, Artificial Neural Network (Multivariate) model show the least error for every error analysis. Thus the model was chosen to forecast the future oil demand in transportation sector in Malaysia for the years 2020, 2025 and 2030.

Table 7:
Comparison of forecasting models for oil consumption in transportation sector in Malaysia

No	Forecasting Model	Root Mean Square Error (RMSE)	Mean Percentage Error (MPE) %	Mean Absolute Percentage Error (MAPE)	Correlation Coefficient (R^2)
1	Linear Model ($y = 12.942t + 120.37$)	31.16955	-3.76893	6.587234	0.952479
2	Exponential Model ($y = 155.69e^{0.0412t}$)	45.04872	-3.49365	8.043056	0.896549
3	Power Model ($y = 96.545x^{0.4599}$)	42.41769	-5.48693	11.3304	0.922863
4	Quadratic Model ($y = -0.0936x^2 + 16.218x + 100.71$)	30.2105	-4.43066	7.855017	0.956306
5	ARIMA Model (0,1,0)	21.55421	-2.66135	3.826108	0.974674
6	Artificial Neural Network Model (Univariate)	0.870031	0.156976	0.166166	0.99998
7	Artificial Neural Network Model (Multivariate)	**0.440739**	**-0.112**	**0.153859**	**0.999996**
9	Double Moving Average Method ($y = 13.187x + 92.023$)	46.86917	-12.411	12.77605	0.952479
10	Double Exponential Smoothing Method ($y = 13.609x + 86.041$)	46.49413	-12.658	13.24366	0.952479
11	Triple Exponential Smoothing Method ($y = -0.0213x^2 + 15.119x + 34.661$)	77.25064	-25.5798	25.66361	0.953883

Figure 11 shows the forecast oil demand using the Artificial Neural Network (Multivariate) model. The predicted oil demand in transportation sector in Malaysia for the years 2020, 2025 and 2030 are 559.444, 581.779 and 609.41 kg of oil equivalent respectively.

