

Remote Sensing Study of The Temperature Profile Over Penang Island Using Nasa Giovanni System

Kamal Fiqre Md Kamaludin¹, Mohd Muzafa Jumidali², Fathunul Najib Ahmad Sa'ad³ and Abd Rahman Mat Amin⁴

¹Faculty of Chemical Engineering, Universiti Teknologi MARA, Cawangan Pulau Pinang, 13500, Pulau Pinang, Malaysia

²Faculty of Applied Sciences, Universiti Teknologi MARA, Cawangan Pulau Pinang, 13500 Permatang Pauh, Pulau Pinang, Malaysia

³Faculty of Applied Sciences, University Teknologi MARA, 40450 Shah Alam, Selangor Malaysia

⁴Faculty of Applied Sciences, Universiti Teknologi MARA, 23200 Bukit Besi, Terengganu, Malaysia

*corresponding author: ²mohdmuza433@uitm.edu.my

ARTICLE HISTORY

ABSTRACT

Received
17 June 2022

Accepted
6 August 2022

Available Online
2 September 2022

NASA GIOVANNI is an advanced technology of remote sensing system. It is able to provide with Earth Science information within only a few minutes. The focus of this study is to obtain a reliable information via NASA GIOVANNI regarding the temperature trend in Penang Island, Malaysia, from year 2016 to 2018 and to verify the validity and reliability of the data obtained from NASA GIOVANNI and therefore being able to replace with the conventional method of using AccuWeather. From the system, MODIS Aqua is selected as it is the most suitable satellite among others to capture the temperature reading at the desired area. The data obtained from AccuWeather are mostly accurate and clear even though there are several limitations regarding the parameter involved as it is quite general. Meanwhile the temperature data obtained from NASA GIOVANNI are more specific and detailed as it consists of several parameters. For example, NASA GIOVANNI can access the sea surface and land surface temperature for the desired location. Apart from that, NASA GIOVANNI also able to separate between daytime and nighttime temperature data. After obtaining both data from NASA GIOVANNI and AccuWeather, an effective comparison is made via cross validation or to be exact, linear regression. From the graph of linear regression, R-Squared value can be calculated and hence can determine the correlation of both data. From the R-Squared calculations, the values calculated are 0.75 for year 2016, 0.65 for year 2017 and 0.71 for year 2018. According to the R-Squared value, it is concluded that the temperature data from NASA GIOVANNI is having a positive and strong correlation with AccuWeather as the values are exceeding 0.5 for each year. Therefore, the temperature data via NASA GIOVANNI is possible to be obtained and that the obtained temperature data are almost as accurate as AccuWeather.

Keywords: NASA GIOVANNI; AccuWeather; time series study; linear regression; MODIS Aqua

1. INTRODUCTION

The study of surface temperature trends in the post-industrial era is extremely crucial in the science of global warming and climate change and this essential element of the global warming hypothesis is subjected to a great deal of controversy and confusion [1]. Climate change can be defined as a global phenomenon and is particularly evident and has existed in the past three

decades [2]. The average global land and ocean temperature is expected to be increased by 0.85°C from 1880 until 2012, stated by the Intergovernmental Panel on Climate Change (IPCC), in its Fifth Assessment Report. In addition, the IPCC mentioned that within the period of 1983 and 2012 was the warmest in the past 800 years.

There are a large number of research that are related to global warming public opinion. Most of the research focused on the distributions of people's beliefs about the existence and threat of global warming, the predictors and behavioral consequences of these beliefs and how to engage people towards the issues [3]. It is believed that an increased amount of temperature is highly related to this global warming issue. There are several factors that lead to the occurrence of global warming and need to be prevented from happening for a better future. One of the factors is greenhouse gasses. Basically, greenhouse gasses will lead to the greenhouse effect. It consists of concentrations of gas such as water vapor, carbon dioxide (CO₂), methane and ozone. The largest contributing source of greenhouse gas is the burning of fossil fuels leading to the emission of carbon dioxide. The increasing amount of carbon dioxide gas and other gasses that trap heat in the atmosphere especially have warmed the Earth and as a result, it proposes a severe event such as rising sea level, melting snow and ice, higher occurrence of extreme heat events, fires, extreme storms, unstable rainfall and floods [4].

