

**SOLAR IRRADIANCE FORECASTING USING GLOBAL
POSITIONING SYSTEM (GPS) DERIVED ATMOSPHERIC
WATER VAPOUR**

NAZREENDRAN A/L ELUMALAI



UNIVERSITI TEKNIKAL MALAYSIA MELAKA

**SOLAR IRRADIANCE FORECASTING USING GLOBAL
POSITIONING SYSTEM (GPS) DERIVED ATMOSPHERIC
WATER VAPOR**

NAZREENDRAN A/L ELUMALAI

**This report is submitted in partial fulfilment of the requirements
for the degree of Bachelor of Electronic Engineering with Honours**

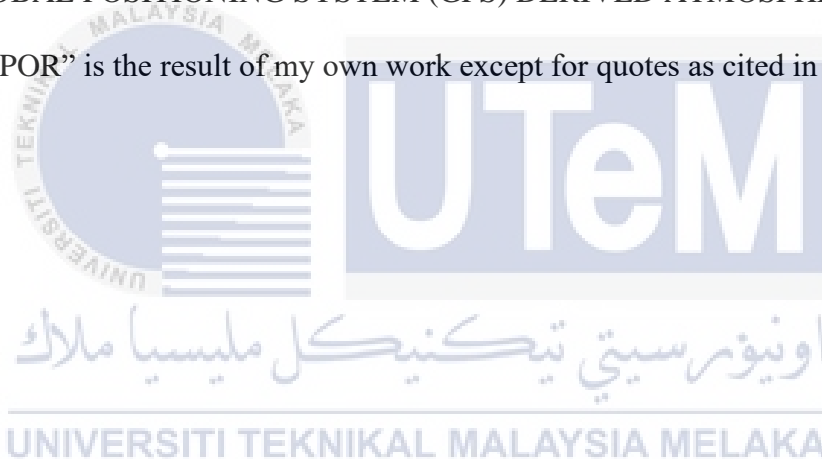


**Faculty of Electronic and Computer Engineering
Universiti Teknikal Malaysia Melaka**
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2020

DECLARATION

I declare that this report entitled “SOLAR IRRADIANCE FORECASTING USING GLOBAL POSITIONING SYSTEM (GPS) DERIVED ATMOSPHERIC WATER VAPOR” is the result of my own work except for quotes as cited in the references.



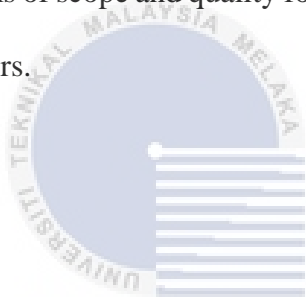
Signature :

Author : NAZREENDRAN A/L ELUMALAI

Date : 26/06/2020

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Bachelor of Electronic Engineering with Honours.



اونيورسيتي تيكنيكل مليسيا ملاك

Signature : _____
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Supervisor Name : DR. HO YIH HWA

Date : 25 JUN 2020

DEDICATION

This thesis is wholeheartedly dedicated to our beloved parents, who have been our source of inspiration and gave us strength when we thought of giving up, who constantly provide their moral, spiritual, emotional, and financial support.

Secondly, to Dr. Ho Yih Hwa our supervisor who have laid out the direction, tried to eliminate the blisters and to open our eyes to recognise the excellence of continuing the education. Thirdly, to our brothers, sisters, relatives, mentor, friends, and classmates who shared their words of advice and encouragement to finish this research. At last, I dedicated this thesis to the Almighty God, thank you for the guidance, strength, power of mind, protection, and skills and for giving us a healthy life. All of these, we offer to you.

ABSTRACT

The increase in contribution of renewable energy resources into the grid is a part of clever grid initiatives. The challenge of grid operators is to integrate the renewables such as solar energy into electrical energy. Because its intermittency due to weather variations. Despite this, the installed capacity of solar PV globally continues to increase. Thus, forecasting is becoming an important tool for the system grid operators to manage the solar photovoltaic (PV) energy production and satisfy the demand of energy consumers. Solar radiation measurement will be decreased when the atmospheric water vapor content is increased. This is due to absorption of solar radiation by water vapor in atmosphere and local weather activity. GPS data was used to find precipitable water vapor (PWV), zenith wet delay (ZWD) and zenith total delay (ZTD). Moreover, this project was investigated the interaction and relation between precipitable water vapor (PWV), weather data and solar irradiance were assessed. Solar irradiance was forecasted using two type of time series forecasting model in MATLAB.

ABSTRAK

Peningkatan sumbangan sumber tenaga boleh diperbaharui ke dalam grid adalah sebahagian daripada inisiatif grid pintar. Cabaran pengendali grid adalah untuk menggabungkan tenaga boleh diperbaharui seperti tenaga suria ke tenaga elektrik. Hal ini menyebabkan perubahan cuaca. Walau begitu, kapasiti pemasangan solar PV di seluruh dunia terus meningkat. Oleh itu, ramalan menjadi alat penting bagi pengendali grid sistem untuk menguruskan pengeluaran tenaga suria fotovoltaik (PV) dan memenuhi permintaan pengguna tenaga. Pengukuran sinaran suria akan berkurang apabila kandungan wap air meningkat. Ini disebabkan oleh penyerapan sinaran matahari oleh wap air di atmosfera dan aktiviti cuaca tempatan. Data GPS akan digunakan untuk mencari wap air yang dapat diendapkan (PWV), zenith wet delay (ZWD) dan zenith total delay (ZTD). Lebih-lebih lagi, projek ini akan menyiasat interaksi dan hubungan antara wap air yang dapat diendapkan (PWV), data cuaca dan penyinaran matahari akan dinilai dan selanjutnya digunakan dalam ramalan radiasi matahari.

ACKNOWLEDGEMENTS

The research discussed in this thesis was not complete without the support and encourage of countless people. First and foremost, I would like to express my deepest gratitude for this limitless and invaluable guidance to my supervisor Dr. Ho Yih Hwa. His supervision and guidance really help me with the project and the thesis advancing and smoothly completing. Even if busy with jobs and duties, he always manages to direct me and inspire me to complete this project.

I should also like to acknowledge the grateful guidance given by others supervisor and lecturer, and panels that discuss their experience and skills recommendation during the presentation of the project and give me the right path towards my project completion.

I would also like to express my sincere thanks and gratitude to my family and friends for their cooperation, motivation, positive advice and complete support not only for the thesis and project but also for my whole life at the Universiti Teknikal Malaysia Melaka. Ultimately, my sincere appreciation around to Universiti Teknikal Malaysia Melaka for giving me a great chance to continue my studies and to be successful completion of my research.

TABLE OF CONTENTS

Declaration	
Approval	
Dedication	
Abstract	i
Abstrak	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures	ix
List of Tables	xi
List of Symbols and Abbreviations	xii
List of Appendices	xiii
CHAPTER 1 INTRODUCTION	14
1.1 Background	14
1.1.1 Atmospheric Water Vapor	14
1.1.2 Atmospheric Propagation Delay	15
1.1.3 Tropospheric delay	15

1.2	Problem Statement	15
1.3	Objectives	16
1.4	Scope of work	16
1.5	Thesis	17
CHAPTER 2 BACKGROUND STUDY		19
2.1	Atmospheric Water Vapor	19
2.1.1	Water vapor	19
2.2	Techniques for Atmospheric Water Vapor Observation	20
2.2.1	Radiosondes	20
2.2.2	Water Vapor Radiometer	21
2.2.3	Global Positioning System	21
2.2.3.1	Ground-based GPS Receiver	21
2.2.3.2	Water Vapor from Satellite Based GPS Receivers	22
2.3	GPS System Architecture	23
2.4	GPS Receiver	25
2.5	GPS Antenna	26
2.6	GPS Signal	26
2.6.1	Coarse/Acquisition code	27
2.6.2	Precision code	27
2.7	GPS Observables	28

2.7.1	Doppler measurements	28
2.7.2	Code-Pseudorange Measurements	28
2.7.3	Carrier phase Measurements	29
2.8	Estimating Water Vapor from Ground-based GPS Observation	30
2.8.1	Zenith Total Delay	30
2.8.1.1	Zenith Wet Delay	31
2.8.1.2	Zenith Hydrostatic Delay	31
2.8.2	Precipitable Water Vapor	32
2.9	Solar Irradiance	33
2.9.1	Solar Smart Grid (PV) System	34
2.10	MATLAB Software	36
CHAPTER 3 METHODOLOGY		37
3.1	Flowchart of research project	38
3.2	Literature Review	39
3.3	Data processing of Rinex GPS Data and Weather Data	40
3.3.1	Research on Atmospheric Water Vapor and Ground Based GPS Receiver	40
3.3.2	Process flowchart for RINEX GPS data	40
3.3.3	RINEX GPS Data	41
3.3.3.1	Observation file	42
3.3.3.2	Navigation File	43

3.3.4	Data Processing of RINEX file	43
3.3.5	Data preparation to Estimate Water Vapor	44
3.3.5.1	RINEX GPS file process	44
3.3.5.2	Data preparation	46
3.3.6	Estimate Water Vapor	47
3.3.6.1	Zenith Total Delay	47
3.3.6.2	Zenith Hydrostatic Delay (Dry Delay)	49
3.3.6.3	Zenith Wet Delay	50
3.3.6.4	Precipitable Water Vapor	50
3.4	MATLAB Platform for Modelling	51
3.4.1	Neural Network Fitting	52
3.4.2	Prediction using LSTM and ARIMA model	54
3.4.2.1	LSTM Model	54
3.4.2.2	ARIMA Model	58
3.4.3	Predict Solar Irradiance using trained Neural Network	60
CHAPTER 4 RESULTS AND DISCUSSION		62
4.1	Introduction	63
4.2	Neural Network Fitting	63
4.3	Comparison between LSTM and ARIMA model of input prediction	67
4.3.1	Humidity Prediction	67

4.3.2	Temperature prediction	68
4.3.3	Pressure prediction	69
4.3.4	Water Vapor prediction	70
4.4	Comparison of solar irradiance prediction using indirectly and directly method	71
4.4.1	Predict the solar irradiance indirectly	72
4.4.1.1	Solar irradiance prediction using LSTM predicted inputs.	72
4.4.1.2	Solar irradiance prediction using ARIMA predicted inputs.	74
4.4.2	Predict the solar irradiance directly	76
4.4.2.1	Predict solar irradiance directly from LSTM model.	76
4.4.2.2	Predict solar irradiance directly from ARIMA model.	78
4.4.3	Comparison between indirectly and directly technique for solar prediction	80
CHAPTER 5 CONCLUSION AND FUTURE WORKS		82
5.1	Conclusion	82
5.2	Future works	83
REFERENCES		84
LIST OF PUBLICATIONS AND PAPERS PRESENTED		88
APPENDIX A		89
APPENDIX B		90
APPENDIX C		92

LIST OF FIGURES

Figure 2.1 Atmospheric surroundings by using a satellite-based GPS receiver	23
Figure 2.2 GPS satellite constellation - Make global or near-global continuous coverage.	24
Figure 2.3 Basic operation of control segment	25
Figure 2.4 Two clear days' solar radiation record	34
Figure 2.5 Global spectral irradiance as a wavelength variable of different water vapor values.	35
Figure 3.1 Flowchart of research project	38
Figure 3.2 Flowchart of estimate ZTD	41
Figure 3.3 RINEX observation file of November 11, 2019	42
Figure 3.4 Navigation file of November 11, 2019	43
Figure 3.5 RTKLIB software option window for double difference technique	44
Figure 3.6 Main interface of RTKLIB software	45
Figure 3.7 Post file of GPS data	46
Figure 3.8 Data collection in Excel file	47
Figure 3.9 Main window of RTKPOST	48
Figure 3.10 STAT file generated from RTKPOST with ZTD	49
Figure 3.11 Data For Estimation Of Water Vapor (PWV)	51
Figure 4.1 Correlation coefficient of Neural Network model	64

Figure 4.2 Performance of trained network model	65
Figure 4.3 Error histogram of trained network model	66
Figure 4.4 Humidity prediction using LSTM model	67
Figure 4.5 Humidity prediction using ARIMA model	68
Figure 4.6 Temperature prediction using LSTM model	68
Figure 4.7 Temperature prediction using ARIMA model	69
Figure 4.8 pressure prediction using LSTM model	69
Figure 4.9 pressure prediction using ARIMA model.	70
Figure 4.10 Water Vapor prediction using LSTM model	70
Figure 4.11 Water Vapor prediction using ARIMA model	71
Figure 4.12 Solar irradiance prediction for 1 hour ahead	72
Figure 4.13 Solar irradiance prediction for 3 hours ahead	73
Figure 4.14 Solar irradiance prediction from 10am to 7pm	73
Figure 4.15 Solar irradiance prediction for 1 hour ahead	74
Figure 4.16 solar irradiance prediction for 3 hours ahead	75
Figure 4.17 Solar irradiance prediction for 10am to 7pm	75
Figure 4.18 Solar irradiance prediction for 1 hour ahead	76
Figure 4.19 solar irradiance prediction for 3 hours ahead	77
Figure 4.20 Solar irradiance prediction for 9 hours (10am – 7pm)	78
Figure 4.21 solar irradiance prediction for 1 hour ahead	78
Figure 4.22 solar irradiance prediction for 3 hours ahead	79
Figure 4.23 solar irradiance prediction for 9 hours ahead	80

LIST OF TABLES

Table 1 comparison between indirectly and directly techniques with different predicted time and RMSE values. 80



LIST OF SYMBOLS AND ABBREVIATIONS

PWV : Precipitable Water Vapor

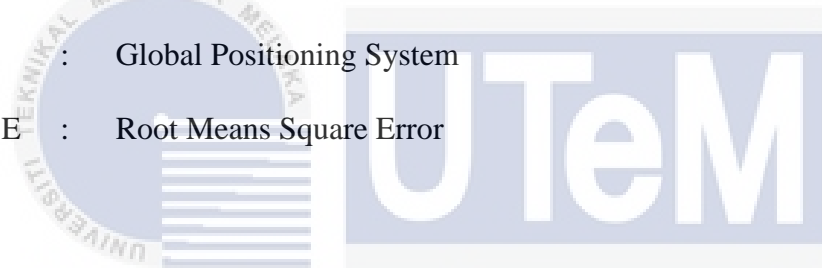
ZTD : Zenith Total Delay

ZHD : Zenith Hydrostatic Delay

ZWD : Zenith Wet Delay

GPS : Global Positioning System

RMSE : Root Means Square Error



اونيورسيتي تيكنيكل مليسيا ملاك

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF APPENDICES

Appendix A:	90
Appendix B:	91
Appendix C:	93



CHAPTER 1

INTRODUCTION



1.1 Background

1.1.1 Atmospheric Water Vapor

More than 99 percent of atmospheric moisture is in the form of water vapor, and this vapor is the main source of atmospheric energy that drives the development of short-term weather systems and influences the climate on longer time scales. Water vapor is a critical component of the climate systems of earth. It is the main greenhouse gas on Earth, trapping more heat than carbon dioxide. Water vapor diffusion and its related latent energy from vaporization is also responsible for nearly 50 percent of water transfer from the tropics to the poles. Typically, water vapor is the major source of differential atmospheric delay for measuring the GPS signals.

1.1.2 Atmospheric Propagation Delay

Delays in the propagation path induced by the atmosphere are major contributors to the error in measuring GPS. As radio waves travel through the atmosphere of the planet, the radio signals are significantly affected by the variation of their refractive index, which primarily induces the delay of delivery, usually referred to as the tropospheric delay.

1.1.3 Tropospheric delay

Refraction has a dry component and a wet component in the troposphere. The dry component that contributes most of the delay is closely correlated with the atmospheric pressure, perhaps 80 percent to 90 percent. It is less difficult to degree the dry component than the wet component. Luckily, the dry component detail leads to the larger part of the troposphere's array failure, as the value of the lag because of the wet component is based on the incredibly variable distribution of water vapor within the environment.

1.2 Problem Statement

The problem faced with PV system is mainly due to the climate condition. It will be affecting the performance of solar PV panel and the output of PV is incompatible values with rated values. Specifically, there are many factors that could influence the power of solar PV such as wind speed, wind direction, temperature, relative humidity and mainly focus on water vapor. This will cause difficult to knowing the actual power

output of solar PV due to multiplicity amount of solar radiation and the ambient environmental factors in certain location. Integration of solar energy into the electricity network is becoming essential because of its continually increasing growth in usage. An efficient use of the fluctuating energy output of photovoltaic (PV) systems requires reliable forecast information. In fact, this integration can offer a better quality of service if the solar irradiance variation can be predicted with great accuracy.

1.3 Objectives

The objectives of the projects are as follows.

- A) To determine PWV, ZWD & ZTD from GPS data
- B) To investigate the interaction between Precipitable Water Vapor (PWV), weather data and solar radiation measurement.
- C) To forecast the short-term solar irradiance with PWV estimated from GPS receiver network.

1.4 Scope of work

This project is intended to forecast solar irradiance by determine the water vapor content in atmosphere using GPS receiver and collect weather data. Determination of water vapor and weather data will help make solar irradiance forecast more reliable and reduce the cost of the system. To estimate these, need some technique to use in this project.

First part, 4 months of GPS data were used to determine the delays that occur and the water vapor in the atmosphere by using GPS data which is zenith total delay (ZTD) and zenith wet delay (ZWD). By using these two parameters, precipitable water vapor (PWV) was calculated.

Second part some parameters of local weather data were determined from weather station. Example parameters are temperature, pressure, solar irradiance, and humidity. Neural network in MATLAB software was used to investigate the relationship of water vapor & solar irradiance by creating a model using statistical approach. By using these steps, the solar irradiance was forecasted based on atmospheric water vapor content from GPS data and weather data.

1.5 Thesis

The material of this thesis is split into five chapters such as introduction, background study, methodology, result and discussion, and conclusion.

In chapter 1, the introduction of this project was explained. This starts by discussing about the project-related background studies. All the issues with the ventures were clarified after the context. The objectives are devised by examining the statement of the problem. It also clearly explains the project's scope or boundaries. In chapter 2, explains about the background studies related to the project that had been conducted. The background study will include the GPS meteorology, atmospheric delays, solar irradiance, water vapor and data processing method of GPS data.

In chapter 3, the flow of project methodology was explained, which is to determine zenith delays that occur in GPS signal and get values from GPS data. Precipitable water

vapor was computed using formula with some parameters includes in GPS data. Then, by determined the weather data, the relationship between water vapor and weather data was investigated. In chapter 4, was explained about the results of the project and briefly discussed about result obtained in project. Lastly, chapter 5 was explained the project conclusion includes with suggestion for future works have been reviewed in end of the chapter.



CHAPTER 2

BACKGROUND STUDY

2.1 Atmospheric Water Vapor

The environment of Earth is classified into five primary layers: the exosphere, the thermosphere, the mesosphere, the troposphere, and the stratosphere. The troposphere is the closest level to the surface of the Earth. It is 4 to 12 miles (7 to 20 km) thick and has half the atmosphere of the Earth [1]. Close to the ground, air is warmer, and it gets colder. Nearly all the atmosphere's water vapor and dirt are in this area, which is why clouds are located here.

2.1.1 Water vapor

Water vapor is a critical component of the environmental processes on Earth. It is the largest greenhouse gas on Earth, generating more energy than carbon dioxide. Knowledge of reliable quantification of water vapor and analysis of its physical characteristics is critical to understanding how the global climate environment, such

as global warming, changes over time [2]. In addition, the water vapor within the environment is an important influence in the solar radiation depletion that comes through the ecosystem to the Earth. The water vapor can soak up more than 10% of solar radiation traveling via the atmosphere, while the quantity of absorption depends on the water vapor inside the atmosphere. The rapid development of ground-based GPS networks provides a new source of information on water vapor. When ambient moisture affects the ozone refractivity, gaps in the satellite-receiver route include unique information about the overall troposphere and stratosphere water vapor. Consequently, GPS has become a standard PWV measuring technique [3].

2.2 Techniques for Atmospheric Water Vapor Observation

2.2.1 Radiosondes

One of the most common strategies for water vapor detection is radiosondes. This provides collection of instruments for simultaneous analysis and radio transmission of meteorological information, including atmospheric temperature, pressure, and humidity. As classic methods of collecting data on ambient water vapor, radiosondes typically do not provide the spatial and temporal precision needed for detailed weather and climate studies [4]. At most locations, these radiosondes are conducted twice a day and hundreds of miles away from each other. In comparison, configuration problems, poor data performance, and long-term performance influence these approaches [5].