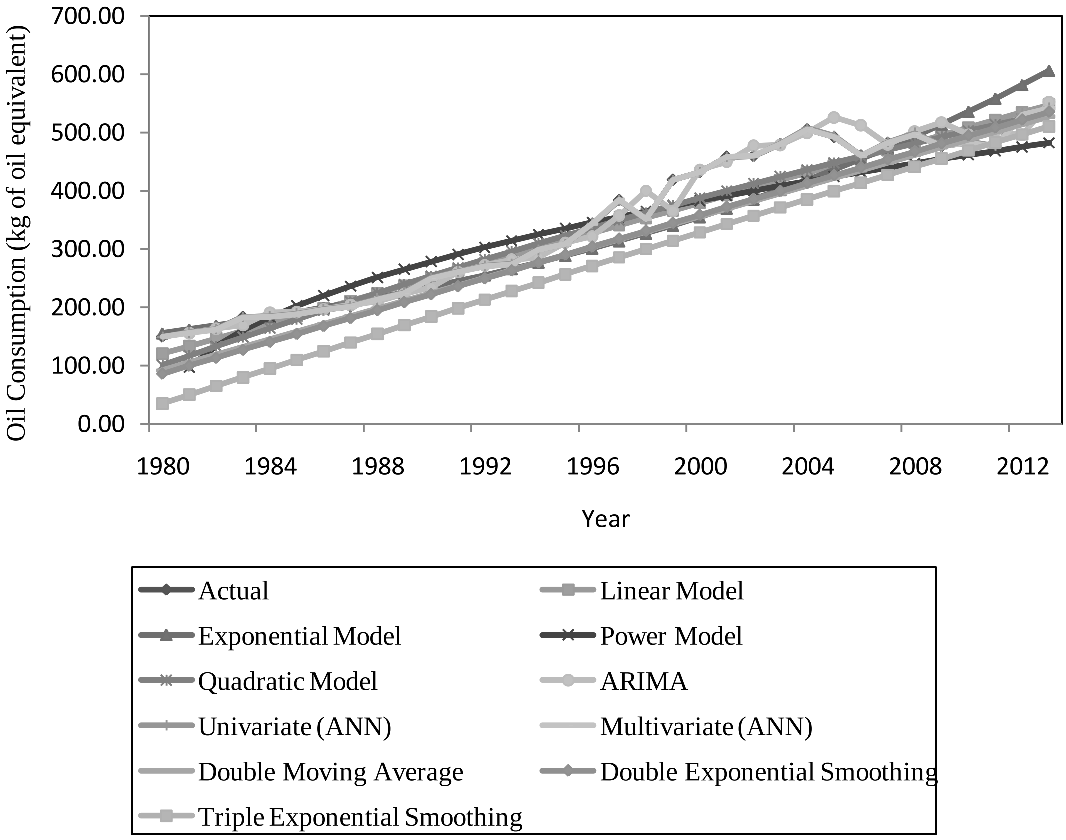

Figure 10: Validation of forecasting models for oil consumption in the transportation sector

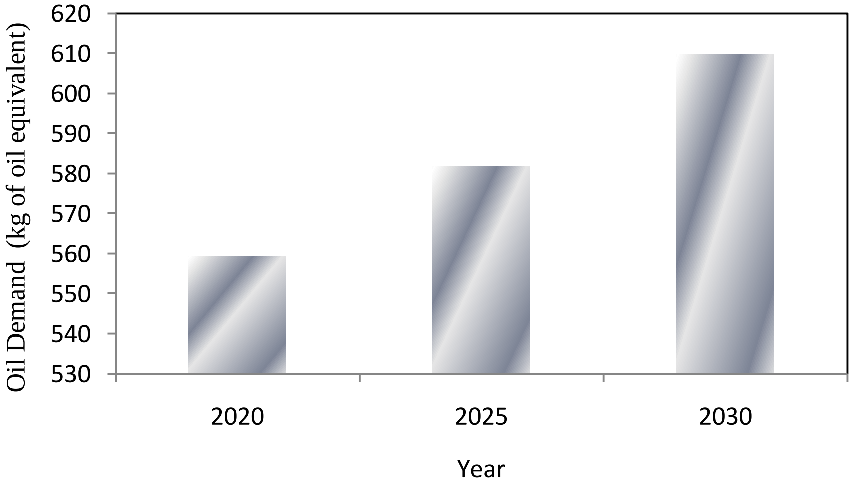

Figure 11: Forecast of oil demand in transportation sector in Malaysia

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

Different forecasting models were developed in this study in order to find the most suitable model to forecast the future oil demand in transportation sector in Malaysia until the year 2030. The input variables used in this study are past oil consumption in transportation sector, number of vehicles registered, population, GDP, imports and exports. Based on the observed result from each model, model validation was done to find the model which produces the least error for Mean Percentage Error (MPE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The Artificial Neural Network (Multivariate) model is found to show the least error and is used to forecast the future oil demand. The forecast oil demand in transportation sector in Malaysia using Artificial Neural Network (Multivariate) model for years 2020, 2025 and 2030 are 559.444, 581.779 and 609.41 kg of oil equivalent respectively. Hopefully this study will provide more insight for the policy makers to develop new policy related to energy consumption in transportation sector in Malaysia.

5.2 Recommendations

The recommendations for this project can be divided into two parts, first to improve the accuracy of the forecasting model output and the widen scope of study of the project. In order to get more accurate result, more different structure of Artificial Neural Network model should be designed and compared the result in model validation. The design of ANN model should be vary in the form of different structure of ANN model such as using different number of layers and varying the number of input variables. The model validation for different structure of artificial neural network model should be well documented. The output result for each of these ANN model should be compared with the other model. Other method is to use different input data to forecast the future demand. In this research, the input future data is estimated using linear regression method. The data should be supported by other method in the future.

The second part is to widen the scope of study of the project. Since this project only considers one sector which is transportation, it is suggested in the future to cover all other sectors that contribute the oil consumption in order to get the total oil demand in Malaysia. Other sectors are industrial sector, commercial sector and residential sector. The input variables would be different for different sectors and new models have to be developed for each sector.

REFERENCES

[1] M. Jahirul, R. Saidur, and H. Masjuki, "Predictability of artificial neural network (ANN) in performance prediction of a retrofitted CNG engine," *International Journal of Mechanical and Materials Engineering,* vol. 5, pp. 268-275, 2010.