A phenomenon extensively related to the effect of temperature has been occurring in every part of the world, including in Malaysia. Regarding the issues related to temperature, several researches done forecasted that there would be constant changes regarding the climate in which would indirectly pose severe effects towards the forests, agriculture, freshwater supplies, coastlines and other natural resources. Malaysia is a country that can be classified as equatorial rainforest and fully humid climate region [5]. Malaysia has experienced a few seasons throughout the year. Unlike any other countries located in west, Malaysia is usually affected by only two monsoon seasons and two inter-monsoon seasons that normally occur at certain period of time and only at certain places will receive the exact weather situation. The two monsoon seasons consist of two different places namely southwest monsoon and northeast monsoon. As for the southwest monsoon, the season occurs from May to August while the northeast monsoon starts from November until February [6]. However, there will always be a slight change of starting month and the period of time of the occurrence of these monsoon which can be either sooner or a bit later than usual.

AccuWeather is one of the famous local and international weather forecasts [7]. It is known to be very familiar and commonly used weather forecast for quite a period of time. Apart from that, AccuWeather also collects information based on the real-time data therefore the data from this weather forecast are safe to say that it is almost accurate and reliable. Despite having an accurate result, this conventional method seems to have several drawbacks especially concerning the manpower involvements and the time consumption as well. Moreover, the limitation of location also contributes to the weaknesses of the current conventional method as it requires one to be on exact location to obtain the information. By that, AccuWeather is a perfect match for NASA GIOVANNI to undergo an effective comparison regarding the information of real-time collection data.

Therefore, the purpose of this study is to obtain a reliable information and data of temperature profile in Penang Island via NASA GIOVANNI and to observe and compare the reading of temperature via NASA GIOVANNI system. This is done with the real time actual reading from

AccuWeather and location of the study will be in Penang Island. Besides, it is also to verify the validity and reliability of the temperature data obtained from NASA GIOVANNI and therefore is able to be replaced with the current method via AccuWeather.

2. METHODOLOGY

According to the reviews, there are several online softwares with remote sensing technology system that are able to give access to the users of obtaining a reliable data regarding the temperature trend of a certain place over a certain period. NOAA and OPeNDAP seem to be inefficient and not suitable for this study as it involved scaling issues over terrestrial surface and automated scripts failure respectively. NASA GIOVANNI is selected to be the method for this study as it is easy to understand the user interface of the system and the data obtained are clear and straightforward as there are no further processes need to be done towards the raw data. The flowchart of overall process for obtaining data and verifying the data of GIOVANNI system is depicted in Figure 1.

2.1 MODIS Aqua and Terra

For this study, it is mainly related to time series study on a temperature effect over Penang Island. Specifically, this study does not involve with any laboratory procedures and hence will be undergo fully using a remote sensing technology, NASA GIOVANNI. The technology is capable of displaying environmental data and analyzing the desired geophysical parameters. Users are able to select the plot as it comes with variety of options such as map choices, comparisons vertical choices, time series choices and miscellaneous choices. As for the platform, MODIS-Aqua and MODIS-Terra are selected as it involved the specified places, Penang Island for the period of time and capable of offering the desired parameters of this study.

2.2 Cross Validation

After analyzing the data of temperature trend using NASA GIOVANNI, an effective comparison needs to be made with the data obtained from local and international weather forecast, AccuWeather hence achieving the second objective of this study. In general, cross validation is a measurement of evaluating or accessing the performance of a predictive model as well as analysis the independent dataset statistically [8]. It is one the important and effective techniques that was used and applied by a data scientist to validate the stability of the model. On the other hand, the concept of applying cross validation is mainly due to the capability of the techniques of having low bias and variance outcomes. This will create a very accurate and transparent result. There are several types of cross validation techniques. For this study, a regression analysis is chosen as the method to compare both data of temperature from NASA GIOVANNI and AccuWeather.

2.3 Linear Regression

Regression is known to be one of the famous methods commonly used nowadays. It is a form of statistical modeling for analyzing the relationship between variables as well as visualizing the correlation between variables by using a scatterplot [9]. In the scatter plot, each axis represents one variable whereby a positive correlation need to be initiated from the bottom left of the graph and end at the top right corner [10]. In more detailed explanation, a linear regression is an extension of Pearson's correlation (Pearson r). It is divided into two different types, simple

linear regression, and multiple linear regression. Basically the difference of these two types is mainly by the number of predictor variables which mean that one predictor variable is known as simple linear regression whereby multiple linear regression is normally dealing with multiple predictor variables [10]. These are the equation for both simple and multiple linear regression.

$$Y = a + bX + u \quad (1)$$

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_tX_t + u \quad (2)$$

After plotting the graph of linear regression via scatter plot, another corresponding value need to be calculated in order to compare both set of temperature data from NASA GIOVANNI and AccuWeather which is the value of R-Squared (R^2). Fundamentally, R^2 is a statistical measure that represents the proportion of the variance for a dependent variable that is explained by an independent variable or variable in a regression model [11]. R^2 values are commonly stated in percentage whereby the value ranges from 0 to 1. For this study, higher amount of R^2 value needs to be expected as it shows that both temperature has a strong correlation.