2.2.2 Water Vapor Radiometer

During certain frequencies, radiometers passively evaluate the electromagnetic energy intensity. By scanning and measuring the temperature and brightness depending on the weather of the surface, the temperature of the sky and the dielectric constant of the surface, water vapor may be removed [6]. Water vapor radiometer uses high frequency power emission, usually between 23.8 and 31.4 GHz, comprising of antenna or scanning device, receiver, and data handling unit. However, water vapor radiometer operates only in certain weather condition with complicated calibration[4][7].

2.2.3 Global Positioning System

The GPS comprises of 28 satellites broadcasting L-band radio signals at 19- and 22-cm wavelengths on six orbital flights [8]. Despite being designed by the United States as a navigation aid. The amount of specialized civilian applications, the Air Force, has gradually increased over time. Civilian organizations from around the world have set up the International GPS Service (IGS) network, which comprises more than 100 international monitoring stations, offering 5-cm precision in orbit certainty to enable geodesic and geophysical testing activities. There are two main approaches for using GPS in the atmospheric sciences.

2.2.3.1 Ground-based GPS Receiver

Dual-frequency signals are obtained on ground-based receivers to acquire the transmission lag and therefore the combined water vapor along the GPS satellite route to the receiver. A potentially useful feature of GPS data is their very high temporal resolution (typically a few minutes), which has not been systematically explored so

far, but is expected to provide detailed information on fronts, squall lines, and other small-scale weather systems in the future [8]. In numerical weather prediction it could be used even more extensively than at present. Through comprehensive data assimilation, it is possible to better determine the high frequency transfer of liquid between the air and the ground of the earth [9].

2.2.3.2 Water Vapor from Satellite Based GPS Receivers

Atmospheric sounding is achieved by means of a radio occultation technique utilizing low-earth orbit satellites (LEOs) which, when they grow and set relative to GPS satellites, calculate the difference in GPS signal frequency. The Doppler-shifted frequency calculations are used to determine the radio wave's bending angles, which are a result of atmospheric refractivity [8]. Refractivity is a function of the ionosphere's electron density and atmospheric temperature, pressure, and water vapor. The occultation soundings offer important information about the global atmosphere's temperature and water vapor environments, and together with ground-based GPS devices can help to better define the regional and mesoscale circulation of water vapor [9].

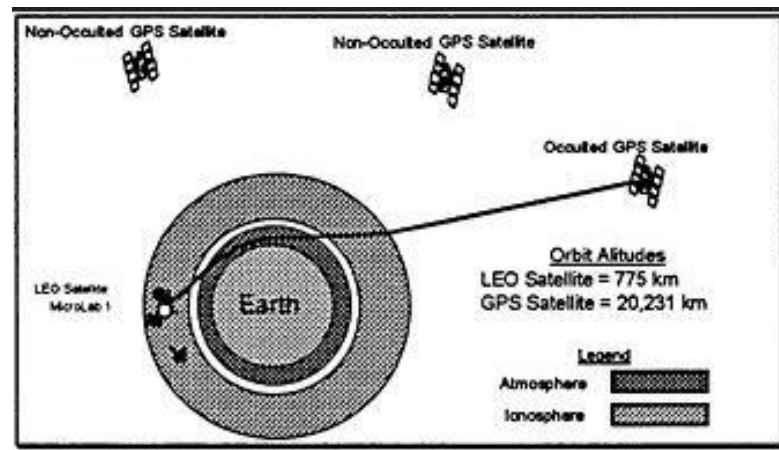


Figure 2.1 Atmospheric surroundings by using a satellite-based GPS receiver

2.3 GPS System Architecture

The GPS system contains of three main segments: Segment Space, Segment Control and Segment User [10].

The Space Segment consists of the satellites, their design, processing, and launch. The reference system consists of 24 operating satellites with a diameter of 26,559.7 km in near-circular orbits. In terms of the equatorial pole, the satellites are grouped in six orbital planes with an angle of inclination of 55 degrees. Currently undergoing satellite design that will provide more efficient signals in addition to improved signal quality, precision and credibility that will enable efficiency, navigation, and timing services. GPS Space Segment's main functions are to relay radio-navigation messages with a different signal layout, and to store and retransmit the Control Segment's navigation text. Such communications were operated on board the satellites by extremely stable atomic clocks.

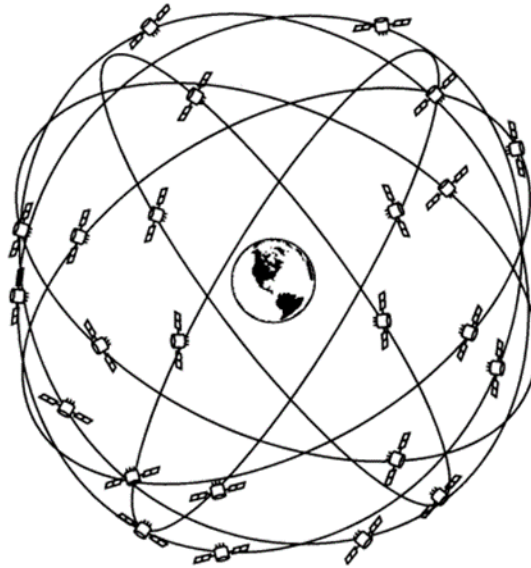


Figure 2.2 GPS satellite constellation - Make global or near-global continuous coverage.

The satellites were monitored and controlled by the ground-based network known as the Control Segment. Control Segment contains a master control station, a regional array of monitor stations (now 16 monitor stations are positioned at different locations on Earth), an intermediate master control station, and 12 monitoring and command antennas [10]. GPS satellite signals are tracked by control stations. Such signals are then sent to the master control station where operating requirements are reviewed and updated before the command signals are returned to the GPS satellites. By ground-based antennas they are sent out.

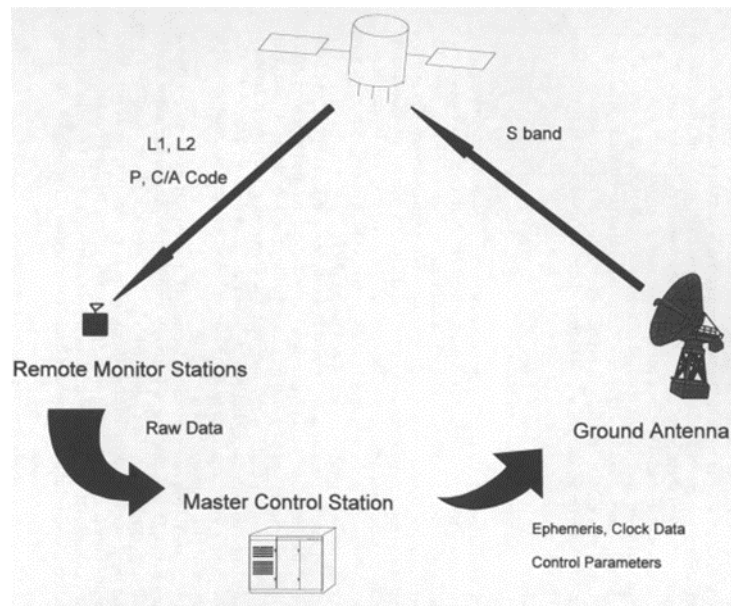


Figure 2.3 Basic operation of control segment

The third section, known as the User Section, is made up of various antennas and receivers that provide commercial and science consumers with location, velocity, and precise timing. In addition to the military use of the GPS system [10]. GPS civil uses are increasing, such as us: navigation, geodesy, scheduling, monitoring, surveillance, transportation, fishing, cadastral, and some more else. Their main function is to collect GPS signals, evaluate pseudo ranges (and other observables), and solve the navigation formulas to acquire their positions and provide a very precise time.

2.4 GPS Receiver

The GPS receivers are designed for high-quality raw GPS data processing, analysis, and transition. Powerful positioning system for GPS receivers which offers several advanced features. First of all, the main feature ensures measurements of high quality using various GNSS constellations: GPS, GLONASS, GALILEO, COMPASS and SBAS [11]. Moreover, The Cloud Server installed provides wireless internet access.

The raw data can be sent to an existing FTP server immediately and registered users can access the raw data files stored in the station directly. Other than that, the receiver frequently acts as a reference station or as a fixed or wireless base station for applications in real-time as well as post-processing. In fact, raw data is available in RINEX (Receiver Independent Exchange) format, the default format for handling GPS receiver measurements.

2.5 GPS Antenna

The GPS transmitter is designed with an antenna that is responsible for detecting the GPS satellite broadcast L-band signals. It contains several types of characteristics, the potential to transmit signals from multiple frequencies and multi constellations [11]. The captured signal is GPS which contains L1 C/A, L1/L2 P, L2C, and L5. In addition, the capacity to ignore multipath stimuli from close sources that are fading and the phase center's stability [9].

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2.6 GPS Signal

Global Positioning System (GPS) satellites relay microwave signals that enable GPS receivers to establish position and coordinated time on or near the Earth's surface. The U.S. runs the GPS system itself. Defense Department (DoD) used by the army as well as the public. GPS signals include signals ranging from satellite distance measurement to navigation messages. The navigation communications contain ephemeris details used to measure the satellite's location in space, as well as knowledge about the period and state of the whole satellite constellation, called the

almanac. The generated signals on board the satellites are based on or derived from the generation of a fundamental frequency of approximately 10.23 MHz. The signal is controlled by an atomic clock and has a range of 10⁻¹³ stability over a single day [12]. Two carrier signals are produced in the L-band, denoted as L1 and L2, by means of integer multiplications of as much as possible. Carriers L1 and L2 are biphasic-modulated codes to provide the receiver with satellite clock readings and to transmit data such as orbital parameters. There are two codes in the original GPS design which are the Coarse / Acquisition (C / A) code, which is freely available to the public, and the limited Precision (P) code, which is usually reserved for military applications [13].

2.6.1 Coarse/Acquisition code

The C / A code is a 1,023-bit deterministic code-named pseudorandom noise (also pseudorandom binary sequence) (PN or PRN code) which repeats each millisecond while transmitted at 1,023 megabits per second (Mbit / s). Also fit, or closely correlate such sequences when they are perfectly matched. That satellite transmits a special PRN code that is inconsistent with the PRN code of any other satellite [13].

2.6.2 Precision code

The P-code is also a PRN. Furthermore, the P-code PRN code for each satellite is 6,1871 ~1012 bits long (6,187,100,000,000 packets, ~720,213 gigabytes) and runs only once a week and transmitted at 10.23 Mbps [13]. The extreme range of the P-code improves its similarity benefit and prevents the variability of distance within the

Solar System. But the software is so long and complex that it was assumed that the receiver could not simply absorb and synchronize the signal on its own.

2.7 GPS Observables

2.7.1 Doppler measurements

Doppler measurements were based on the Doppler effect, a characteristic of the electrical signal frequency change induced by the relative motion of the emitter and receiver and used to calculate the user's velocity vector [14]. The Doppler measurements are typically obtained in the GPS receiver by measuring the inner oscillator frequency setting which measures the incoming signal cycle. The calculated Doppler shift is not the actual calculation itself, but the difference from the raw inner oscillator measurement without any Doppler shift.

2.7.2 Code-Pseudorange Measurements

Code-Pseudorange metric is a function of the length or distance between the GPS satellite and the GPS receiver, more specifically the distance between the antenna of the receiver and the antenna of the satellite. Also known as pseudo ranging is the GPS range determination technique, which is defined as the discrepancy between the time of obtaining the signal (as measured by the receiver clock) and the period of transmission at the satellite. The transmitting time is measured by the peak receiver code and GPS signal connection analysis [14]. When we presume that both the satellite and the receiver clock (which governs the production of the signal) are completely synchronized, the receiver should produce an exact replica of that frequency when the

PRN message is sent from the satellite. After a while, the receiver can "pick up" the transmitted code (which is equal to the signal transit time in space). The receiver will measure the signal travel time by matching the transmitted message with its replication. By combining this travel time with the speed of light, the range between the satellite and the receiver will be given. The simple observational formula for the calculation of pseudo range between satellite s and receiver r in frequency band j can be written as:

$$P_r^s(j) = p_r^s + c(\delta t_r - \delta t^s) + \delta_{iono}(j) + \delta_{trop} + \epsilon_{rx} + \epsilon_{mul} \quad (2.1)$$

where:

$P_r^s(j)$ is measured analysis of pseudo ranges.

δt_r and δt^s are the receiver and satellite clock biases,

ϵ_{rx} is failure attributed to interference from the receiver,

ϵ_{mul} is failure in the multipath calculation,

$\delta_{iono}(j)$ and δ_{trop} are ionosphere and tropospheric effect.

2.7.3 Carrier phase Measurements

The carrier phase is a calculation of the phase gap now of transmission between the transmitted satellite signal and the receiver-generated carrier signal. The range in this calculation is the sum of the cumulative number of full carrier cycles and fractional cycles at the receiver and the transmitter, compounded by the carrier wavelength,

where a loop is a maximum carrier wave and the uncertain amount of cycles of carrier phase measurement is called uncertainty [14]. Because the carrier phase wavelength (e.g. 19 cm for L1 frequency) is much shorter than the coding wavelength, the carrier phase calculation is more reliable, however carrier phase observables are required for an average carrier phase analysis.

2.8 Estimating Water Vapor from Ground-based GPS Observation

Atmospheric water vapor can be measured to refractivity (bent and slowed) by using ground-based GPS receivers. Atmospheric refractivity creates a lag in the GPS signal relative to that predicted if the atmosphere did not interfere. The ionosphere, an area of charged particles between 50 and 1000 km above the surface of the earth, is the greatest source of delays in atmospheric transmission. Furthermore, because the refraction in a loaded medium depends on the distance, this influence can be avoided by using dual-frequency GPS detection. The atmosphere is electrically neuronal (non-ionized) up to 50 km above the earth's surface. The main sources of signal lag, especially in the troposphere, are variations in temperature, pressure, and water vapor. Tropospheric lag (tropospheric refraction) can be measured and transformed into atmospheric water vapor in precise GPS data collection.

2.8.1 Zenith Total Delay

Zenith Total Delay (ZTD) GPS measurement that is connected to the water vapor in the atmosphere. In the vertical direction, it's a delay ZTD is important because it

helps us to predict the weather [15]. Zenith Hydrostatic Delay and Zenith Wet Delay are two major components of the ZTD [16].

2.8.1.1 Zenith Wet Delay

The Zenith Wet Delay is a wet part of the caused environment on the GPS signal and provides a way of detecting the Precipitable Water Vapor [16]. The ZWD refers explicitly to the PWV. The ZWD can thus be translated into the atmosphere's Integrated Water Vapor (PWV) material. ZWD is directly associated with temperature.

2.8.1.2 Zenith Hydrostatic Delay

The Zenith Hydrostatic Delay (ZHD), also known as Zenith Dry Delay, can be measured by using equation below from the local surface pressure [17][18].

$$ZHD = (2.779 \pm 0.00024) \frac{P_s}{f(\lambda H)} \quad (2.2)$$

Where,

P_s is the surface pressure and its unit in millibars

$f(\lambda H)$ is a factor accounting for the variation in gravitational acceleration with latitude and height

The ZHD is subtracted from ZTD to get the Zenith Wet Delay (ZWD).

2.8.2 Precipitable Water Vapor

Precipitable water vapor is a depth of water in a vertical column of air in the atmosphere of water vapor. At solar and terrestrial wavelengths, the water molecule creates strong bands of absorption. Gaseous absorption and release of thermal airborne radiation, for instance, is mainly due to bands of water vapor absorption below 7.6 μm , between 13 μm and 16 μm , and above 22 μm . Additionally, in the hydrological cycle, the vapor phase plays an important role, affecting the climate and weather systems. The existence attenuates airborne electromagnetic radiation, which is critical for astronomy, radar, telecommunications, and remote sensing. The temperature of the atmosphere is strongly affected by precipitable water vapor at all levels.

Assuming that water vapor is a perfect product, the absolute humidity q at sounding point z was obtained through a classic formula extracted from an ideal gas state equation [19]:

$$P_v(z) = \frac{217 \times RH(z) \times e(z, T)}{T} \quad (2.3)$$

Where.

T represents atmospheric temperature.

$RH(z)$ represents relative humidity of atmosphere.

$e(z, T)$ represents atmospheric pressure.

By using these parameters, can calculate precipitable water vapor (PWV) using standard formula that given below.

$$PWV = \int P_v(z) dz \quad (2.4)$$

2.9 Solar Irradiance

Solar irradiance is the Sun's source of energy. It can be measured over a square meter by energy per unit area. Such radiant power is calculated and called the radiance of the sun. Specific sunlight wavelengths hit specific Earth's surface. It thus induces different temperatures on the Earth's different surface [20]. We feel hot, warm, or cold owing to the visible and ultraviolet radiation that enters the Earth's surface. However, the measures of Solar Radiation are not easily available due to the cost of calculating instruments and the criteria of repair and adjustment. In current research, very important to develop methods for measuring solar radiation in order to use it for solar energy applications, including developing and evaluating energy conversion devices and architectural design [21]. The three elements of solar radiation are: Global Solar Radiation (GSR), Clear Solar Radiation and Diffuse Solar Radiation. By conduct research of Unit of Applied Research in Renewable Energy in Ghardaia, Algeria, using Global Solar Radiation (GSR) component collect some data of clear days and come out with solar irradiance versus time series.

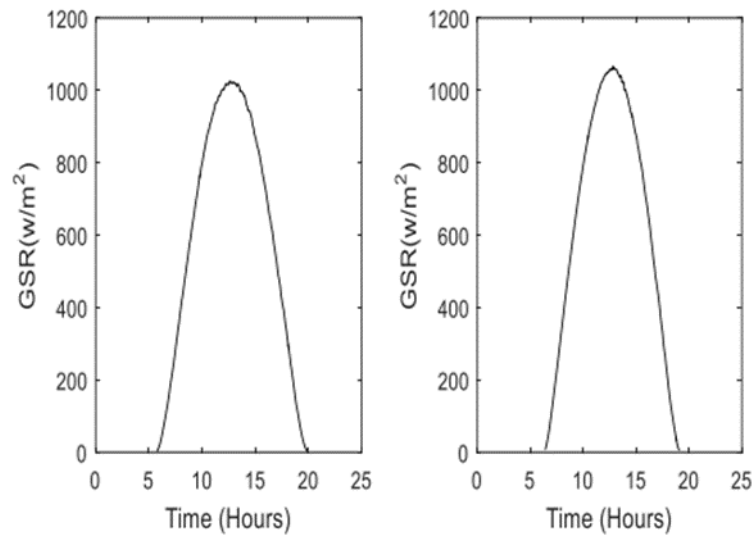


Figure 2.4 Two clear days' solar radiation record

Global surface solar radiation has increased by about 1.5 W m^{-2} over 30 years (in line with a solar power station's lifetime), representing a small percentage of the global average radiation of 190 W m^{-2} . The reanalysis shows some discrepancy for solar radiation, but the trend sign is consistent (positive over land, global negative) [22]. In recent times, solar energy supplies have changed, with changes in solar activity that change the global rate of solar production. This can have an impact on photovoltaic and concentrated solar power generation of electricity.

2.9.1 Solar Smart Grid (PV) System

Solar-grid integration is a network that makes a significant penetration of PV in the regional power grid. This is an important technology because the incorporation of regulated PV systems in grids optimizes the energy balance of homes, improves the economy of the PV sector, lowers operating costs, and adds value for customers and

utilities. Solar-grid integration is now a common practice in many countries around the world as demand for alternative clean energy as opposed to fossil fuel is increasing.

Variability of insolation is a major challenge. The amount of generation from Photovoltaic or PV systems at any location and time depends on the amount of insolation or sunlight. Instability on the grid could be induced by both under generation and over generation. Climate changes are main factor to the variability of insolation [23]. It will be affecting the performance of solar PV panel and also the output of PV is incompatible values with rated values. Specifically, there are many factors that could influence the power of solar PV such as wind speed, wind direction, temperature, relative humidity and mainly focus on water vapor. Figure 2.9.1.1, based on research, proved that the effect of water vapor on solar cells performance [24].