[2] T. Limanond, S. Jomnonkwao, and A. Srikaew, "Projection of future transport energy demand of Thailand," *Energy Policy,* vol. 39, pp. 2754-2763, 5// 2011.

[3] F. Kaytez, M. C. Taplamacioglu, E. Cam, and F. Hardalac, "Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines," *International Journal of Electrical Power & Energy Systems,* vol. 67, pp. 431-438, 5// 2015.

[4] A. Azadeh, S. F. Ghaderi, and S. Sohrabkhani, "Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors," *Energy Conversion and Management,* vol. 49, pp. 2272-2278, 8// 2008.

[5] N.-m. Xie, C.-q. Yuan, and Y.-j. Yang, "Forecasting China's energy demand and self-sufficiency rate by grey forecasting model and Markov model," *International Journal of Electrical Power & Energy Systems,* vol. 66, pp. 1-8, 3// 2015.

[6] C. Gorria, J. Jimeno, I. Laresgoiti, M. Lezaun, and N. Ruiz, "Forecasting flexibility in electricity demand with price/consumption volume signals," *Electric Power Systems Research,* vol. 95, pp. 200-205, 2// 2013.

[7] Z. Dilaver and L. C. Hunt, "Modelling and forecasting Turkish residential electricity demand," *Energy Policy,* vol. 39, pp. 3117-3127, 6// 2011.

[8] A. El-Shazly, "Electricity demand analysis and forecasting: A panel cointegration approach," *Energy Economics,* vol. 40, pp. 251-258, 11// 2013.

[9] R. V. Kale and S. D. Pohekar, "Electricity demand and supply scenarios for Maharashtra (India) for 2030: An application of long range energy alternatives planning," *Energy Policy,* vol. 72, pp. 1-13, 9// 2014.

[10] K. P. Amber, M. W. Aslam, and S. K. Hussain, "Electricity consumption forecasting models for administration buildings of the UK higher education sector," *Energy and Buildings,* vol. 90, pp. 127-136, 3/1/ 2015.

[11] Y. S. Murat and H. Ceylan, "Use of artificial neural networks for transport energy demand modeling," *Energy Policy,* vol. 34, pp. 3165-3172, 2006.

[12] A. Kialashaki and J. R. Reisel, "Development and validation of artificial neural network models of the energy demand in the industrial sector of the United States," *Energy,* 2014.

[13] A. Kialashaki and J. R. Reisel, "Modeling of the energy demand of the residential sector in the United States using regression models and artificial neural networks," *Applied Energy,* vol. 108, pp. 271-280, 8// 2013.

APPENDICES

Appendix A

Output calculation for univariable network

```
%Obtain the weight value for each layers and neurons
weight = getwb(network);

%Initialize the number of input data
n = numel(input);

i = 1;
while i<n+1;
    Calculate the value for neuron 1 in hidden layer
    v(i,1) = logsig (input(i)*weight(1)+weight(6));

    Calculate the value for neuron 2 in hidden layer
    v(i,2)= logsig (input(i)*weight(2)+weight(7));

    Calculate the value for neuron 3 in hidden layer
    v(i,3) = logsig (input(i)*weight(3)+weight(8));

    Calculate the value for neuron 4 in hidden layer
    v(i,4) = logsig (input(i)*weight(4)+weight(9));

    Calculate the value for neuron 5 in hidden layer
    v(i,5) = logsig (input(i)*weight(4)+weight(10));

    Calculate the value of output neuron in output layer
    y(i) = v(i,1)*weight(11) + v(i,2)*weight(12) + v(i,3)*weight(13)
+...
    v(i,4)*weight(14) + v(i,5)*weight(15) + weight(16);
    i = i+1;
end
```