3. RESULT AND DISCUSSION

3.1 AccuWeather

AccuWeather is one of the famous local and international weather forecasts [7]. It is known to be very familiar and commonly used weather forecast for quite a period. Apart from that, AccuWeather also collects information based on the real-time data. Therefore, the data from this weather forecast are safe to say that it is almost accurate and reliable. By that, AccuWeather is a perfect match for NASA GIOVANNI to undergo an effective comparison regarding the information of real-time collections data. Hence, we are able to verify the validity and reliability of NASA GIOVANNI as mentioned in the third objective of this study.

3.1.1 Temperature Profile via AccuWeather

From this study, AccuWeather is selected as the main reference that will later be compared with NASA GIOVANNI. From this weather forecast, the location of the study is selected to be at Penang Island from year 2016 to 2018. Data of monthly temperature reading are obtained from all three years involved. Figure 2 shows the temperature profile via AccuWeather for year 2016, 2017 and 2018.

3.2 NASA GIOVANNI

According to the first objective of this study, it is clearly seen that the main purpose of this study is to obtain a collection of accurate data of temperature profile in Penang Island. Several variables such as the period time are remained constant whereby the collection date started from year 2016 until 2018. Unlike AccuWeather, the data obtained from NASA GIOVANNI are more detailed and specific as it is not only capable of reading the land surface temperature (LST), but also able to obtain the sea surface temperature (SST) as the Penang Island is surrounded by the ocean. Apart from that, NASA GIOVANNI is also able to separate between the temperature for daytime and nighttime within the collection date. For this study, MODIS

Aqua is chosen over MODIS Terra as the data from MODIS Aqua entirely cover both LST and SST meanwhile MODIS Terra only for LST parameter [12].