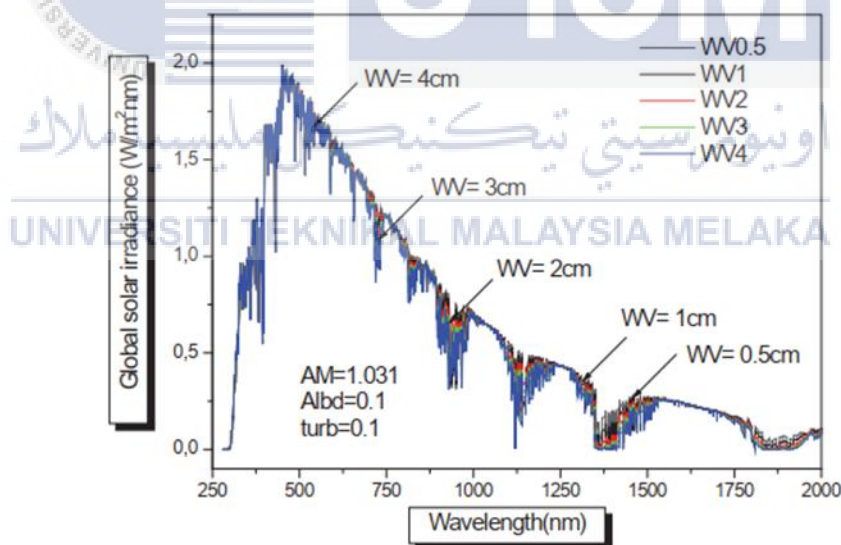


Figure 2.5 Global spectral irradiance as a wavelength variable of different water vapor values.

2.10 MATLAB Software

MATLAB is a high-performance programming language. This incorporates the framework of modeling, visualizing, and scripting. MATLAB is also a modern language framework for computing: it has complex data structures, provides built-in editing and debugging tools, and embraces object-oriented programming. These factors make MATLAB's perfect teaching and study platform. It has powerful built-in routines that allow for a wide range of computations [25]. It also has easy-to-use graphics commands that provide direct analysis of performance. Different programs are provided in the toolbox boxes. Signal processing toolboxes, mathematical computation, control theory, simulation, optimization, and several other areas in applied science and engineering are available.



CHAPTER 3

METHODOLOGY



The section would explain the approach specifics used to complete the task. Approach is described as the study of the concepts of processes, guidelines and a process or collection of procedures that were challenged for the completion of the project. Methodology is, in addition to research design, one of the key factors in determining the effectiveness of loss research.

There are many observations or approaches that are collected from the paper for future studies which contribute to this project. The method is used to accomplish the project's goal to achieve a successful outcome.

3.1 Flowchart of research project

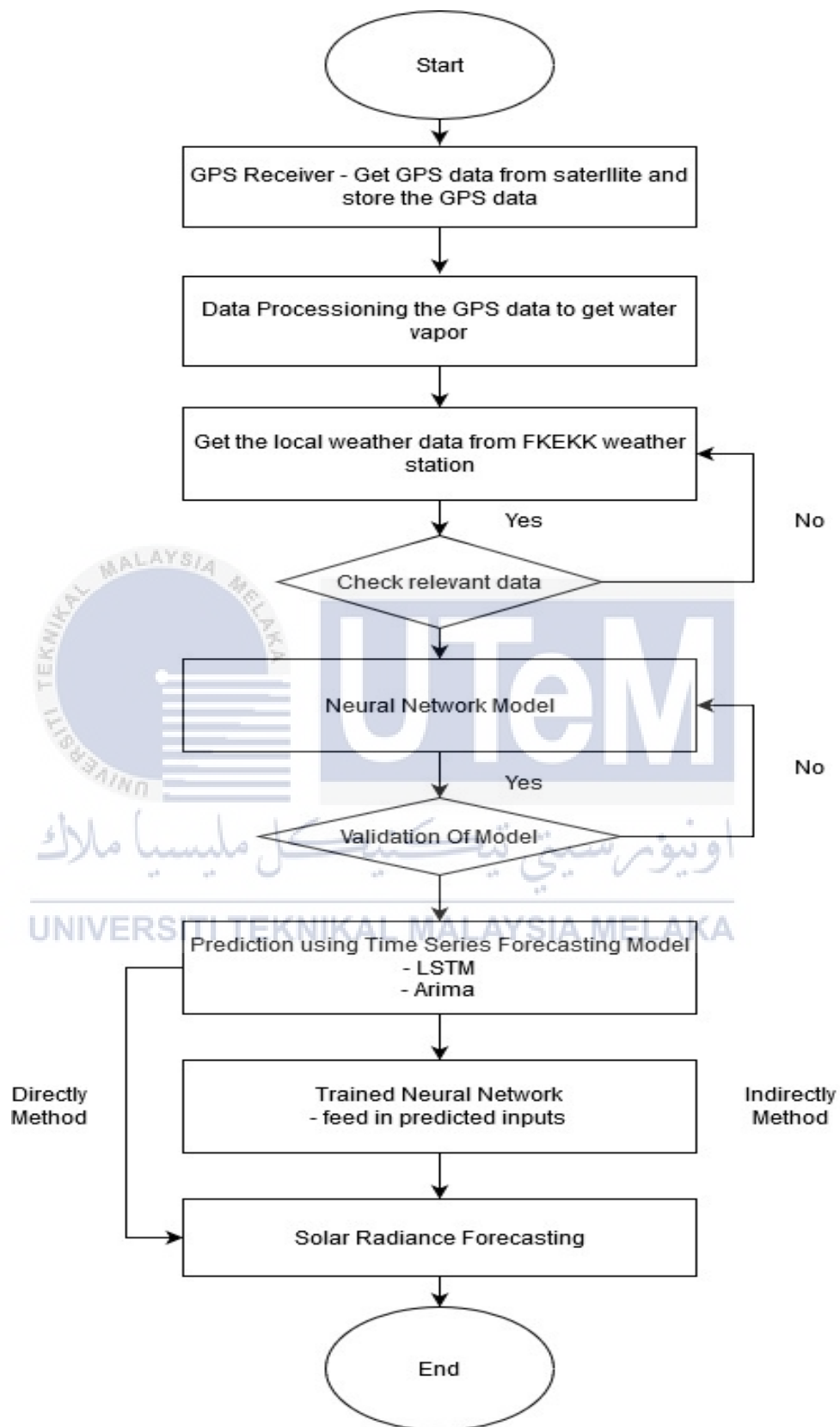


Figure 3.1 Flowchart of research project

The operation of interaction between solar irradiance and water vapor was start by receiving and storing GPS data by the GPS Receiver. The GPS data was proceeded to estimate the water vapor content such as PWV, ZWD and ZTD parameters. At the same time, local weather data like temperature, solar irradiance, relative humidity, and pressure is also were measured and stored the data. To gain comprehensive result, the obtained water vapor values and local weather data was represented in graphical form by using statistical method which is MATLAB software and Neural network model was created. Model accuracy or validity was checked to forecast. After that, data that used to create model was predicted using time series forecasting model. Finally, the solar irradiance was forecasted by indirectly and directly. Indirectly method, which feed in the predicted data into trained neural network to forecast solar irradiance. For directly, that solar irradiance was forecasted by using time series forecasting model.

3.2 Literature Review

Literature review will proceed the project to the further steps after identifying the objectives, problem statement and scope of work. Literature reviews help to understand fully about every element of this project. Focused to this project, mainly for identifying the relationship between solar irradiance and atmospheric water vapor using ground-based GPS. Hence, background studies help to understanding about how GPS performs to identifying and analyzing data needed to estimate precipitable water vapor. Temperature, pressure, humidity, pressure, zenith total delay (ZTD), zenith wet delay (ZWD) and zenith dry delay (ZHD) are the main parameters that help to determine precipitable water vapor. Moreover, literature review will guide to get GPS

data by the GPS techniques to determine zenith delay, estimate water vapor and forecast solar irradiance.

3.3 Data processing of Rinex GPS Data and Weather Data

3.3.1 Research on Atmospheric Water Vapor and Ground Based GPS Receiver

Beginning of this project, some research is carried out on water vapor and GPS section. In this project, most of the research documents and reference articles are getting from outsides. By referring the documents and articles, some of the researchers use same method to be identifying and analyzing the atmospheric water vapor. To start this project by research on analyzation water vapor, generally can understand the working principle of ground-based GPS, important role of water vapor to global climate and effect to the solar power generate.

For this project, firstly focus on determine some parameters on GPS signal delay and estimating the water vapor. In UTEm, local weather data and GPS observation data were collected.

3.3.2 Process flowchart for RINEX GPS data

Firstly, the GPS data was collected from FKEKK weather station in form of RINEX file. RTKLIB software was used for processing the data from the RINEX file with specification of double difference technique. By using the technique, the Zenith Total Delay was estimated and stored in a file format. Figure below was showing the flowchart to estimate ZTD.

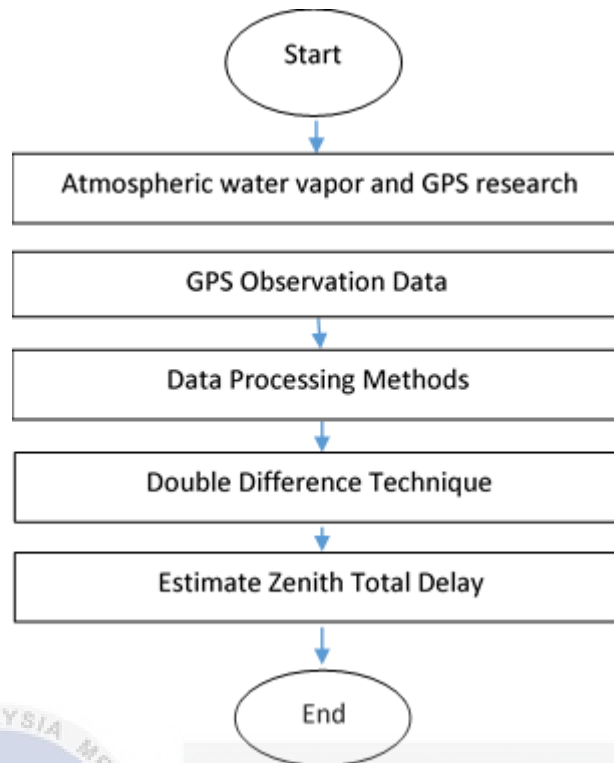


Figure 3.2 Flowchart of estimate ZTD

3.3.3 RINEX GPS Data

This project begins with doing research on atmospheric water vapor and the ground-based GPS. For the 1st step, GPS data is collected from the GPS receiver in form of raw file and format as RINEX file. This GPS file is compressed in one folder. It contains three types of files which is observation file (. O), navigation file (. N) and meteorological (. G) file. To estimate the ZTD value, only observation and navigation file needed.

3.3.3.1 Observation file

The following figures show the sample observation file and navigation file of November 11, 2019. In RINEX observation file observed by using RTKLIB software. This RINEX GPS data is get from UTeM. By using RTKLIB software, will process the data to get furthermore parameters. In this observation file, contains code/ pseudo range, phase error, ionosphere phase delay, receiver channel numbers and antenna position. Moreover, this observation data will record in every 30 seconds. This RINEX file in version 2.11 and the observer for this file is Micheal W. Walsh and agency of SERIS. For the first observation recorded on 11/11/2019 pure GPS file.

```

2.11 MALAYSIA OBSERVATION DATA M (MIXED) RINEX VERSION / TYPE
NetR9 5.15 Receiver Operator 20191101 000001 UTC PGM / RUN BY / DATE
UTEM MARKER NAME
0002 MARKER NUMBER
Micheal W. Walsh SERIS OBSERVER / AGENCY
5702R51131 Trimble NetR9 5.15 REC # / TYPE / VERS
TRM57971.00 NONE ANT # / TYPE
-1359662.8397 6226312.4223 255808.1633 APPROX POSITION XYZ
0.0000 0.0000 0.0000 ANTENNA: DELTA H/E/N
1 1 WAVELENGTH FACT L1/2
8 CI L1 S1 P1 C2 L2 S2 P2 # / TYPES OF OBSERV
30.000 INTERVAL
2019 11 11 0 0 0.00000000 GPS TIME OF FIRST OBS
GLONASS C/A & P. PHASE MATCH: PHASE SHIFTS REMOVED COMMENT
END OF HEADER
19 11 1 0 0 0.00000000 0 21G 2R14R17G19G13G 6G28R 2R16G 5G 9R15 0.000000000
G15R24G30G 4G24R 3G17G12R18
23729405.617 7 124698744.199 7 42.100
97167930.18247 27.500 23729404.31647
22459211.320 9 119720174.073 9 48.100 22459210.672 8 22459208.988 7
93115732.887 7 41.900
22738973.188 7 121681281.541 7 43.000 22738972.723 7 22738969.648 7
94640783.420 7 41.100
21930397.328 8 115245000.965 8 46.800
89801295.18749 37.300 21930395.89849
23298610.609 6 122435155.584 6 39.400
95403896.18946 22.600 23298610.48846
23988222.008 6 126058834.995 6 39.100
98227755.70747 27.600 23988225.49247
22475637.320 7 118110309.193 7 41.700
92033981.85747 28.300 22475636.71147

```

Figure 3.3 RINEX observation file of November 11, 2019

3.3.3.2 Navigation File

Other than that, navigation file (. N) also, important to estimate water vapor. This file contains the satellite positions and satellite numbers that transmit signal. This navigation message provides all the basic information to active the ability of positioning service. This navigation data includes the Ephemeris parameter, which used to calculate the position of each satellite with high accuracy. Figure 3.2 shows navigation file on November 11, 2019.

```

2.11          N: GPS NAV DATA          RINEX VERSION / TYPE
NetR9 5.15    Receiver Operator          PGM / RUN BY / DATE
.1211D-07   -.7451D-08   -.1192D-06   .5960D-07      ION ALPHA
.9626D+05   -.3277D+05   -.1966D+06   .1966D+06      ION BETA
.000000000000D+00 .266453525910D-14   61440   2078 DELTA-UTC: A0,A1,T,W
18          LEAP SECONDS
          END OF HEADER
2 19 11 1 0 0 0.0 -.337534118444D-03 -.773070496507D-11 .000000000000D+00
.830000000000D+02 -.101156250000D+03 .413017203824D-08 .289183803600D+01
-.510551035404D-05 .193008601200D-01 .903010368347D-05 .515357162666D+04
.432000000000D+06 .193715095520D-06 .474671234179D+00 -.117346644402D-06
.956877388847D+00 .199062500000D+03 -.170526392526D+01 -.751495588522D-08
.871464871438D-10 .100000000000D+01 .207700000000D+04 .000000000000D+00
.240000000000D+01 .000000000000D+00 -.176951289177D-07 .830000000000D+02
.424818000000D+06 .400000000000D+01
19 19 11 1 0 0 0.0 -.245993956923D-03 .522959453519D-11 .000000000000D+00
.500000000000D+02 -.696875000000D+02 .417053086221D-08 .116493992081D+01
-.392645597458D-05 .926696474198D-02 .374577939510D-05 .515365291405D+04
.432000000000D+06 -.372529029846D-07 -.405149676110D+00 -.126659870148D-06
.982409256283D+00 .325156250000D+03 .152274410973D+01 -.816962601200D-08
-.307155651409D-09 .100000000000D+01 .207700000000D+04 .000000000000D+00
.240000000000D+01 .000000000000D+00 -.153688224812D-07 .500000000000D+02
.424818000000D+06 .400000000000D+01
13 19 11 1 0 0 0.0 -.355597585440D-04 .227373675443D-11 .000000000000D+00
.840000000000D+02 -.144375000000D+02 .464376486002D-08 -.237654158431D+01
-.875443220139D-06 .399330747314D-02 .722892582417D-05 .515366081810D+04
.432000000000D+06 .726431608200D-07 .275986727589D+01 .745058059692D-08
.968089225191D+00 .242500000000D+03 .119244254664D+01 -.793033033009D-08
-.251081887140D-09 .100000000000D+01 .207700000000D+04 .000000000000D+00
.240000000000D+01 .000000000000D+00 -.111758708954D-07 .840000000000D+02
.424926000000D+06 .400000000000D+01

```

Figure 3.4 Navigation file of November 11, 2019

3.3.4 Data Processing of RINEX file

GPS signal propagation used two type of frequencies which is L1 and L2. The signal frequencies will be recorded in GPS observation data (. O) file. By using these

two frequencies, the delay of the signal will be observed. The usage of two difference frequencies is called double difference technique. RTKLIB is one of the software process the GPS data by using double difference technique.

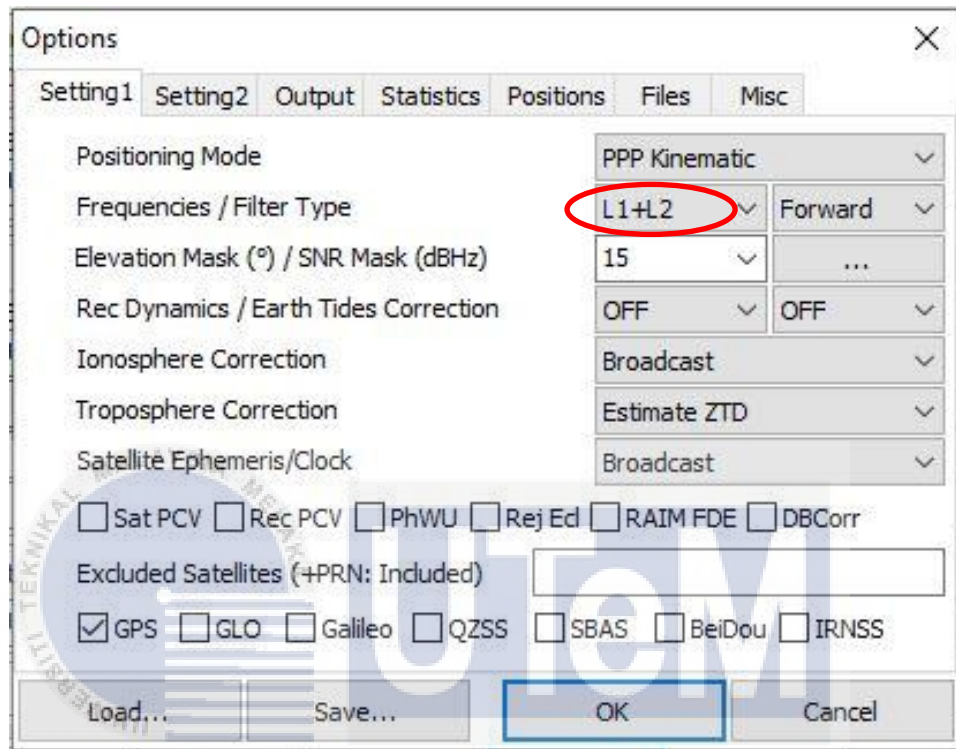


Figure 3.5 RTKLIB software option window for double difference technique

3.3.5 Data preparation to Estimate Water Vapor

3.3.5.1 RINEX GPS file process

The RINEX data is the one of the main parameters that help to estimate water vapor day by day. By using the RINEX GPS data, will get some parameters which are longitude, latitude, height of the GPS Satellite and Zenith Total Delay. The following figures are the steps to determine these parameters using RINEX GPS data.