Appendix B

Output calculation for multivariable network

```
%Obtain the weight value for each layers and neurons
weight = getwb(network2);

%Initialize the number of input data
n = numel(input)/6;

i = 1;
while i<n+1;
    %Calculate the value for neuron 1 in hidden layer
    v(i,1) = logsig (input(1,i)*weight(1)+input(2,i)*weight(6)+...
        input(3,i)*weight(11)+input(4,i)*weight(16)+...
        input(5,i)*weight(21)+input(6,i)*weight(26)+weight(31));

    %Calculate the value for neuron 2 in hidden layer
    v(i,2) = logsig (input(1,i)*weight(2)+input(2,i)*weight(7)+...
        input(3,i)*weight(12)+input(4,i)*weight(17)+...
        input(5,i)*weight(22)+input(6,i)*weight(27)+weight(32));

    %Calculate the value for neuron 3 in hidden layer
    v(i,3) = logsig (input(1,i)*weight(3)+input(2,i)*weight(8)+...
        input(3,i)*weight(13)+input(4,i)*weight(18)+...
        input(5,i)*weight(23)+input(6,i)*weight(28)+weight(33));

    %Calculate the value for neuron 4 in hidden layer
    v(i,4) = logsig (input(1,i)*weight(4)+input(2,i)*weight(9)+...
        input(3,i)*weight(14)+input(4,i)*weight(19)+...
        input(5,i)*weight(24)+input(6,i)*weight(29)+weight(34));

    %Calculate the value for neuron 5 in hidden layer
    v(i,5) = logsig (input(1,i)*weight(5)+input(2,i)*weight(10)+...
        input(3,i)*weight(15)+input(4,i)*weight(20)+...
        input(5,i)*weight(25)+input(6,i)*weight(30)+weight(35));

    %Calculate the value output neuron in output layer
    y(i) = v(i,1)*weight(36) + v(i,2)*weight(37) +
v(i,3)*weight(38)+...
        v(i,4)*weight(39) + v(i,5)*weight(40) + weight(41);
    i = i+1;
end
```

Appendix C

Forecasted Oil Consumption data from 1980 to 2013 in Transportation sector using Multiple Linear Regression Model

Year	Oil consumption in Transportation Sector in Malaysia (kg of oil equivalent)				
	Actual data	Forecasted data from the model			
		Linear model	Exponential model	Power model	Quadratic model
1980	149.49	120.37	155.69		100.71
1981	156.70	133.31	162.24	96.55	116.83
1982	162.41	146.25	169.06	132.79	132.77
1983	183.16	159.20	176.17	160.01	148.52
1984	184.17	172.14	183.58	182.65	164.08
1985	188.78	185.08	191.30	202.39	179.46
1986	195.72	198.02	199.35	220.09	194.65
1987	201.51	210.96	207.74	236.26	209.65
1988	214.04	223.91	216.47	251.22	224.46
1989	224.71	236.85	225.58	265.21	239.09
1990	249.85	249.79	235.07	278.37	253.53
1991	261.14	262.73	244.95	290.85	267.78
1992	270.45	275.67	255.26	302.72	281.85
1993	274.19	288.62	265.99	314.07	295.73
1994	298.47	301.56	277.18	324.96	309.42
1995	308.80	314.50	288.84	335.44	322.92
1996	343.18	327.44	300.99	345.54	336.24
1997	383.80	340.38	313.65	355.31	349.37
1998	351.02	353.33	326.84	364.78	362.31
1999	418.50	366.27	340.59	373.96	375.06
2000	431.75	379.21	354.91	382.89	387.63
2001	458.00	392.15	369.84	391.58	400.01
2002	459.41	405.09	385.39	400.05	412.20
2003	479.78	418.04	401.60	408.31	424.21
2004	505.02	430.98	418.50	416.38	436.03
2005	491.92	443.92	436.10	424.27	447.66
2006	459.94	456.86	454.44	431.99	459.10
2007	481.99	469.80	473.55	439.56	470.36
2008	496.99	482.75	493.47	446.97	481.43
2009	476.53	495.69	514.23	454.24	492.31
2010	482.74	508.63	535.86	461.38	503.01
2011	476.09	521.57	558.40	468.39	513.52
2012	530.50	534.51	581.88	475.28	523.84
2013	541.98	547.46	606.36	482.05	533.97

Appendix D

Forecasted Oil Consumption data from 1980 to 2013 in Transportation sector using double moving average method and double and triple exponential smoothing method.