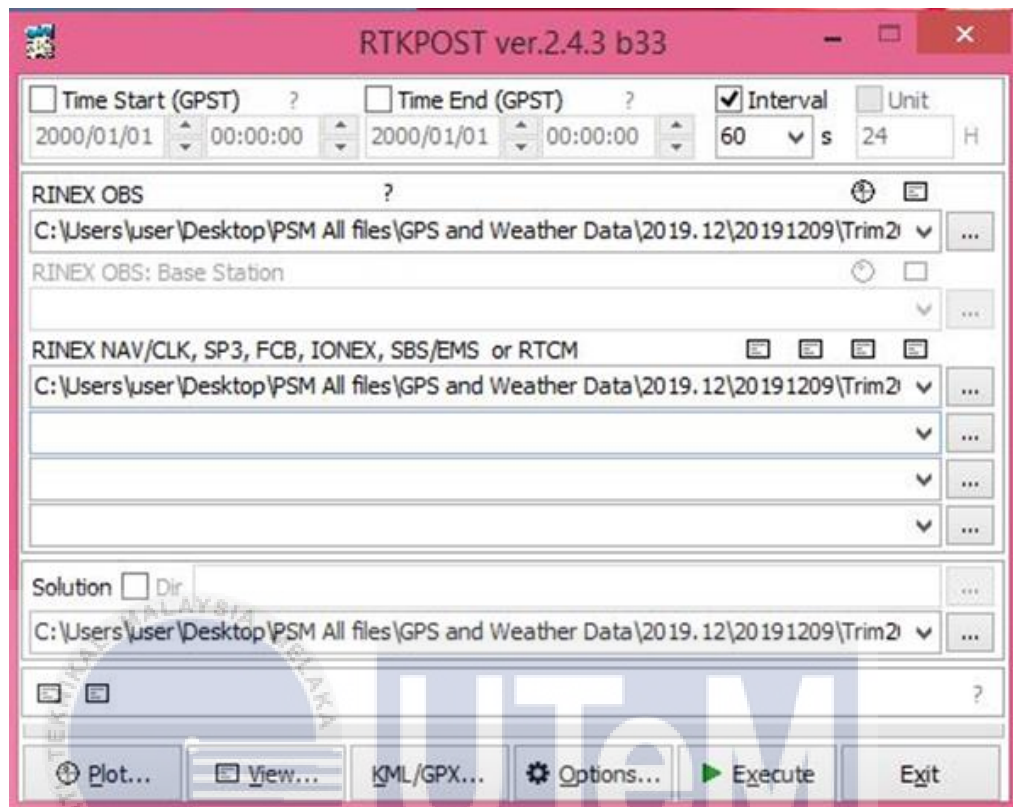


Figure 3.6 Main interface of RTKLIB software

Firstly, using RTKLIB software, include observation (.obs) and navigation (.nav) files of same date recorded. Secondly, verify the process using dual difference frequencies (L1-L2) and set interval to record data for every 60 secs. After the verification, execute the process and come out with post file and stat file. Post file contains longitude, latitude, height, slant delay from the GPS data and the Stat file contain the tropospheric delays which declare as zenith total delay. This process will repeat again to record each, and every day data collected from FKEKK Lab.

```

% program : RTKPOST ver. 2.4.3 b33
% inp file : C:\Users\user\Desktop\GPS and Weather Data\2019.10\20191025\Trim201910250000C.190
% inp file : C:\Users\user\Desktop\GPS and Weather Data\2019.10\20191025\Trim201910250000C.19N
% obs start : 10/25/2019 00:00.0 GPST (week2076 432000.0s)
% obs end : 10/25/2019 59:00.0 GPST (week2076 518340.0s)
% pos mode : ppp-kinematic
% solution : forward
% elev mask : 15 deg
% dynamics : off
% tidecorr : off
% tropo opt : est ztd
% ephemeris : broadcast
% antennal : ( 0 0 0.0000)
%
% "(lat/lon/height=WGS84/ellipsoidal,Q=1:fix,2:float,3:sbas,4:dgps,5:single,6:ppp,ns=#" of satellites)
% GPST latitude(deg) longitude(deg) height(m) Q ns sdn(m) sde(m) sdu(m) sdne(m) sdeu(m) sdu(m) age(s)
10/25/2019 0:00:00 2.314033601 102.3185039 59.3033 6 8 1.8216 1.9658 6.4791 -0.7097 0.2016 2.2946 0
10/25/2019 0:01:00 2.314033438 102.3185036 59.1749 6 8 1.5703 1.6931 5.5632 -0.5904 0.3945 1.9541 0
10/25/2019 0:02:00 2.314034103 102.3185039 59.5161 6 8 1.4729 1.5878 5.1931 -0.5335 0.4885 1.8076 0
10/25/2019 0:03:00 2.314034309 102.3185041 59.7195 6 8 1.4192 1.5299 4.9778 -0.4943 0.5581 1.7161 0
10/25/2019 0:04:00 2.314034114 102.3185043 59.6171 6 8 1.3842 1.4922 4.8287 -0.4623 0.615 1.6481 0
10/25/2019 0:05:00 2.314034006 102.3185041 59.5703 6 8 1.359 1.465 4.7145 -0.4341 0.6637 1.5925 0
10/25/2019 0:06:00 2.314034059 102.318504 59.5417 6 8 1.3397 1.444 4.6215 -0.4078 0.7062 1.5442 0
10/25/2019 0:07:00 2.314034169 102.3185038 59.8202 6 8 1.3243 1.4271 4.5423 -0.3826 0.744 1.5007 0
10/25/2019 0:08:00 2.314034333 102.3185037 59.8996 6 8 1.3115 1.4129 4.473 -0.3578 0.778 1.4605 0
10/25/2019 0:09:00 2.314034881 102.3185037 60.1676 6 8 1.3008 1.4008 4.4111 -0.3331 0.8088 1.4229 0
10/25/2019 0:10:00 2.314035048 102.3185035 60.1869 6 8 1.2917 1.3903 4.3551 -0.308 0.8368 1.3872 0
10/25/2019 0:11:00 2.314035345 102.3185035 60.2993 6 8 1.2838 1.3809 4.3038 -0.2821 0.8626 1.353 0
10/25/2019 0:12:00 2.314035242 102.3185039 60.3694 6 8 1.277 1.3725 4.2565 -0.2549 0.8863 1.3202 0
10/25/2019 0:13:00 2.314035305 102.3185033 60.3636 6 8 1.271 1.3649 4.2127 -0.2259 0.9082 1.2884 0
10/25/2019 0:14:00 2.314035604 102.3185032 60.4711 6 8 1.2659 1.3579 4.1719 -0.1938 0.9286 1.2576 0
    
```

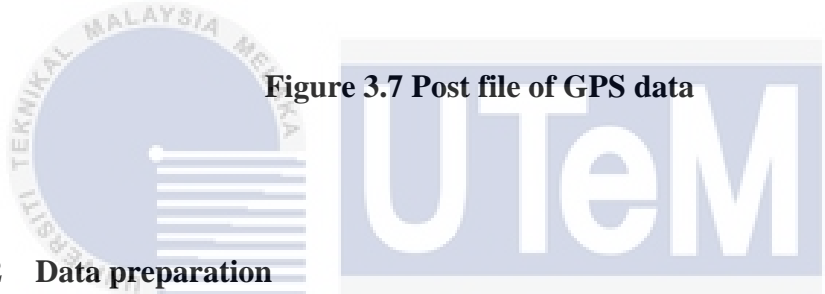


Figure 3.7 Post file of GPS data

3.3.5.2 Data preparation

After the post file generated, analyzed the relevant data into Excel file. By creating the Excel file, it will help to calculate the data of water vapor (PWV). The data of GPS has been collected for 5 months and recorded into several sheets. Relevant weather data (temperature, humidity, pressure, and solar irradiance) also collected and compiled into every sheets. Each day, the data has been collected from 7am to 7pm. The analyzed data was obtained from each minute of the duration stated earlier.

	A	B	C	D	E	F	G	H	I	J
1	Date	Time	latitud	longitu	height	Solar R	Humidi	Surface	Pressur	Mean T
422	11/26/2019	7:00 AM	2.314068	102.3185	63.2428	7	91	23.4	1003.5	87.048
423	11/26/2019	7:01 AM	2.314068	102.3185	63.1927	9	91	23.5	1003.5	87.12
424	11/26/2019	7:02 AM	2.314068	102.3185	63.3498	11	91	23.5	1003.5	87.12
425	11/26/2019	7:03 AM	2.314069	102.3185	63.3821	12	91	23.5	1003.6	87.12
426	11/26/2019	7:04 AM	2.314068	102.3185	63.338	12	91	23.5	1003.5	87.12
427	11/26/2019	7:05 AM	2.314068	102.3185	63.3022	15	91	23.5	1003.6	87.12
428	11/26/2019	7:06 AM	2.314068	102.3185	63.3743	16	91	23.6	1003.6	87.192
429	11/26/2019	7:07 AM	2.314068	102.3185	63.4157	16	91	23.6	1003.6	87.192
430	11/26/2019	7:08 AM	2.314067	102.3185	63.426	18	91	23.6	1003.5	87.192
431	11/26/2019	7:09 AM	2.314066	102.3185	63.4453	18	91	23.6	1003.6	87.192
432	11/26/2019	7:10 AM	2.314066	102.3185	63.653	21	91	23.6	1003.6	87.192
433	11/26/2019	7:11 AM	2.314066	102.3185	63.5204	24	91	23.6	1003.5	87.192
434	11/26/2019	7:12 AM	2.314066	102.3185	63.4718	26	91	23.6	1003.6	87.192
435	11/26/2019	7:13 AM	2.314066	102.3185	63.4875	28	91	23.6	1003.5	87.192
436	11/26/2019	7:14 AM	2.314065	102.3185	63.4442	32	91	23.6	1003.6	87.192
437	11/26/2019	7:15 AM	2.314065	102.3185	63.4344	35	91	23.6	1003.6	87.192
438	11/26/2019	7:16 AM	2.314065	102.3185	63.3646	37	91	23.7	1003.6	87.264
439	11/26/2019	7:17 AM	2.314064	102.3185	63.486	41	90	23.7	1003.6	87.264
440	11/26/2019	7:18 AM	2.314064	102.3185	63.497	42	90	23.7	1003.6	87.264
441	11/26/2019	7:19 AM	2.314064	102.3185	63.5335	40	90	23.7	1003.6	87.264
442	11/26/2019	7:20 AM	2.314064	102.3185	63.5177	39	90	23.8	1003.6	87.336
443	11/26/2019	7:21 AM	2.314063	102.3185	63.415	39	90	23.8	1003.6	87.336

Figure 3.8 Data collection in Excel file

3.3.6 Estimate Water Vapor

3.3.6.1 Zenith Total Delay

ZTD is a significant parameter for estimating the PWV, as discussed in literature review. We will evaluate the ZTD by using the RTKLIB code. To generate ZTD, which is a raw data file for GPS observation and GPS navigation data file, two files are needed. Both files have some parameters that help calculate the total zenith delay such as duration, pseudo range, period, angle, and intensity of the satellite.

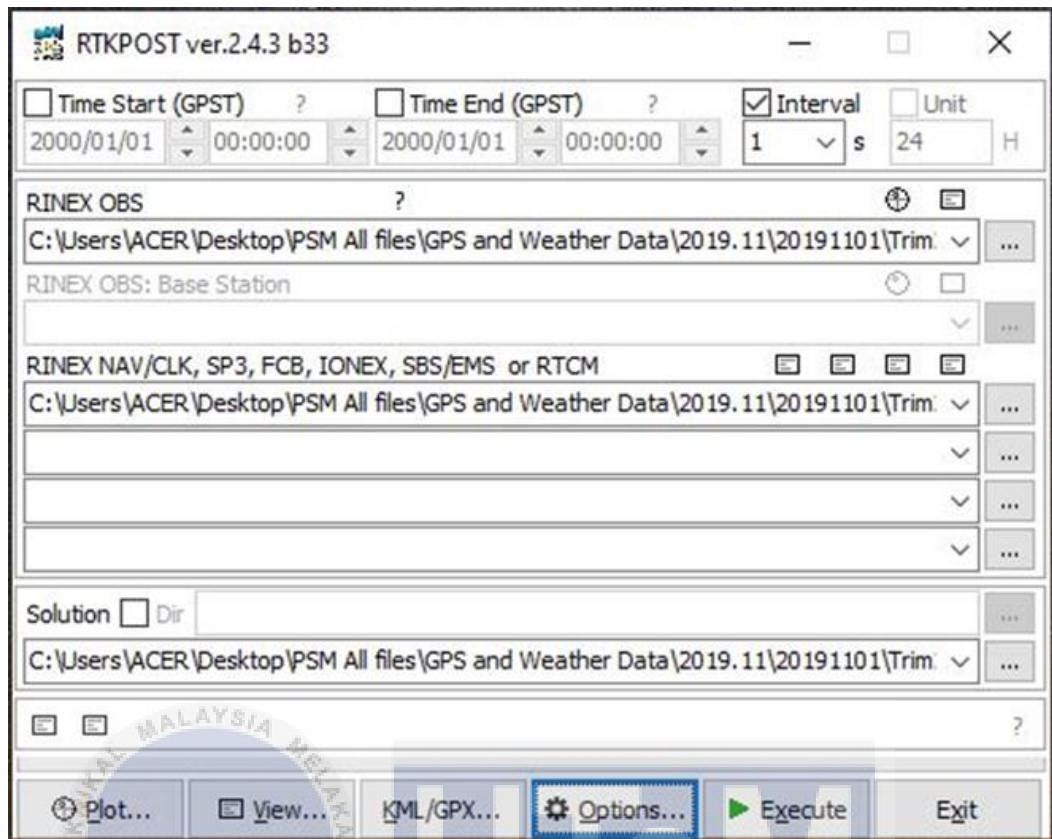


Figure 3.9 Main window of RTKPOST

RTKPOST is one of the applications that generate the ZTD values. By add and execute observation file (. O) and navigation file (. N) in this application, the STAT file generates automatically. From the STAT file, can observe the tropospheric delay (\$TROP) which is ZTD. Moreover, from the STAT file, can determine the coordinate position of satellite.

```

C:\Users\ACER\Desktop\PSM All files\GPS and Weather Data\2019.11\20191101\Trim201911010000C.pos.stat
Find
$POS,2077,432000.000,6,-1359661.0018,6226302.4445,255806.9299,2.3553,5.9605,1.8781
$CLK,2077,432000.000,6,1,-33.162,-1.553,12.707,200.138
$TROP,2077,432000.000,6,1,2.5595,0.1200
$POS,2077,432060.000,6,-1359660.9689,6226302.4513,255806.8948,2.0419,5.1527,1.6262
$CLK,2077,432060.000,6,1,-33.144,-1.362,11.026,200.138
$TROP,2077,432060.000,6,1,2.5619,0.1199
$POS,2077,432120.000,6,-1359660.9499,6226302.3694,255806.8830,1.9270,4.8446,1.5323
$CLK,2077,432120.000,6,1,-33.195,-1.313,10.397,200.138
$TROP,2077,432120.000,6,1,2.5642,0.1199
$POS,2077,432180.000,6,-1359660.8298,6226302.1586,255806.8880,1.8682,4.6779,1.4831
$CLK,2077,432180.000,6,1,-33.573,-1.423,10.066,200.138
$TROP,2077,432180.000,6,1,2.5666,0.1199
$POS,2077,432240.000,6,-1359660.7957,6226302.1236,255806.8495,1.8331,4.5714,1.4530
$CLK,2077,432240.000,6,1,-33.679,-1.372,9.862,200.138
$TROP,2077,432240.000,6,1,2.5690,0.1198
$POS,2077,432300.000,6,-1359660.8046,6226302.2575,255806.8555,1.8105,4.4965,1.4328
$CLK,2077,432300.000,6,1,-33.483,-0.937,9.724,200.138
$TROP,2077,432300.000,6,1,2.5713,0.1198
$POS,2077,432360.000,6,-1359660.8019,6226302.2692,255806.8436,1.7954,4.4404,1.4187
$CLK,2077,432360.000,6,1,-33.357,-0.689,9.627,200.138
$TROP,2077,432360.000,6,1,2.5736,0.1197
$POS,2077,432420.000,6,-1359660.7798,6226302.4161,255806.8508,1.7850,4.3966,1.4083
$CLK,2077,432420.000,6,1,-33.192,-0.345,9.555,200.138
$TROP,2077,432420.000,6,1,2.5760,0.1197
$POS,2077,432480.000,6,-1359660.7703,6226302.4663,255806.8802,1.7780,4.3614,1.4007
$CLK,2077,432480.000,6,1,-32.916,0.006,9.502,200.138
$TROP,2077,432480.000,6,1,2.5782,0.1197

```

Figure 3.10 STAT file generated from RTKPOST with ZTD

3.3.6.2 Zenith Hydrostatic Delay (Dry Delay)

After observed ZTD value from STAT file, ZHD need to be calculated. Some parameters from the equation ZHD taken from the local weather data that collected from FKEKK. By using the equation below, the ZHD can be calculated.

$$ZHD = (2.779 \pm 0.00024) \frac{P_s}{f(\lambda H)} \quad (3.1)$$

Where, the $f(\lambda H)$ represents function of latitude (λ) and ellipsoid height (H) which get from the generated file in RTKPOST. P_s represents atmospheric pressure which can get from the local weather data.

3.3.6.3 Zenith Wet Delay

From the observation above, using ZTD and ZHD values can determine the ZWD value using equation that get from overall research summarized. The equation for ZWD is given below:

$$ZWD = ZTD - ZHD \quad (3.2)$$

By refer to the equation, about to extract the value of ZWD by subtracting ZHD observed from atmospheric pressure with GPS measured ZTD.

3.3.6.4 Precipitable Water Vapor

After obtained the values that has been discussed earlier, the precipitable water vapor can be calculated. The following equation used to calculate precipitable water vapor.

$$PWV = k \cdot ZWD \quad (3.3)$$

Where the ZWD value calculated from the previous steps. k represents the global climatology which can estimate from specific equations using atmospheric parameters. The equation mentioned below:

$$k = \frac{10^8}{[Rv \left(\frac{k^3}{Tm + k'_2} \right)]} \quad (3.4)$$

Where $Rv = (461.525 \pm 0.013) \text{ J / (kmol K)}$ is gas constant for water vapor, $k^3 = (373900 \pm 1200) \text{ K}^2 \text{ Pa}^{-1}$ and $k'_2 = (0.221 \pm 0.023) \text{ KPa}^{-1}$ which represents

refractivity coefficient T_m is mean temperature in atmosphere. T_m can be estimated from surface temperature (T_s) measurement using formula given below:

$$T_m = 70.2 + 0.72 (T_s) \quad (3.5)$$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Date	Time	latitud	longitu	height	Solar R	Humid	Surface	Pressur	Mean T	ZTD	ZHD	ZWD	k (glob	IWV
422	11/26/2019	7:00 AM	2.314068	102.3185	63.2428	7	91	23.4	1003.5	87.048	2.4497	0.023253	2.426447	50.24498	121.9168
423	11/26/2019	7:01 AM	2.314068	102.3185	63.1927	9	91	23.5	1003.5	87.12	2.4481	0.023252	2.424848	50.28637	121.9368
424	11/26/2019	7:02 AM	2.314068	102.3185	63.3498	11	91	23.5	1003.5	87.12	2.4465	0.023253	2.423247	50.28637	121.8563
425	11/26/2019	7:03 AM	2.314069	102.3185	63.3821	12	91	23.5	1003.6	87.12	2.445	0.023256	2.421744	50.28637	121.7807
426	11/26/2019	7:04 AM	2.314068	102.3185	63.338	12	91	23.5	1003.5	87.12	2.4435	0.023253	2.420247	50.28637	121.7054
427	11/26/2019	7:05 AM	2.314068	102.3185	63.3022	15	91	23.5	1003.6	87.12	2.4419	0.023255	2.418645	50.28637	121.6249
428	11/26/2019	7:06 AM	2.314068	102.3185	63.3743	16	91	23.6	1003.6	87.192	2.4404	0.023256	2.417144	50.32777	121.6495
429	11/26/2019	7:07 AM	2.314068	102.3185	63.4157	16	91	23.6	1003.6	87.192	2.4389	0.023256	2.415644	50.32777	121.574
430	11/26/2019	7:08 AM	2.314067	102.3185	63.426	18	91	23.6	1003.5	87.192	2.4374	0.023254	2.414146	50.32777	121.4986
431	11/26/2019	7:09 AM	2.314066	102.3185	63.4453	18	91	23.6	1003.6	87.192	2.436	0.023256	2.412744	50.32777	121.428
432	11/26/2019	7:10 AM	2.314066	102.3185	63.653	21	91	23.6	1003.6	87.192	2.4345	0.023258	2.411242	50.32777	121.3524
433	11/26/2019	7:11 AM	2.314066	102.3185	63.5204	24	91	23.6	1003.5	87.192	2.4331	0.023255	2.409845	50.32777	121.2821
434	11/26/2019	7:12 AM	2.314066	102.3185	63.4718	26	91	23.6	1003.6	87.192	2.4317	0.023257	2.408443	50.32777	121.2116
435	11/26/2019	7:13 AM	2.314066	102.3185	63.4875	28	91	23.6	1003.5	87.192	2.4303	0.023254	2.407046	50.32777	121.1412
436	11/26/2019	7:14 AM	2.314065	102.3185	63.4442	32	91	23.6	1003.6	87.192	2.429	0.023256	2.405744	50.32777	121.0757
437	11/26/2019	7:15 AM	2.314065	102.3185	63.4344	35	91	23.6	1003.6	87.192	2.4277	0.023256	2.404444	50.32777	121.0103
438	11/26/2019	7:16 AM	2.314065	102.3185	63.3646	37	91	23.7	1003.6	87.264	2.4264	0.023256	2.403144	50.36916	121.0444
439	11/26/2019	7:17 AM	2.314064	102.3185	63.486	41	90	23.7	1003.6	87.264	2.4251	0.023257	2.401843	50.36916	120.9788
440	11/26/2019	7:18 AM	2.314064	102.3185	63.497	42	90	23.7	1003.6	87.264	2.4238	0.023257	2.400543	50.36916	120.9133
441	11/26/2019	7:19 AM	2.314064	102.3185	63.5335	40	90	23.7	1003.6	87.264	2.4226	0.023257	2.399343	50.36916	120.8529
442	11/26/2019	7:20 AM	2.314064	102.3185	63.5177	39	90	23.8	1003.6	87.336	2.4214	0.023257	2.398143	50.41056	120.8917
443	11/26/2019	7:21 AM	2.314063	102.3185	63.415	39	90	23.8	1003.6	87.336	2.4202	0.023256	2.396944	50.41056	120.8313

Figure 3.11 Data For Estimation Of Water Vapor (PWV)

After precipitable water vapor calculated, two excel files prepared (input1 and output1). The input1 file contains of time, humidity, pressure, temperature, and water vapor. The output1 file contains solar irradiance. All the data has been collected and prepared for the data analysis process in MATLAB software.

3.4 MATLAB Platform for Modelling

After the data has been prepared, was imported into MATLAB software for data preprocess. The data was imported to MATLAB workspace and converted into numeric matrix value. The data was preprocessed by removing the zero values to

reduce the error of the Artificial Neural Network (ANN) model. Moreover, the data was transposed the variable from row to column.

3.4.1 Neural Network Fitting

Neural network fitting known as feedforward network. Feedforward networks are made of a combination of layers. The first layer has a Network Input relation. Through layer subsequently has a connection from the previous layer. The final layer generates output for the network. Feedforward network with sufficient hidden layers and enough hidden layers of neurons will suit any finite input-output mapping problem. The feedforward network coding in MATLAB was explained below.

```
x = input1';
t = output1';
```

The x (input) value is representing the input1 and t (target) value is representing output1 from the MATLAB workspace which are imported from Excel file.

```
trainFcn = 'trainlm'; % Levenberg-Marquardt backpropagation.

% Create a Fitting Network
hiddenLayerSize = 1000;
net = fitnet(hiddenLayerSize,trainFcn);
```

For training, the network was used the default Levenberg-Marquardt algorithm (trainlm). The trainlm is regularly the quickest backpropagation calculation in the tool compartment and is strongly suggested as a first-decision managed calculation, although it requires more memory than different calculations. The network hidden layer size was set to 1000 value.

```
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
```

With these settings, the input vectors and target vectors were irregularly separated, with 70% utilized for training, 15% for validation and 15% for testing.

```
% Train the Network
[net,tr] = train(net,x,t);
```

The network was trained with the input and target data by using Levenberg-Marquardt algorithm.

```
Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)
```

The code above explained that after network trained, the network outputs, error and overall performance was calculated.

```
% View the Network
view(net)
```

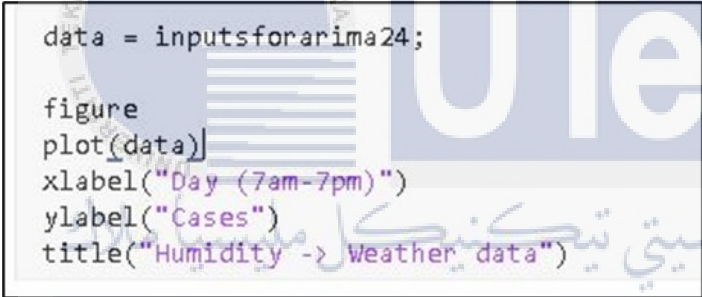
The code above explained after trained network, it was view in network diagram which are include with layers.

3.4.2 Prediction using LSTM and ARIMA model

The input file that has been used in neural network fitting was imported to workspace. The input variables which are humidity, temperature, pressure, and water vapor were separated into several files in form of double matrices. By using LSTM and ARIMA model, the future value has been predicted for 12 hours ahead. The different kind of models the predicted data was generated with varying values.


3.4.2.1 LSTM Model

The Long Short-Term Memory (LSTM) process helps a deep learning on long- and short-term dependencies in time series and sequence data between time stages. The LSTM model code in MATLAB was specified below.



```
data = inputsforarima24;
figure
plot(data)
xlabel("Day (7am-7pm)")
ylabel("Cases")
title("Humidity -> Weather data")
```

The measured humidity data was imported to workspace and declared as data. The humidity data was plotted which are recorded data from 7am-7pm.



```
numTimeStepsTrain = floor(0.9*numel(data));
dataTrain = data(1:numTimeStepsTrain+1);
dataTest = data(numTimeStepsTrain+1:end);
```

Next step, the data was separated into training data and testing data. The training data was separated for 90 percent which identified as dataTrain in code. The remained 10 percent data for testing which can identified as dataTest in code.

```

mu = mean(dataTrain);
sig = std(dataTrain);

dataTrainStandardized = (dataTrain - mu) / sig;

```

Before the LSTM network trained, the training data was standardized by calculate mean and standard deviation of timeseries data.

```

XTrain = dataTrainStandardized(1:end-1);
YTrain = dataTrainStandardized(2:end);

```

The code explained that the input data (XTrain) and output data (YTrain) was declared by separated the standardized data, respectively.

```

numFeatures = 1;
numResponses = 1;
numHiddenUnits = 100;

layers = [ ...
sequenceInputLayer(numFeatures)
lstmLayer(numHiddenUnits)
fullyConnectedLayer(numResponses)
regressionLayer];

```

The LSTM regression network created, and layers was defined as regression layer and specified the input and output layer as 'sequenceInputLayer' and 'fullyConnectedLayer'. LSTM layer was specified to required 100 hidden units.

```

options = trainingOptions('adam', ...
    'MaxEpochs',150, ...
    'GradientThreshold',1, ...
    'InitialLearnRate',0.0001, ...
    'LearnRateSchedule','piecewise', ...
    'LearnRateDropPeriod',125, ...
    'LearnRateDropFactor',0.2, ...
    'Verbose',0, ...
    'Plots','training-progress');

```

The training option was detailed to train the LSTM regression network. The solver was set to 'adam' and trained for 250 epochs. To avoid the gradients from exploding, the threshold value was set to 1 and initial learn rate was specified to 0.005.

```

net = trainNetwork(XTrain,YTrain,layers,options);

```

The LSTM regression network was trained using the input output standardized data, defined layers, and the training options.

```

dataTestStandardized = (dataTest - mu) ./ sig;
XTest = dataTestStandardized(1:end-1);

```

In forecasting and prediction section, the testing data was standardized by using the previous calculation of mean and standard deviation.

```

net = predictAndUpdateState(net,XTrain);
[net,YPred] = predictAndUpdateState(net,YTrain(end));

numTimeStepsTest = numel(XTest);
for i = 2:numTimeStepsTest
    [net,YPred(:,i)] = predictAndUpdateState(net,YPred(:,i-1),'ExecutionEnvironment','cpu');
end

```

The trained LSTM regression network has been used to predict the future data. The XTrain data has been feed into the net. The predict data was stored as 'YPred'

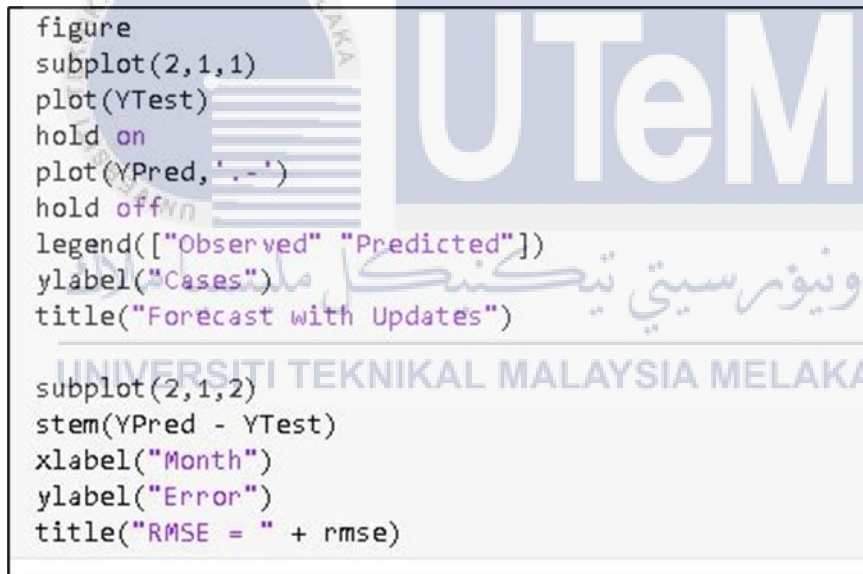
according to number time step of test data. The predict data was computed by CPU using the 'ExecutionEnvironment'

```
YPred = sig*YPred + mu;
```

The predicted data was unstandardized using the parameters calculated.

```
rmse = sqrt(mean((YPred-YTest).^2))
```

The root means square error (rmse) was calculated by compare the predicted data and test data.



```
figure
subplot(2,1,1)
plot(YTest)
hold on
plot(YPred,'.-')
hold off
legend(["Observed" "Predicted"])
ylabel("Cases")
title("Forecast with Updates")

subplot(2,1,2)
stem(YPred - YTest)
xlabel("Month")
ylabel("Error")
title("RMSE = " + rmse)
```

Figure was created with two subplots. In subplot (2,1,1) plotted the measured data and predicted data in one graph. In subplot (2,1,2) plotted difference between the predicted data and measured data and the calculated root mean square error (rmse) was displayed as title.

Moreover, the net was reset and update the data was predicted using the observed data to get more accurate. The same LSTM regression network code was used respectively to predict the temperature, pressure, water vapor and solar irradiance. The predicted data was collected from workspace to excel file.

3.4.2.2 ARIMA Model

ARIMA, short for 'Auto Regressive Integrated Moving Average' is really a class of models that 'clarifies' a given time arrangement dependent on its own past qualities, that is, its own slacks and the slacked figure mistakes, with the goal that condition can be utilized to estimate future qualities. The MATLAB code that used for ARIMA model was explained below.

```
numObs = length(Humidity);
logGDP = (Humidity(1:end-1408));
cpi = Humidity(1:end-1408);
T = length(logGDP); % Effective sample size
frstHzn = T+1:numObs; % Forecast horizon
hoCPI = Humidity(frstHzn); % Holdout sample
```

The humidity data was imported to workspace and included into code. The length of data exhibit measurement was calculated and stored as 'numObs' in workspace. The humidity data was separated to training data as 'logGDP' and set predictor as 'cpi'. The length 'logGDP' data was measured and stored as T variable. The time series forecast of data was declared as 'frstHzn'. The test data for prediction was declared as 'hoCPI'.

```
ToEstMdl = regARIMA('ARLags',1,'MALags',1,'D',1);
```

The regARIMA model was created as 'ToEstMdl' and specified using lags value (1,1,1). The autoregressive polynomial degree ('ARLags') was set as 1, which the vector must be in positive integer. The moving average polynomial degree ('MALags') was set to 1 and the nonseasonal polynomial degree ('D') was set to 1.

```
Reg4Int = [ones(T,1), cpi]\logGDP;
intercept = Reg4Int(1);
```

For a process of combined error, the intercept is not visible. So, correct its value before estimation. Simple linear regression has been used to estimate the intercept.

```
ToEstMdl.Intercept = intercept;
EstMdl = estimate(ToEstMdl,logGDP,'X',cpi,...
'Display','off')
```

The predict value was estimated by using the created model ('ToEstMdl') and the predictor ('cpi').

```
[gdpF,gdpMSE] = forecast(EstMdl,1408,'Y0',logGDP,...
'X0',cpi,'XF',hoCPI);
```

The predicted data was forecasted using the estimate model according to the time series forecast (hoCPI). 'gdpF' contains the predicted value in specific period and 'gdpMSE' contains the predicted errors in specific period.


```

figure
plot((gdpF),'.-')

xlim([0 1500])
ylim([50.0 95.0])
legend('Predicted Humidity')
xlabel('predicted time')
ylabel('solar irradiance')

```

The figure was created and the predicted data in specific period was plotted. The regression ARIMA steps were repeated for predict data of temperature, pressure, water vapor and solar irradiance.

3.4.3 Predict Solar Irradiance using trained Neural Network

All predicted inputs from LSTM and ARIMA model was imported to MATLAB workspace. Using the trained Neural Network and predicted inputs, the solar irradiance was estimated. The flow of predict solar irradiance code was explained below.

```

net.layers
net. efficiency
net.trainParam
net.performFcn
net.gradientFcn
net
view(net)

```

The trained Neural Network was verified by their layers, performance functions, training parameters, and the net was viewed.

```

ypredictsolar_521 = net(lstm521minsdata)

```

The predicted inputs using the LSTM model was imported as 'lstm251minsdata' into MATLAB workspace. The predicted inputs were feed into the trained network.

The predicted solar irradiance was stored in MATLAB workspace as 'ypredictsolar_521'.

```
ypredictsolar = net(regArimapredictedinput)
```

The predicted inputs using regArima was imported as 'regArimapredictedInput' into MATLAB workspace. The predicted inputs were feed into the trained neural network and the solar irradiance was predicted and stored as 'ypredictsolar'.



CHAPTER 4

RESULTS AND DISCUSSION



The result obtained from the previous methodology will be discussed here. All results will be displayed in the form of a graph, table, and image to show the relationship and trend of the collected data. All the data collected are studied and analyzed accordingly. This section findings of the study will be used to review the aims of the project.

4.1 Introduction

The aim of this chapter is to present the outcomes of the project in statistical method. This chapter was started with results from the neural network fitting which known as feedforward network. Secondly, the comparison of predicted input value which are humidity, temperature, pressure, and water vapor was done by the LSTM and ARIMA model trained. Next, the prediction of solar irradiance indirectly by the trained neural network model using the predicted value from the LSTM and ARIMA model was compared in short-term and long-term forecasting. Moreover, the solar irradiance was directly predicted using LSTM and ARIMA model and was compared it. Finally, the directly and indirectly methods of solar irradiance prediction were compared.



4.2 Neural Network Fitting

The network parameters are different each time it initializes a feedforward network and may generate different solutions. If the network is not reliable enough, it should try again to configure the network and the training. Accuracy of the model is more important in this research. Accuracy of model was analyzed by observe overall R-value that clarified the relationship between target and output value of the model trained and the R-value must nearer to 1. Several approaches were observed to improve the model accuracy. The network model was trained from the minimum value of hidden neurons in layers which in default value of 10. While trained the network in minimum value of the hidden neurons, the R-value was observed. In this research, 1000 hidden neurons were implemented for network trained.

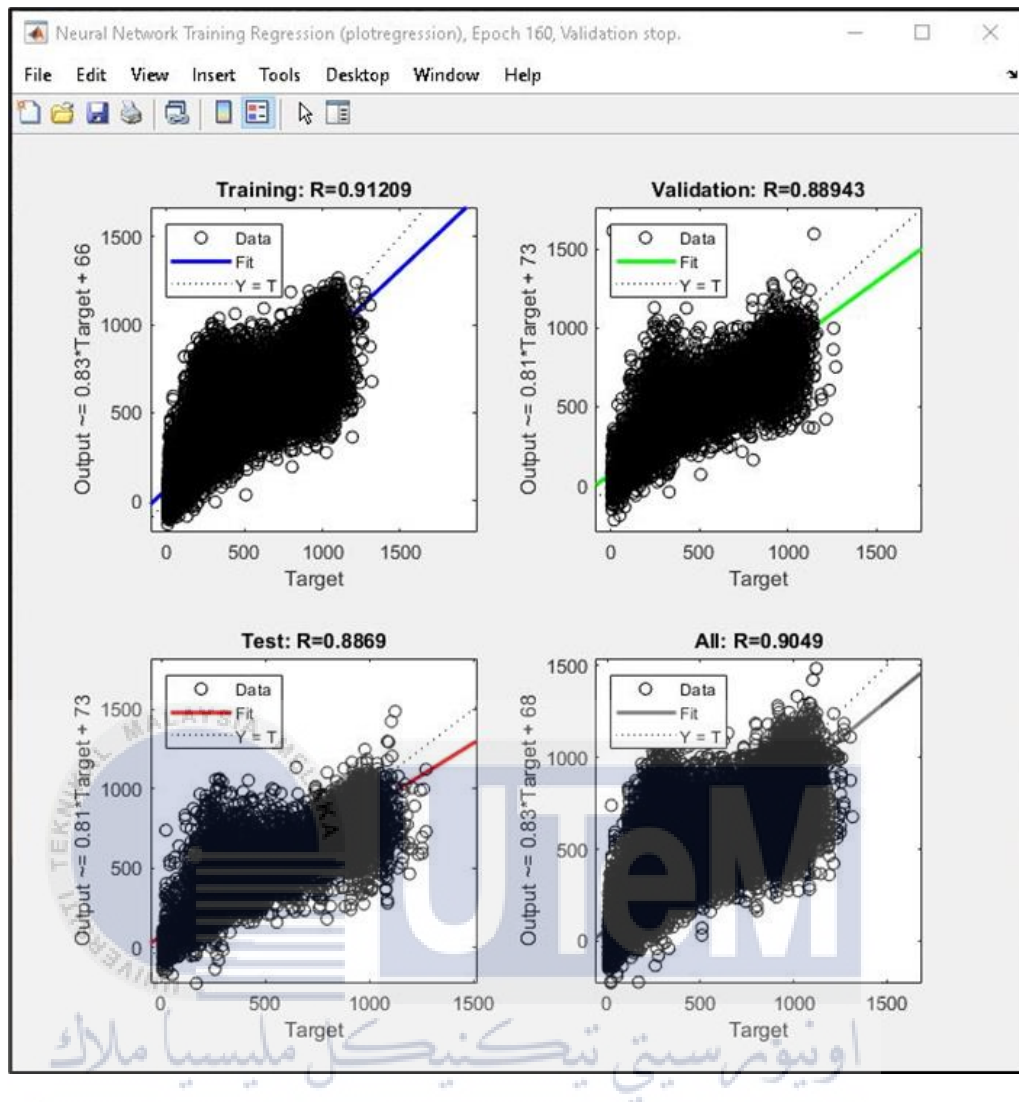


Figure 4.1 Correlation coefficient of Neural Network model

The figure above representing the relationship between target value and output value of the model trained using 1000 hidden neurons. The four plots represent the training, testing, validation, and overall data. The solid line reflects the linear regression line that best matches between outputs and goals. The training and all over result indicated a good fit for the model. The validation and testing results also shown the large value of R which are nearer to 1. In this scatter plot was effective in explaining poor fits to some of the other data points. The scatter plot of testing model, some of the data were not fit to the target.