Year	Oil consumption in Transportation Sector in Malaysia (kg of oil equivalent)			
	Actual data	Forecasted data from the model		
		Double Moving Average Method	Double Exponential Smoothing Method	Triple Exponential Smoothing Method
1980	149.49	92.02	86.041	34.661
1981	156.70	105.21	99.65	49.7587
1982	162.41	118.40	113.259	64.8138
1983	183.16	131.58	126.868	79.8263
1984	184.17	144.77	140.477	94.7962
1985	188.78	157.96	154.086	109.7235
1986	195.72	171.15	167.695	124.6082
1987	201.51	184.33	181.304	139.4503
1988	214.04	197.52	194.913	154.2498
1989	224.71	210.71	208.522	169.0067
1990	249.85	223.89	222.131	183.721
1991	261.14	237.08	235.74	198.3927
1992	270.45	250.27	249.349	213.0218
1993	274.19	263.45	262.958	227.6083
1994	298.47	276.64	276.567	242.1522
1995	308.80	289.83	290.176	256.6535
1996	343.18	303.02	303.785	271.1122
1997	383.80	316.20	317.394	285.5283
1998	351.02	329.39	331.003	299.9018
1999	418.50	342.58	344.612	314.2327
2000	431.75	355.76	358.221	328.521
2001	458.00	368.95	371.83	342.7667
2002	459.41	382.14	385.439	356.9698
2003	479.78	395.32	399.048	371.1303
2004	505.02	408.51	412.657	385.2482
2005	491.92	421.70	426.266	399.3235
2006	459.94	434.89	439.875	413.3562
2007	481.99	448.07	453.484	427.3463
2008	496.99	461.26	467.093	441.2938
2009	476.53	474.45	480.702	455.1987
2010	482.74	487.63	494.311	469.061
2011	476.09	500.82	507.92	482.8807
2012	530.50	514.01	521.529	496.6578
2013	541.98	527.19	535.138	510.3923

Appendix E

Forecasted Oil Consumption data from 1980 to 2013 in Transportation sector using ARIMA model, ANN Univariate model and ANN Multivariate model.

Year	Oil consumption in Transportation Sector in Malaysia (kg of oil equivalent)			
	Actual data	Forecasted data from the model		
		ARIMA model	ANN Univariate model	ANN Multivariate Model
1980	149.49		149.5005	148.9661
1981	156.70	155.67	156.7335	156.1343
1982	162.41	163.18	162.4879	161.7954
1983	183.16	169.12	183.4511	182.5136
1984	184.17	190.73	184.4781	183.5215
1985	188.78	191.78	189.14	188.1223
1986	195.72	196.58	196.1674	195.0646
1987	201.51	203.81	202.0249	200.862
1988	214.04	209.84	214.6831	213.3938
1989	224.71	222.89	225.4268	224.07
1990	249.85	234	250.5557	249.272
1991	261.14	260.17	261.7582	260.6336
1992	270.45	271.93	270.9501	269.9964
1993	274.19	281.63	274.6427	273.7674
1994	298.47	285.52	298.5475	298.2189
1995	308.80	310.8	308.7434	308.6917
1996	343.18	321.56	342.9757	343.3139
1997	383.80	357.36	383.8641	384.2866
1998	351.02	399.66	350.8477	351.0406
1999	418.50	365.53	418.4803	418.7768
2000	431.75	435.79	431.6482	432.1985
2001	458.00	449.59	458.173	458.4672
2002	459.41	476.93	459.6289	459.8833
2003	479.78	478.39	480.9598	480.0829
2004	505.02	499.61	507.3866	505.2107
2005	491.92	525.89	493.7922	491.8909
2006	459.94	512.25	460.1784	460.1532
2007	481.99	478.95	483.3002	482.0846
2008	496.99	501.91	499.0992	496.4467
2009	476.53	517.53	477.5291	476.3533
2010	482.74	496.22	484.0983	482.707
2011	476.09	502.69	477.0669	476.082
2012	530.50	495.76	531.9458	530.2608
2013	541.98	552.42	541.9749	541.7067

Appendix F

Forecasted Oil Consumption data from 2014 to 2030 in Transportation sector using Artificial Neural Network Multivariate model.

Year	Final Demands for Oil Consumption in Transportation Sector in Malaysia (kg of oil equivalent)
2014	544.099
2015	546.184
2016	548.112
2017	555.169
2018	555.512
2019	557.080
2020	559.444
2021	561.419
2022	565.688
2023	569.324
2024	577.909
2025	581.779
2026	584.968
2027	586.252
2028	594.581
2029	598.148
2030	609.941