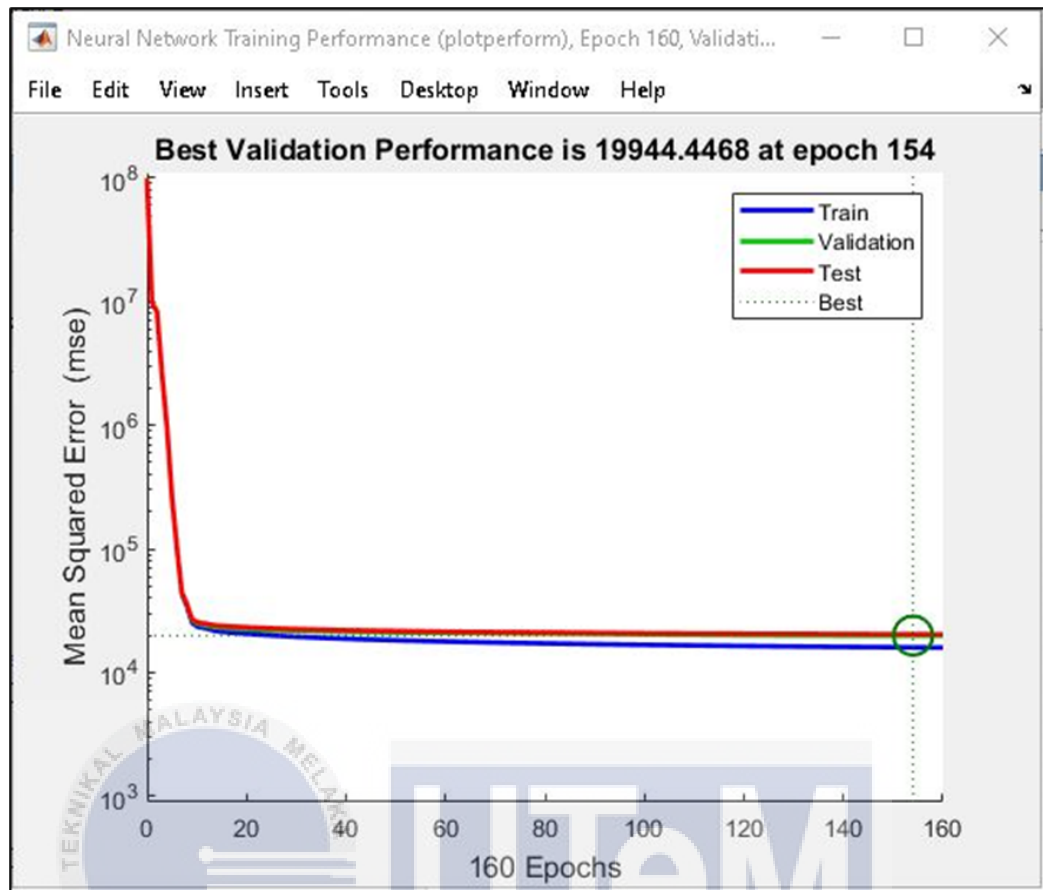


Figure 4.2 Performance of trained network model

The figure above was shown the best validation performance of trained neural network model. The graph was plotted with training errors, validation error and test error occurred while the model was trained. This training stopped since its validation error for six iterations increased which happened at iteration 160. The validation performance was occurred at epochs 154. Moreover, the training error and testing error had similar characteristics.

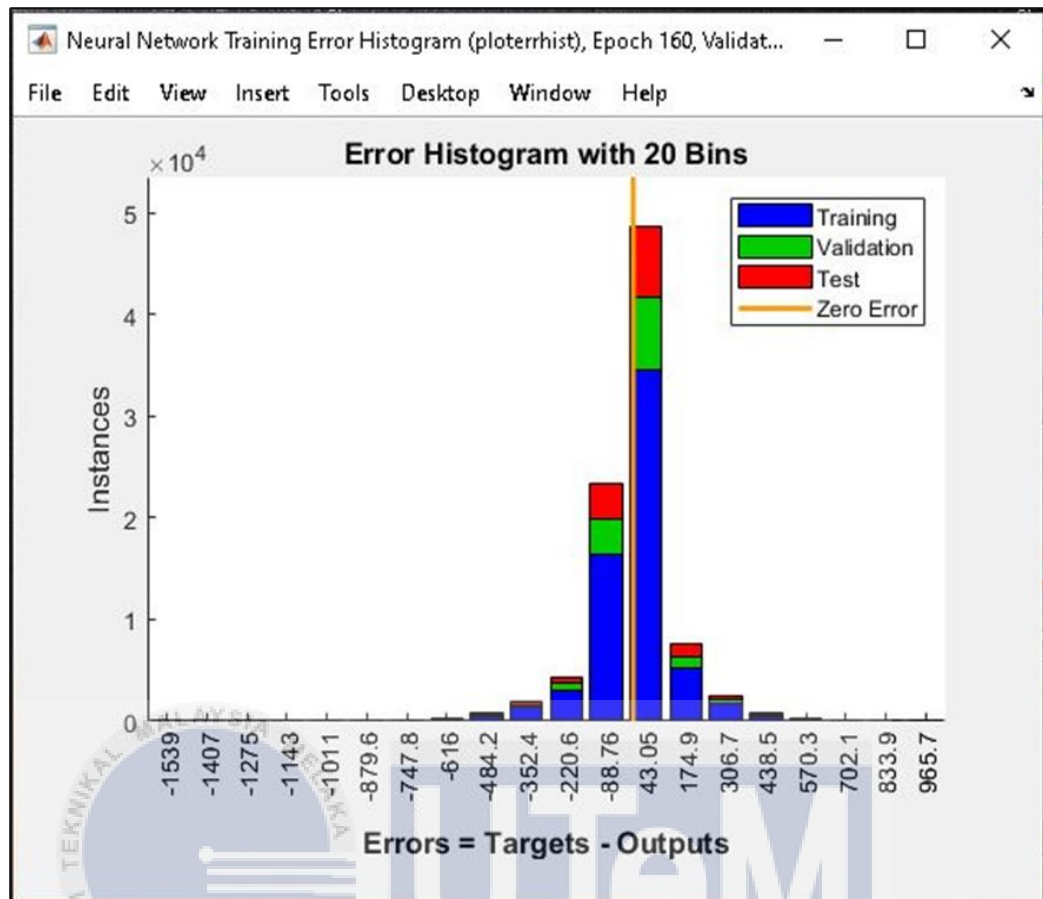


Figure 4.3 Error histogram of trained network model

The figure above was shown the error histogram that occurred in difference between target value and output value (predicted value from neural network model). The error histogram consists of blue, green, and red bars which are represents training data, validation data and testing data, respectively. The histogram could provide with an indication of extreme values, which are data points where fit is considerably worse than most data. From the histogram, the error was fall in between -500 to 500 value.

4.3 Comparison between LSTM and ARIMA model of input prediction

To forecast the solar irradiance in trained neural network model, the inputs that feed into network model need to be predicted. LSTM and ARIMA model were worked for predict every input which are feed into network model to forecast the solar irradiance. Humidity, temperature, pressure, and water vapor are the inputs of network model that were predicted using LSTM and ARIMA model. To compare the performance of LSTM and ARIMA model prediction, the Root Mean Square Error (RMSE) was calculated. The RMSE will measure the differences between actual value and predicted value. Moreover, RMSE used for examining the accuracy of the model prediction. 20 days data was analyzed for the input prediction and the data was separated for training and testing.

4.3.1 Humidity Prediction

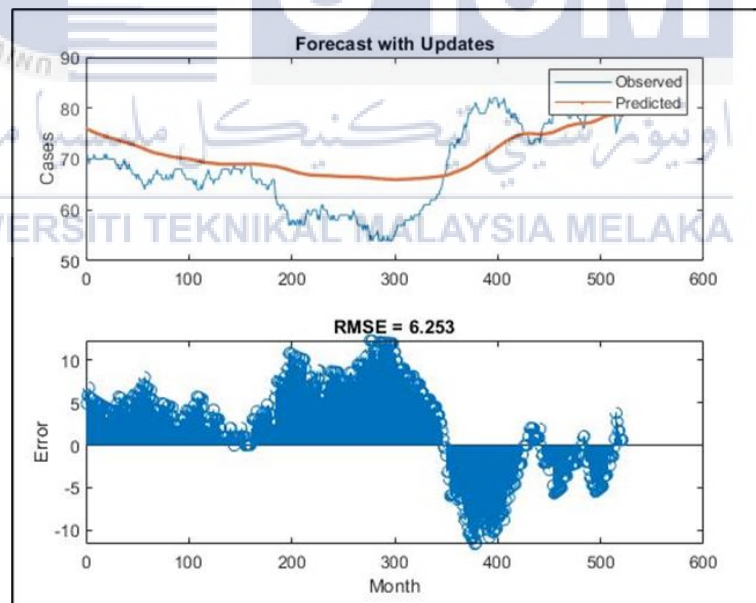


Figure 4.4 Humidity prediction using LSTM model

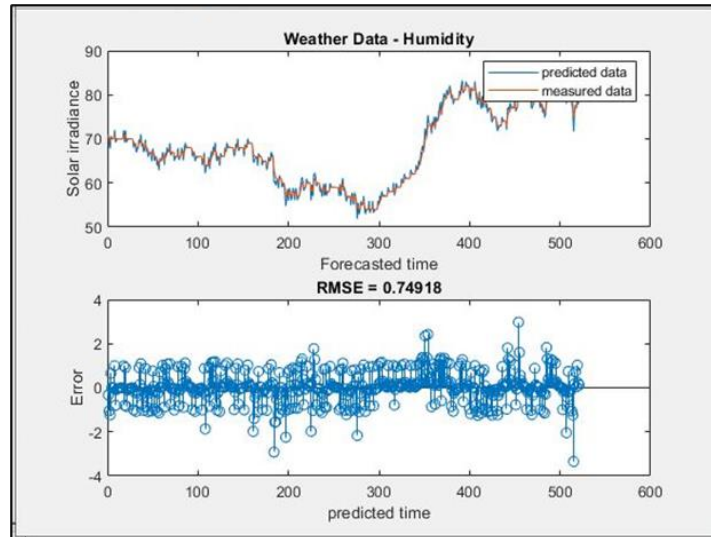


Figure 4.5 Humidity prediction using ARIMA model

Figure 4.3 and figure 4.5 was shown the humidity prediction using LSTM and ARIMA model. The humidity data was predicted start from 10am to 7pm. The RMSE of the LSTM model for humidity prediction is about 6.253 and RMSE of the ARIMA model prediction about 0.74918. By comparing the RMSE of both models, RMSE of ARIMA model is lower than the RMSE of LSTM model.

4.3.2 Temperature prediction

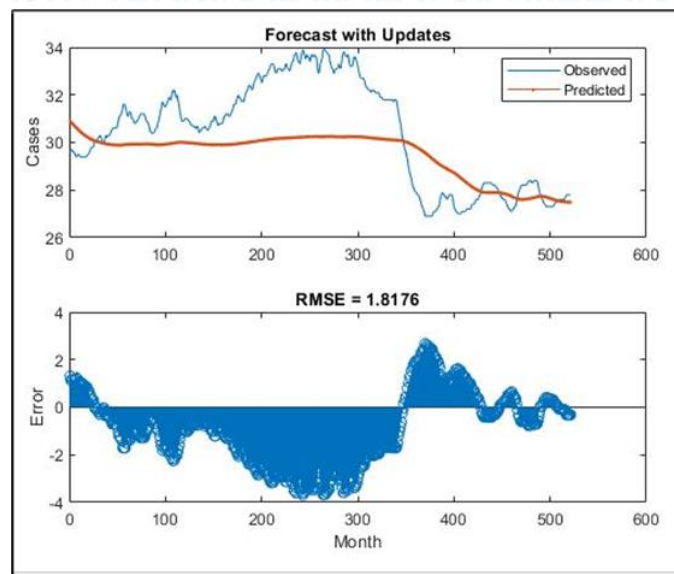


Figure 4.6 Temperature prediction using LSTM model

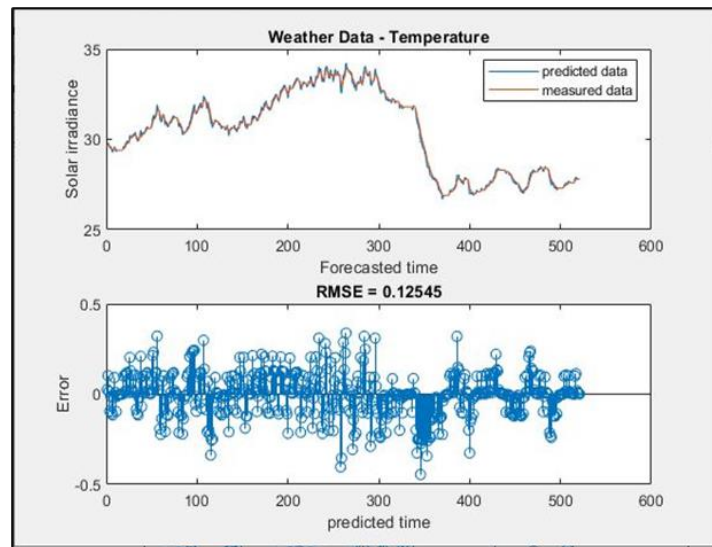


Figure 4.7 Temperature prediction using ARIMA model

Figure 4.6 and figure 4.7 was shown the temperature prediction using LSTM model and ARIMA model. The data was predicted from 10am to 7pm. The RMSE of the LSTM model is 1.8176 and RMSE for the ARIMA model is 0.12545. For LSTM model, the predicted data was good in range of 0 to 170 and 400 to 520. By comparing the RMSE of both models, ARIMA model more accurate than LSTM model.

4.3.3 Pressure prediction

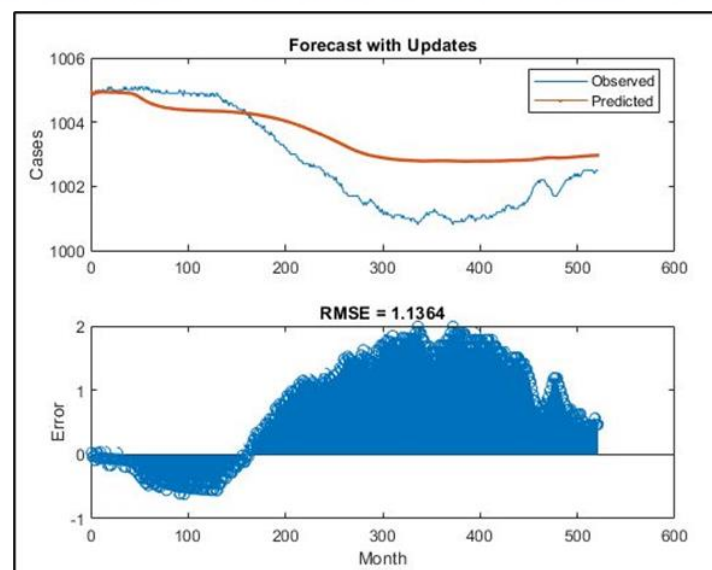


Figure 4.8 pressure prediction using LSTM model

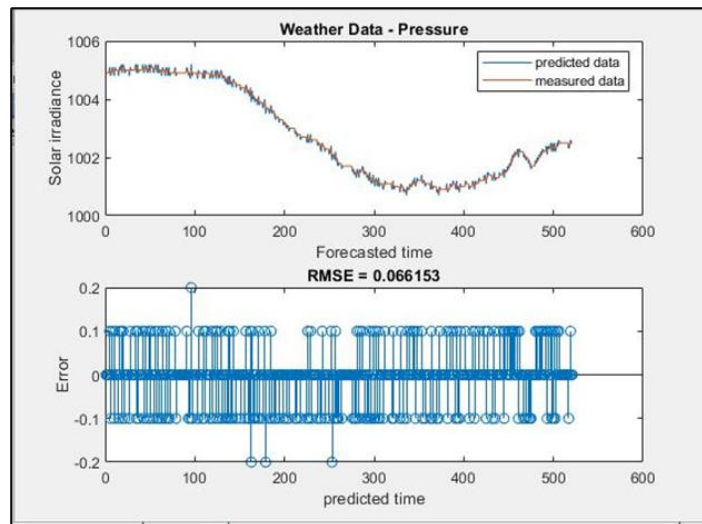


Figure 4.9 pressure prediction using ARIMA model.

Figure 4.8 and figure 4.9 was shows the pressure prediction using LSTM model and ARIMA model. The data was predicted from 10am to 7pm. The RMSE of the LSTM model is about 1.1364 and RMSE of ARIMA model is about 0.066153. By analyzing the RMSE of both models, ARIMA model prediction closer to the actual value compare to LSTM model.

4.3.4 Water Vapor prediction

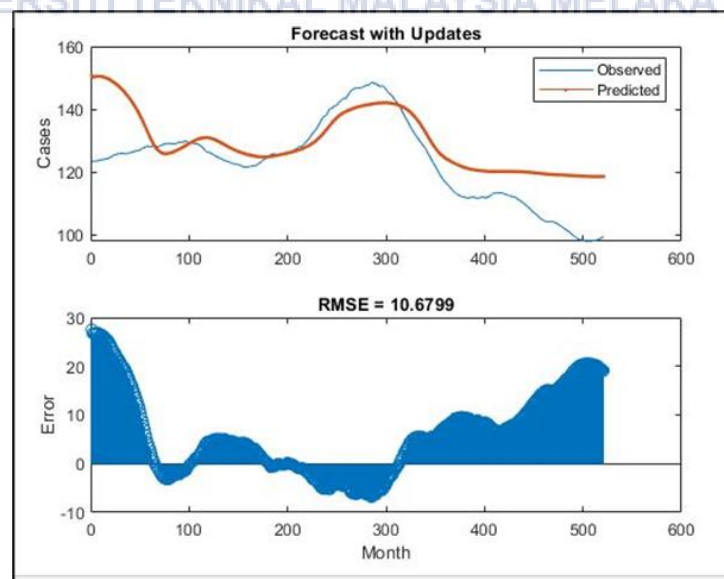


Figure 4.10 Water Vapor prediction using LSTM model

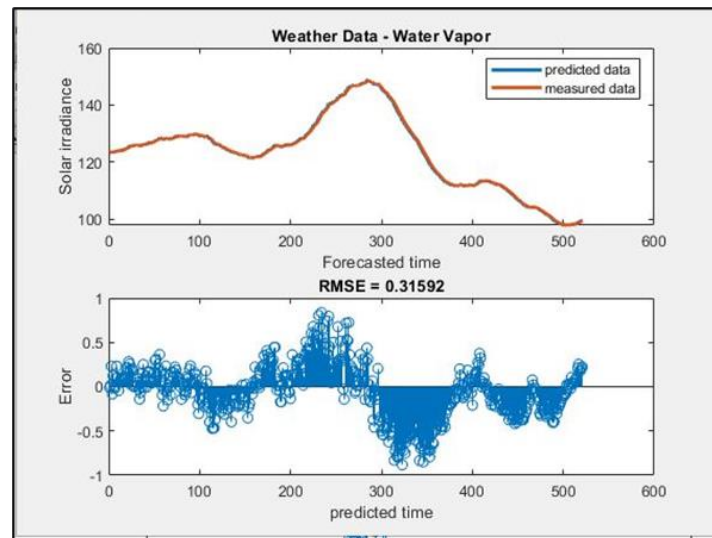


Figure 4.11 Water Vapor prediction using ARIMA model

Figure 4.10 and figure 4.11 was shown the water vapor prediction using LSTM model and ARIMA model. The data was predicted from 10am to 7pm. The RMSE of the LSTM model is about 10.6799 and RMSE of ARIMA model is about 0.31592. By examining the RMSE of both models, ARIMA model prediction closer to the actual value compare to LSTM model.

اونيورسيتي تيكنيكل مليسيا ملاك
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

4.4 Comparison of solar irradiance prediction using indirectly and directly method

The solar irradiance was predicted using indirectly and directly method. The indirectly method is predicting the solar irradiance using the trained neural network model with predicted inputs from LSTM and ARIMA model. The directly method is predicting the solar irradiance directly from LSTM and ARIMA model.

4.4.1 Predict the solar irradiance indirectly

According to the indirectly method, the predicted input data from LSTM and ARIMA model was feed into trained network model to predict the solar irradiance. The predicted input data was collected from 10am to 7pm. The predicted solar irradiance was observed for 1 hour ahead, 3 hours ahead and a day ahead (10am to 7pm).

4.4.1.1 Solar irradiance prediction using LSTM predicted inputs.

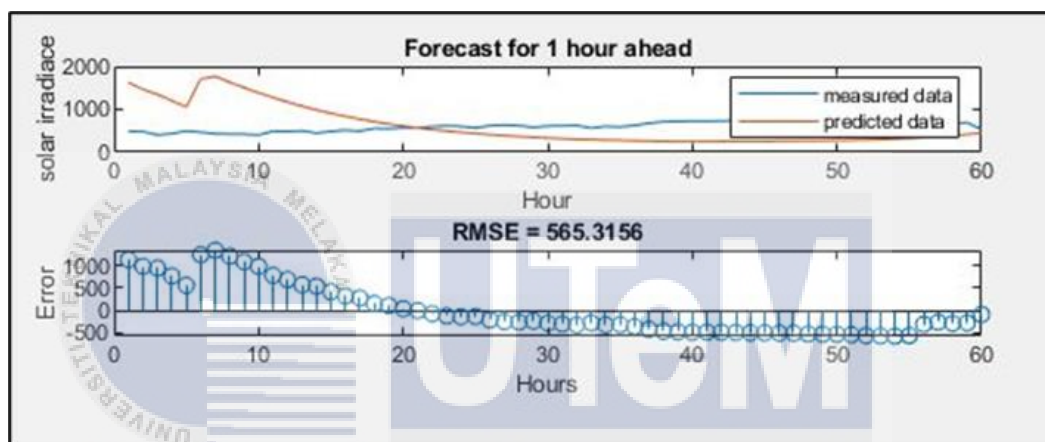


Figure 4.12 Solar irradiance prediction for 1 hour ahead

Figure 4.12 was shown the prediction of solar irradiance in trained network model with LSTM model predicted inputs. The graph was showing the solar irradiance predicted value for 1 hour. The RMSE for 1 hour is about 565.3156. From 0 to 15, the data was not predicted accurately by the trained network model. Because of that data point, the RMSE become higher.

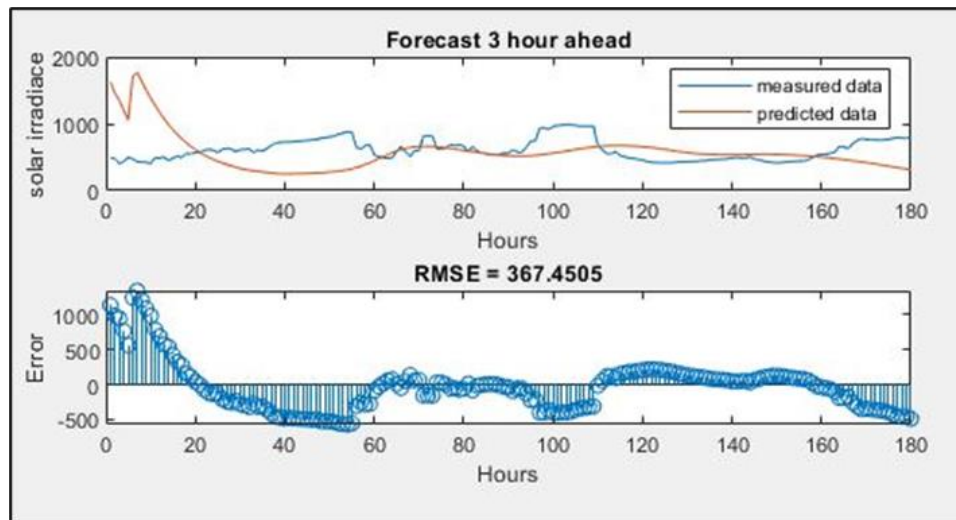


Figure 4.13 Solar irradiance prediction for 3 hours ahead

Figure 4.13 was shown the prediction of solar irradiance in trained network model with LSTM model predicted inputs. The graph was showing the solar irradiance predicted value for 3 hours. The RMSE for 3 hours is about 367.4505. The RMSE was improved compared to RMSE for 1 hour, most of data was predicted nearer to actual values.

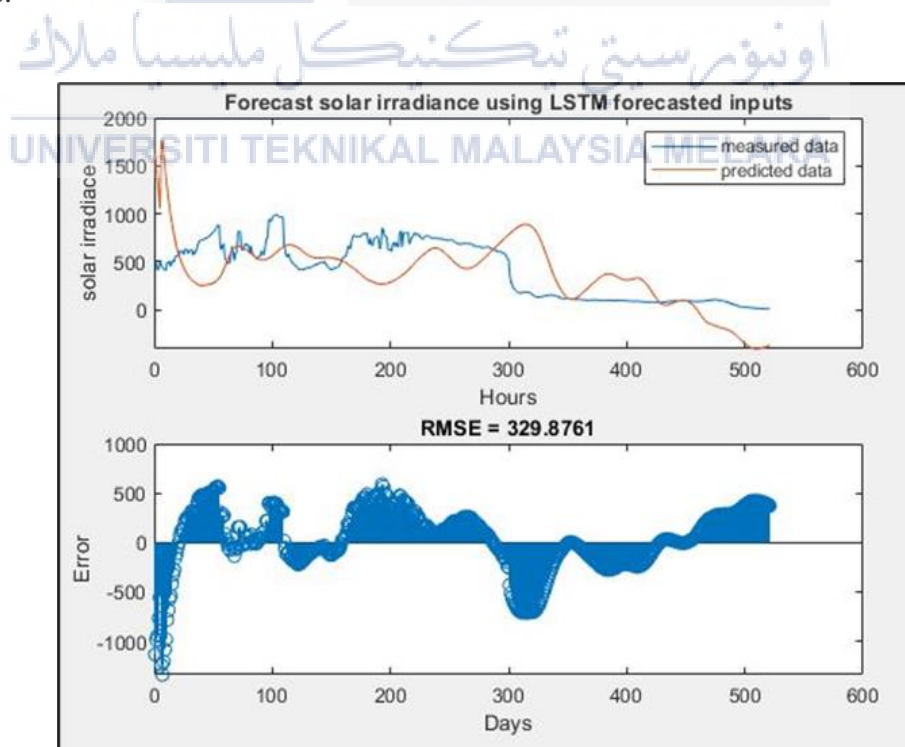


Figure 4.14 Solar irradiance prediction from 10am to 7pm

Figure 4.14 was shown the prediction of solar irradiance in trained network model with LSTM model predicted inputs. The graph was showing the solar irradiance predicted value for 9 hours which from 10 am to 7pm. The RMSE for 3 hours is about 329.8761. The RMSE for 9 hours was improved compared to RMSE for 1 hour and 3 hours, most of data was predicted nearer to actual values.

4.4.1.2 Solar irradiance prediction using ARIMA predicted inputs.

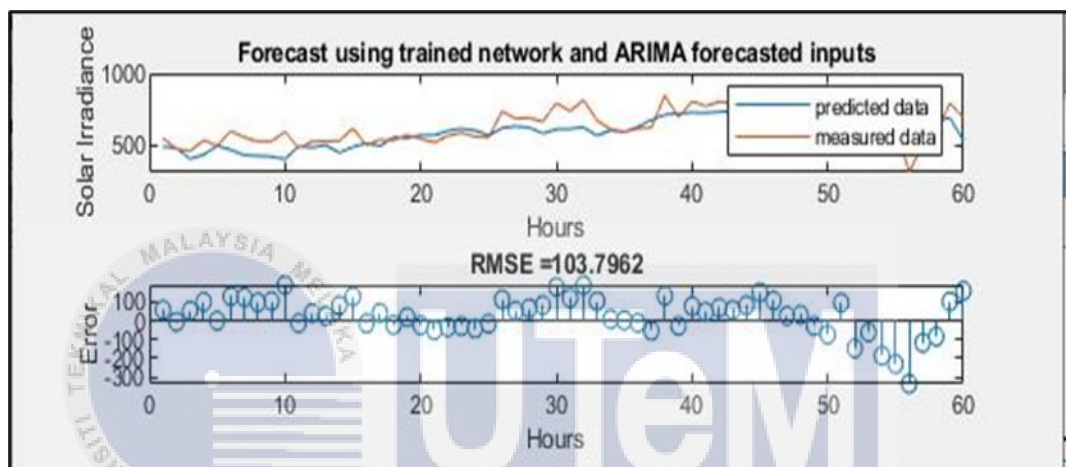


Figure 4.15 Solar irradiance prediction for 1 hour ahead

Figure 4.15 was shown the solar irradiance prediction in trained neural network using ARIMA model predicted inputs. The graph was showing the solar irradiance prediction for 1 hour. The RMSE of 1hour prediction is about 103.7962. Most of the error between predicted value and actual value was occurred in range of 0 to 100. At data point 56, the difference between predicted and actual value was recorded for 300.

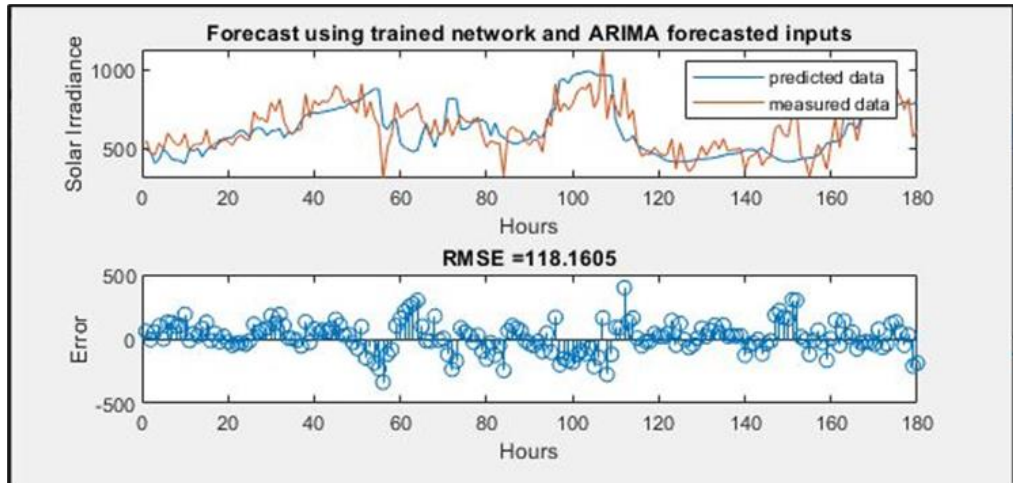


Figure 4.16 solar irradiance prediction for 3 hours ahead

Figure 4.16 was shown the solar irradiance prediction in trained neural network using ARIMA model predicted inputs. The graph was showing the solar irradiance prediction for 3 hours. The RMSE of 3 hours prediction is about 118.1605. The highest error was recorded in a range of 400 at data point of 112.

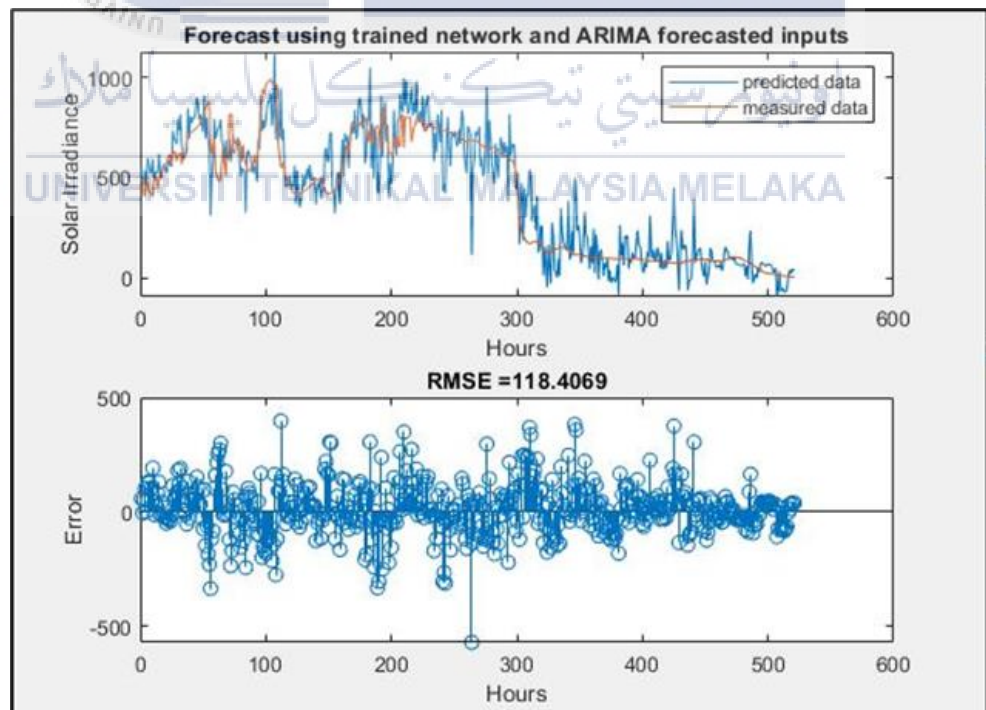


Figure 4.17 Solar irradiance prediction for 10am to 7pm

Figure 4.17 was shown the solar irradiance prediction in trained neural network using ARIMA model predicted inputs. The graph was showing the solar irradiance prediction for 9 hours which from 10 am to 7 pm. The RMSE of 3 hours prediction is about 118.4069. The highest error was recorded in a range of 500 at data point of 260.

4.4.2 Predict the solar irradiance directly

According to the directly method, the solar irradiance was predict directly using LSTM and ARIMA model. The solar irradiance was predicted from 10am to 7pm. The directly predicted solar irradiance was observed for 1 hour ahead, 3 hours ahead and a day ahead (10am to 7pm).

4.4.2.1 Predict solar irradiance directly from LSTM model.

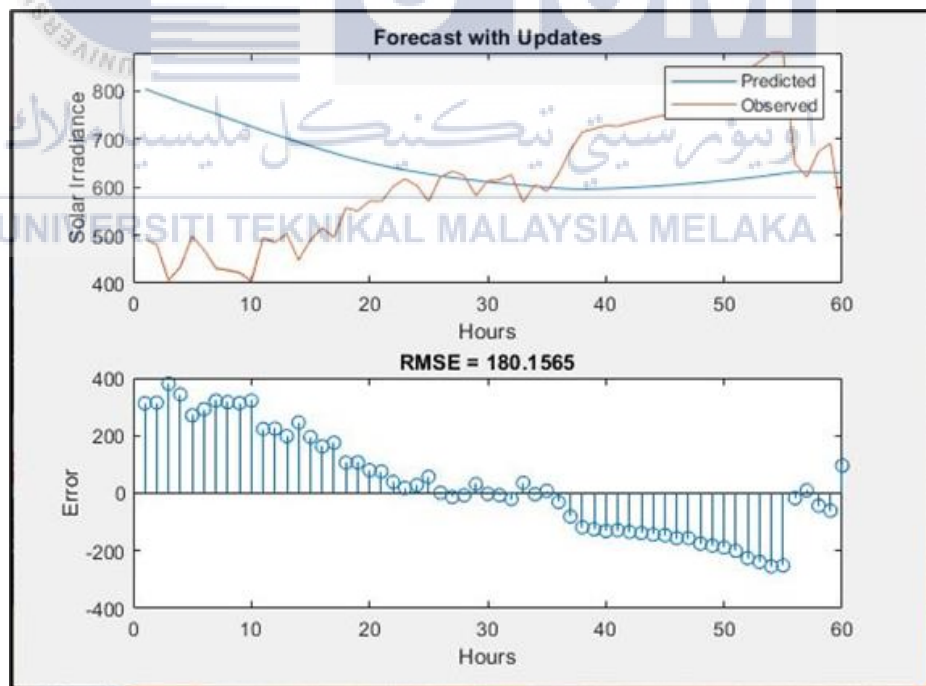


Figure 4.18 Solar irradiance prediction for 1 hour ahead

Figure 4.18 shows the prediction of solar irradiance directly from the LSTM model. The statistical graph was shown the solar irradiance predicted for 1 hour which start from 10am to 11am. The RMSE of the 1-hour prediction is about 180.1565. From the graph was observed that the LSTM model was predicted the solar irradiance nearer to actual value at data point from 24 to 36 and 55 to 60 which was intercept.

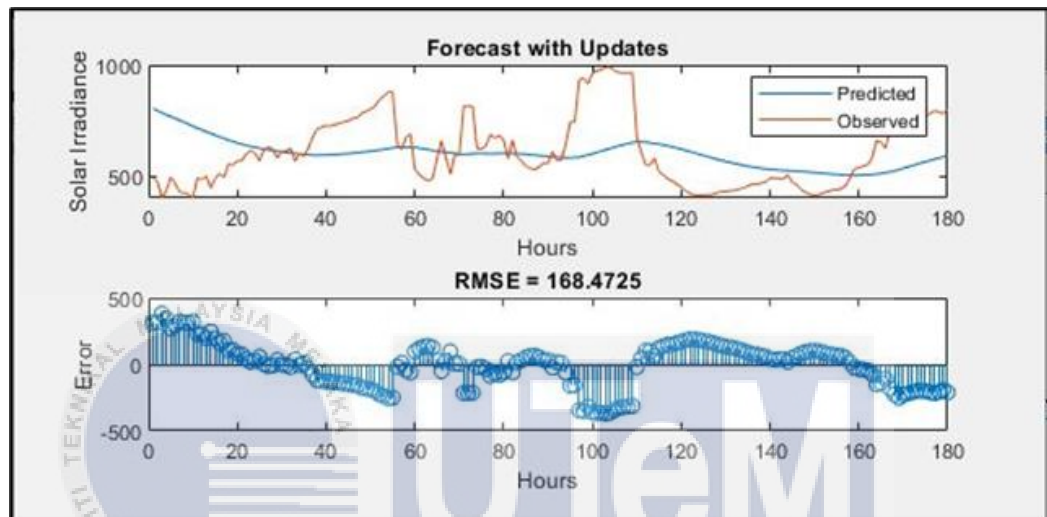


Figure 4.19 solar irradiance prediction for 3 hours ahead

Figure 4.19 shows the prediction of solar irradiance directly from the LSTM model. The statistical graph was shown the solar irradiance predicted for 3 hours which start from 10am to 1pm. The RMSE of the 3-hours prediction is about 168.4725. From the graph was observed that at several data points the LSTM predicted solar irradiance was intercept with actual values which means at some of time model predicted solar irradiance accurately.

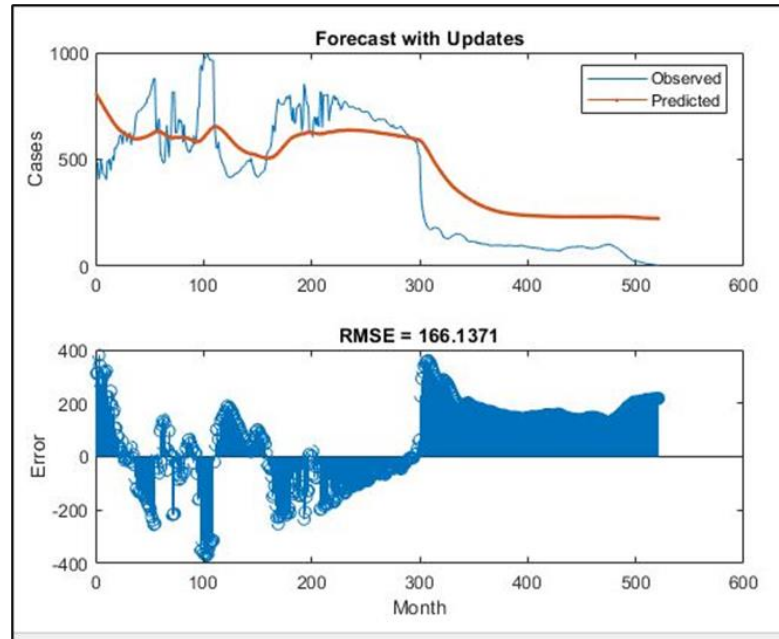


Figure 4.20 Solar irradiance prediction for 9 hours (10am – 7pm)

Figure 4.20 shows the prediction of solar irradiance directly from the LSTM model. The statistical graph was shown the solar irradiance predicted for 9 hours which start from 10am to 7pm. The RMSE of the 9-hours prediction is about 166.1376. From the graph was observed overall that the predicted solar irradiance was followed the trend shaped of actual values.

4.4.2.2 Predict solar irradiance directly from ARIMA model.

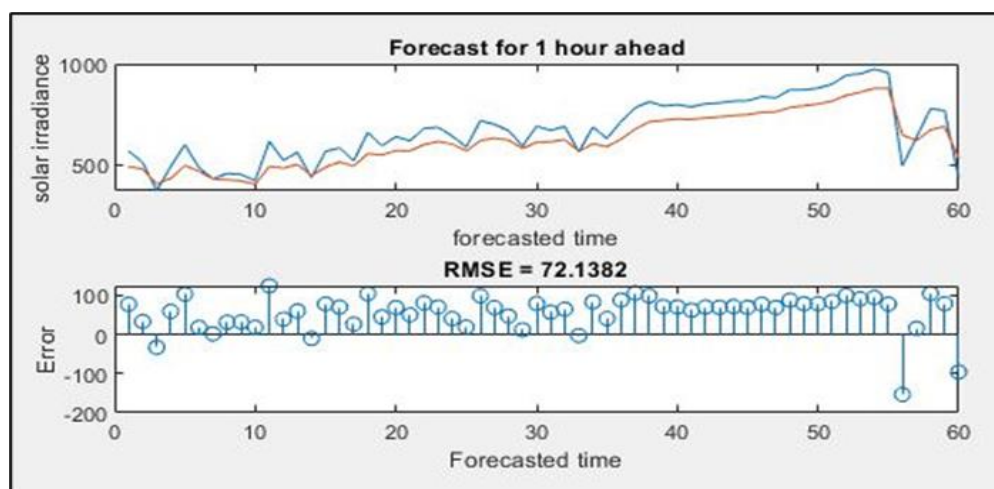


Figure 4.21 solar irradiance prediction for 1 hour ahead

Figure 4.21 was shown the prediction of solar irradiance directly from the ARIMA model. The statistical graph was shown the solar irradiance predicted for 1 hour which start from 10am to 11am. The RMSE of the 1-hour prediction is about 72.1382. From the graph was observed that at data point 56, the difference between predicted value and actual value was recorded for 150.

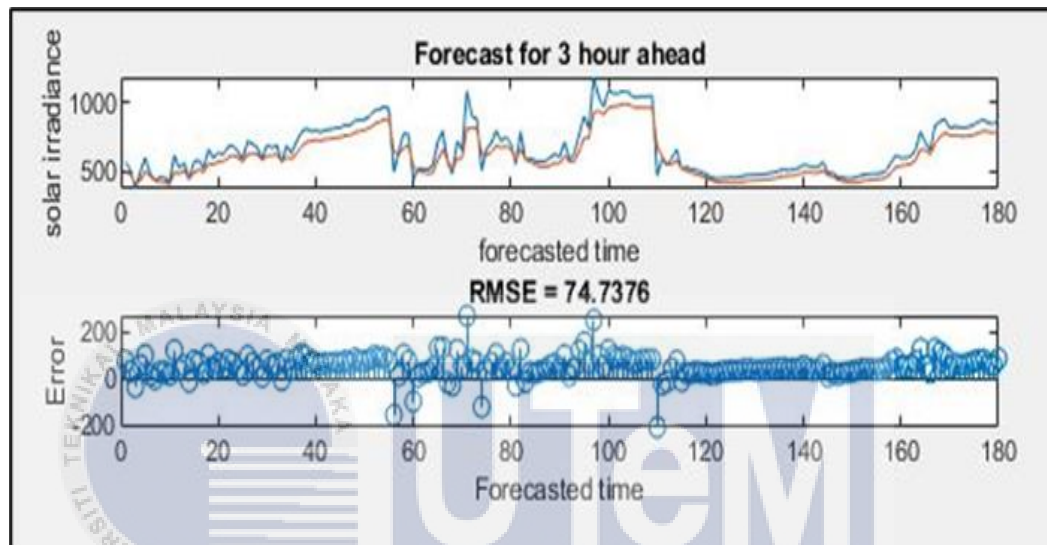


Figure 4.22 solar irradiance prediction for 3 hours ahead

Figure 4.22 was shown the prediction of solar irradiance directly from the ARIMA model. The statistical graph was shown the solar irradiance predicted for 3 hours which start from 10am to 1pm. The RMSE of the 3-hour prediction is about 74.7376. From the statistical observation the maximum error reached around value of 220.

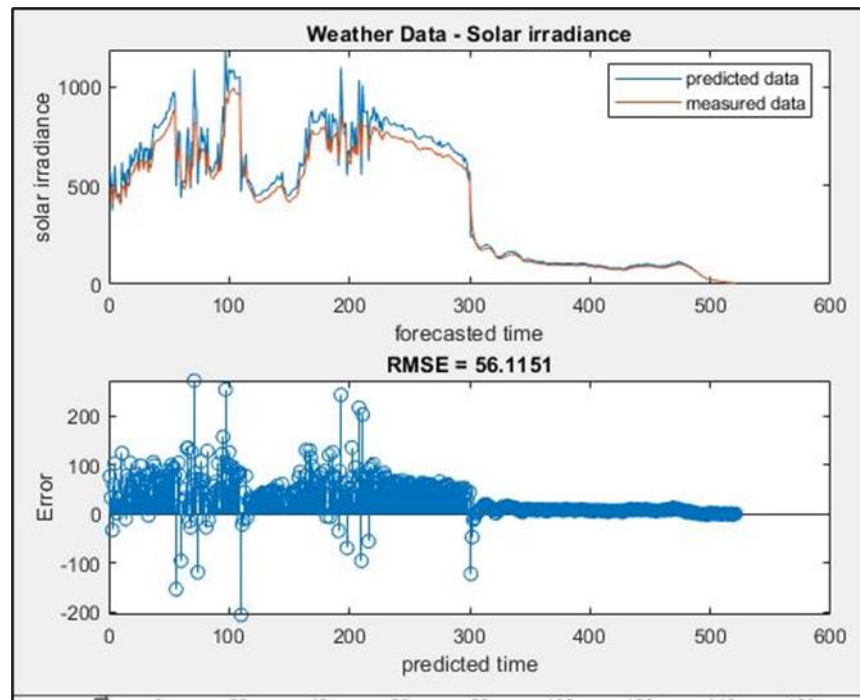


Figure 4.23 solar irradiance prediction for 9 hours ahead

Figure 4.23 was shown the prediction of solar irradiance directly from the ARIMA model. The statistical graph was shown the solar irradiance predicted for 9 hours which start from 10am to 7pm. The RMSE of the 3-hour prediction is about 56.1151. Data point from 300 to the end, the model predicted value almost nearer to actual value.

4.4.3 Comparison between indirectly and directly technique for solar prediction

Table 1 comparison between indirectly and directly techniques with different predicted time and RMSE values.

Technique	Model	Predicted time	RMSE
Indirectly	Neural net – LSTM predicted inputs	1 Hour	565.3156
		3 Hours	367.4505
		9 Hours	329.8761
		1 Hour	103.7962
		3 Hours	118.1605

	Neural net – ARIMA predicted inputs	9 Hours	118.4069
Directly	LSTM	1 Hour	180.1565
		3 Hours	168.4725
		9 Hours	166.1371
	ARIMA	1 Hour	72.1382
		3 Hours	74.7376
		9 Hours	56.1151

The table was shown above, represents the comparison of RMSE values between indirectly and directly methods in different predicted time. The RMSE values were computed using the predicted values and actual values. The aim of determining RMSE values was choose the best accurate model for prediction.

From the comparison table, in indirectly method, RMSE values of predicted solar irradiance using trained network with LSTM predicted inputs were decreased substantially, and RMSE values of predicted solar irradiance using trained network model with ARIMA predicted inputs were increased with longer time prediction. As shown in table, neural network model with ARIMA predicted inputs were used for short-term prediction and long-term prediction compared to LSTM model.

For the comparison of LSTM and ARIMA model in directly method, RMSE of both models were decreased significantly with different time prediction. In this case, overall discussed that the ARIMA model RMSE lower than the LSTM model.

For the overall comparison between indirectly and directly technique, ARIMA model prediction RMSE was lower than the LSTM prediction. Moreover, the solar irradiance predicted directly from ARIMA model more accurate compared to predicted using the trained neural network model with ARIMA predicted inputs.

CHAPTER 5

CONCLUSION AND FUTURE WORKS



The intension of this chapter is to summarize and the recommendation for the future development of this project.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

5.1 Conclusion

The purpose of this project is to forecast the solar irradiance in short-term and long-term according to the weather and water vapor. The first objective of this project is to determine water vapor from the GPS data. To archive this objective, the theory of water vapor calculation was studied well. The GPS data was used into RTKLIB software to find the zenith total delay. Then, zenith total delay (ZTD) was used and further calculation was done to estimate water vapor.

The second objective of this project is to investigate the interaction between water vapor, weather data and solar irradiance measurement. To archive this objective, the neural network model was designed with correlation coefficient ('R') of 0.92 overall. Moreover, the model was created and trained with input values of weather data and water vapor, and target values of solar irradiance. 4 months data which from 23/10/2019 to 29/2/2020 was used to create the model.

The third objective of this project is to forecast the short-term solar irradiance with weather data and water vapor estimated from the GPS data. To archive this objective, 2 types of time series prediction model were used which are LSTM model and ARIMA model. The weather data and water vapor were predicted and feed into the trained network model to forecast the solar irradiance. Moreover, the solar irradiance was predicted directly from the LSTM and ARIMA model and observed with trained network. By compare the RMSE value from indirectly and directly method, ARIMA model was performed well which the predicted data nearer to the actual data. As conclusion, the objectives stated was achieved successfully.

5.2 Future works

Currently the GPS data was retrieved from FKEKK weather station then process it. So, it makes the result shown might be area dependent. In future get some GPS data from outside weather station and observe the water vapor. Moreover, in future work, re-applying this method to other geographical location for model accuracy in a less regular climate. Other than that, longer period of data to gather studies in different seasons and climates such as sunny, cloudy, rainy, and etc. Moreover, forecast the solar irradiance for medium and long term.

REFERENCES

- [1] R. A. ANTHES, H. A. PANOFSKY, J. J. CAHIR, and A. RANGO, “the Atmosphere.,” no. (1975), 1975.
- [2] S. Phokate, “Atmospheric water vapor: Distribution and Empirical estimation in the atmosphere of Thailand,” *J. Phys. Conf. Ser.*, vol. 901, no. 1, 2017, doi: 10.1088/1742-6596/901/1/012051.
- [3] O. Bock *et al.*, “Accuracy assessment of water vapour measurements from in situ and remote sensing techniques during the DEMEVAP 2011 campaign at OHP,” *Atmos. Meas. Tech.*, vol. 6, no. 10, pp. 2777–2802, 2013, doi: 10.5194/amt-6-2777-2013.
- [4] A. E. Niell *et al.*, “Comparison of measurements of atmospheric wet delay by radiosonde, water vapor radiometer, GPS, and VLBI,” *J. Atmos. Ocean. Technol.*, vol. 18, no. 6, pp. 830–850, 2001, doi: 10.1175/1520-0426(2001)018<0830:COMOAW>2.0.CO;2.
- [5] H. Namaoui, S. Kahlouche, A. H. Belbachir, R. Van Malderen, H. Brenot, and E. Pottiaux, “GPS water vapor and its comparison with radiosonde and ERA-Interim data in Algeria,” *Adv. Atmos. Sci.*, vol. 34, no. 5, pp. 623–634, 2017, doi: 10.1007/s00376-016-6111-1.

- [6] S. J. Keihm and J. Liljegren, “Water Vapor Radiometer – Global Positioning System Comparison Measurements and Calibration of the 20 to 32 Gigahertz Tropospheric Water Vapor Absorption Model,” *Jet Propuls.*, no. July 2015, pp. 1–23, 2001, [Online]. Available: https://tmo.jpl.nasa.gov/progress_report/42-144/144E.pdf.
- [7] J. Wang and Z. Liu, “Improving GNSS PPP accuracy through WVR PWV augmentation,” *J. Geod.*, vol. 93, no. 9, pp. 1685–1705, 2019, doi: 10.1007/s00190-019-01278-2.
- [8] L. Bengtsson *et al.*, “The use of GPS measurements for water vapor determination,” *Bull. Am. Meteorol. Soc.*, vol. 84, no. 9, pp. 1249-1258+1162, 2003, doi: 10.1175/BAMS-84-9-1249.
- [9] C. Adams and E. Cohen, *Understanding G4*. 2018.
- [10] L. F. Wiederholt and E. D. Kaplan, “GPS system segments,” *Kaplan (ed.), Understanding GPS Princ. Appl. Artech House, Bost.*, pp. 59–81, 1996.
- [11] F. Hdidou, S. Mordane, and S. Sbi, “Installation and processing of ground-based GPS stations of the Moroccan National Meteorological Service,” pp. 1–8.
- [12] E. A. Sholarin and J. L. Awange, “Global navigation satellite system (GNSS),” *Environ. Sci. Eng. (Subseries Environ. Sci.)*, no. 9783319276496, pp. 177–212, 2015, doi: 10.1007/978-3-319-27651-9_9.
- [13] “GPS signal,” *Appl. Numer. Harmon. Anal.*, no. 9780817643904, pp. 17–30, 2007, doi: 10.1007/978-0-8176-4540-3_2.

- [14] A. Sušnik and K. Bergant, "GPS for Weather and Space Weather Studies," p. 149, 2013.
- [15] L. Pan and F. Guo, "Real-time tropospheric delay retrieval with GPS, GLONASS, Galileo and BDS data," *Sci. Rep.*, vol. 8, no. 1, pp. 1–17, 2018, doi: 10.1038/s41598-018-35155-3.
- [16] M. Igondová, "Analysis of precision and accuracy of Precipitable Water Vapour derived from GPS observations," *Contrib. to Geophys. Geod.*, vol. 39, no. 2, pp. 121–132, 2009, doi: 10.2478/v10126-009-0005-5.
- [17] H. Brenot *et al.*, "A GPS network for tropospheric tomography in the framework of the Mediterranean hydrometeorological observatory Cévennes-Vivarais (southeastern France)," *Atmos. Meas. Tech.*, vol. 7, no. 2, pp. 553–578, 2014, doi: 10.5194/amt-7-553-2014.
- [18] S. Hagemann, "Report No. 340 On the determination of atmospheric water vapour from GPS measurements Authors," *Syst. Sci.*, no. 340, 2002.
- [19] A. Maghrabi and H. M. Al Dajani, "Estimation of precipitable water vapour using vapour pressure and air temperature in an arid region in central Saudi Arabia," *J. Assoc. Arab Univ. Basic Appl. Sci.*, vol. 14, no. 1, pp. 1–8, 2013, doi: 10.1016/j.jaubas.2012.11.001.
- [20] A. E. Lawin, M. Niyongendako, and C. Manirakiza, "Solar irradiance and temperature variability and projected trends analysis in Burundi," *Climate*, vol. 7, no. 6, pp. 1–20, 2019, doi: 10.3390/CLI7060083.
- [21] M. Zaiani, D. Djafer, and F. Chouireb, "New approach to establish a clear sky

- Global Solar irradiance model,” *Int. J. Renew. Energy Res.*, vol. 7, no. 3, pp. 1454–1462, 2017.
- [22] J. Ebinger and W. Vergara, *A World Bank Study: Climate Impacts on Energy Systems*. 2011.
- [23] K. N. Nwaigwe, P. Mutabilwa, and E. Dintwa, “An overview of solar power (PV systems) integration into electricity grids,” *Mater. Sci. Energy Technol.*, vol. 2, no. 3, pp. 629–633, 2019, doi: 10.1016/j.mset.2019.07.002.
- [24] A. Guechi, M. Chegaar, and A. Merabet, “The effect of water vapor on the performance of solar cells,” *Phys. Procedia*, vol. 21, no. May, pp. 108–114, 2011, doi: 10.1016/j.phpro.2011.10.016.
- [25] D. Houcque, “Introduction To Matlab for Engineering Students,” no. August, 2005.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Published works as well as papers presented at conferences, seminars, symposiums etc pertaining to the research topic of the research report/ dissertation/ thesis are suggested be included in this section. The first page of the article may also be appended as reference.



APPENDIX A

```

x = input1';
t = output1';

trainFcn = 'trainlm'; % Levenberg-Marquardt backpropagation.

hiddenLayerSize = 1000;
net = fitnet(hiddenLayerSize,trainFcn);

net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;

[net,tr] = train(net,x,t);

y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y);
view(net);

% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotregression(t,y)
%figure, plotfit(net,x,t)

```

APPENDIX B

```

data = inputsforarima24;

figure
plot(data)
xlabel("Day (7am-7pm)")
ylabel("Cases")
title("Humidity -> weather data")

numTimeStepsTrain = floor(0.963*numel(data));

dataTrain = data(1:numTimeStepsTrain+1);
dataTest = data(numTimeStepsTrain+1:end);

mu = mean(dataTrain);
sig = std(dataTrain);

dataTrainStandardized = (dataTrain - mu) / sig;

XTrain = dataTrainStandardized(1:end-1);
YTrain = dataTrainStandardized(2:end);

numFeatures = 1;
numResponses = 1;
numHiddenUnits = 150;

layers = [ ...
    sequenceInputLayer(numFeatures)
    lstmLayer(numHiddenUnits)
    fullyConnectedLayer(numResponses)
    regressionLayer];
options = trainingOptions('adam', ...
    'MaxEpochs',150, ...
    'GradientThreshold',1, ...
    'InitialLearnRate',0.0001, ...
    'LearnRateSchedule','piecewise', ...
    'LearnRateDropPeriod',125, ...
    'LearnRateDropFactor',0.2, ...
    'Verbose',0, ...
    'Plots','training-progress');

```

```

net = trainNetwork(XTrain,YTrain, layers,options);

dataTestStandardized = (dataTest - mu) / sig;
XTest = dataTestStandardized(1:end-1);
net = predictAndUpdateState(net,XTrain);
[net,YPred] = predictAndUpdateState(net,YTrain(end));

numTimeStepsTest = numel(XTest);
for i = 2:numTimeStepsTest
    [net,YPred(:,i)] = predictAndUpdateState(net,YPred(:,i-1),'ExecutionEnvironment','cpu');
end

YPred = sig*YPred + mu;

YTest = dataTest(2:end);
rmse = sqrt(mean((YPred-YTest).^2))

figure
plot(dataTrain(1:end-1))
hold on
idx = numTimeStepsTrain:(numTimeStepsTrain+numTimeStepsTest);
plot(idx,[data(numTimeStepsTrain) YPred],'.-')
hold off
xlabel("Hours")
ylabel("Prediction")
title("Forecast")
legend(["Observed" "Forecast"])

figure
subplot(2,1,1)
plot(YTest)
hold on
plot(YPred, '-.')
hold off
legend(["Observed" "Forecast"])
ylabel("Cases")
title("Forecast")

subplot(2,1,2)
stem(YPred - YTest)
xlabel("Month")
ylabel("Error")
title("RMSE = " + rmse)

net = resetState(net);
net = predictAndUpdateState(net,inputsforarima24);

YPred = [];
numTimeStepsTest = numel(XTest);
for i = 1:numTimeStepsTest
    [net,YPred(:,i)] = predictAndUpdateState(net,XTest(:,i),'ExecutionEnvironment','cpu');
end

YPred = sig*YPred + mu;

rmse = sqrt(mean((YPred-YTest).^2))

figure
subplot(2,1,1)
plot(YTest)
hold on
plot(YPred, '-.')
hold off
legend(["Observed" "Predicted"])

```


APPENDIX C

```

humObs = length(Humidity);
logGDP = (Humidity(1:end-521));
cpi = Humidity(1:end-521);
T = length(logGDP);
frstHzn = T+1:numObs;
hoCPI = Humidity(frstHzn); % Holdout sample

ToEstMdl = regARIMA('ARLags',1,'MALags',1,'D',1);

Reg4Int = [ones(T,1), cpi]\logGDP;
intercept = Reg4Int(1);

ToEstMdl.Intercept = intercept;
EstMdl = estimate(ToEstMdl,logGDP,'X',cpi,...
    'Display','off')

[gdpF,gdpMSE] = forecast(EstMdl,521,'Y0',logGDP,...
    'X0',cpi,'XF',hoCPI);

figure
h2 = plot(gdpF,'b','LineWidth',2);
hold on
h3 = plot(gdpF+1.96*sqrt(gdpMSE),'r:','...',...
    'LineWidth',2);
plot(gdpF-1.96*sqrt(gdpMSE),'r:','...',...
    'LineWidth',2);
title(['\bf Forecasts and Approximate 95% }'...
    '\bf Forecast Intervals for Humidity}']);
legend([h2,{'Forecasted GDP'}]);
axis tight
hold off
figure
subplot(2,1,1)
plot(gdpF)
hold on
plot(hoCPI)
hold off
legend('predicted data','measured data')
title('Weather Data - Humidity')
xlabel('Forecasted time')
ylabel('Solar irradiance')

subplot(2,1,2)
stem((gdpF)-hoCPI)
rmse = sqrt(mean(((gdpF) - hoCPI).^2))
xlabel('predicted time')
ylabel('Error')
title("RMSE = " + rmse)
prediction = gdpF
t = table(hoCPI,prediction)

```