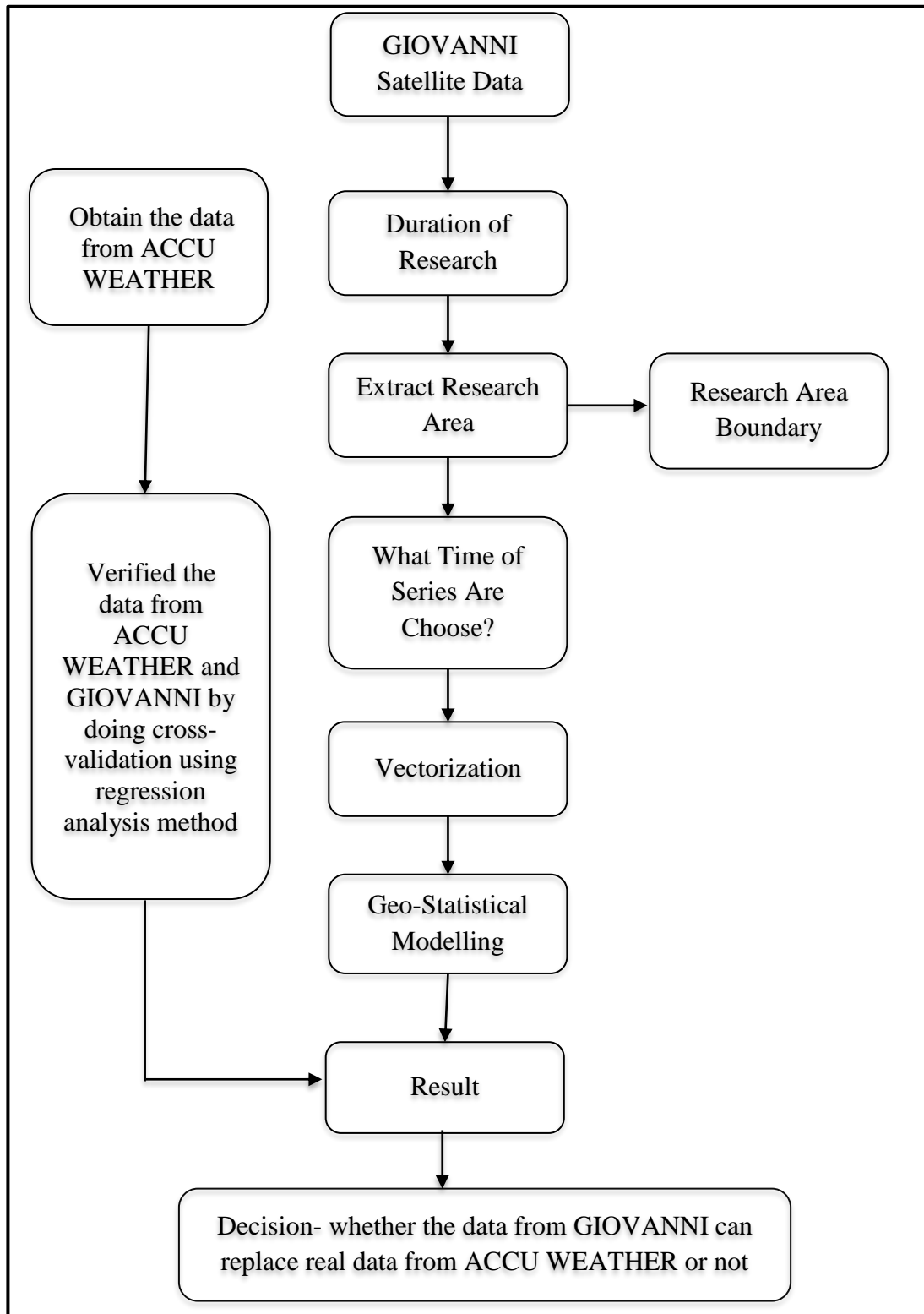


Figure 1: Flowchart process for obtaining and verify the data of GIOVANNI system.

Temperature Profile via AccuWeather

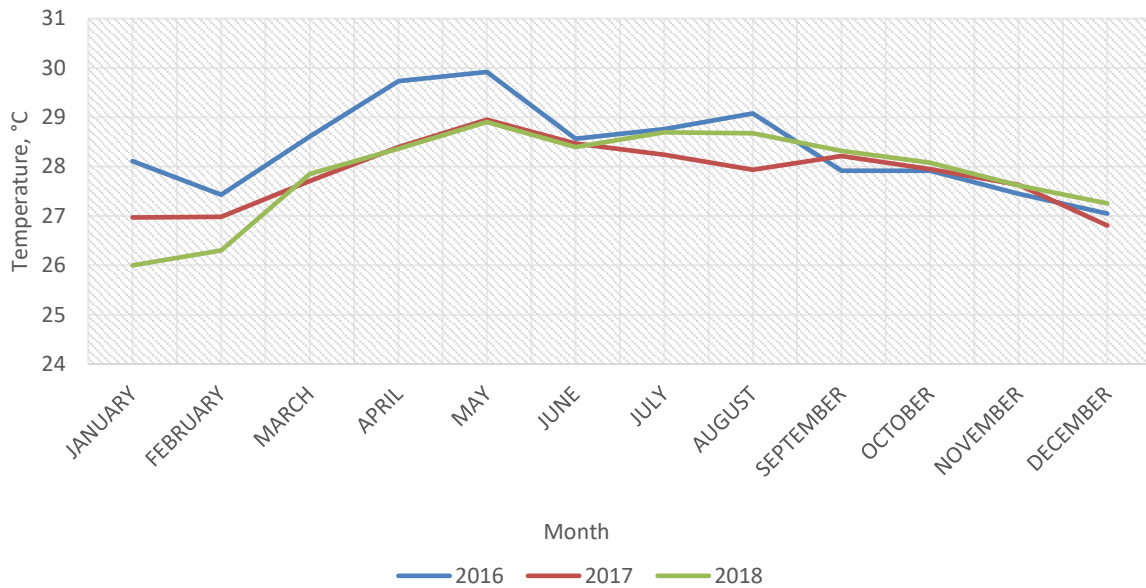


Figure 2 : Temperature Profile via AccuWeather

3.2.1 Sea Surface Temperature (SST)

Sea surface temperature (SST) is basically the temperature of the water located near the surface of the ocean. Since Penang Island is surrounded by the Malacca Strait, NASA GIOVANNI is able to determine the temperature data of that sea surface. Figure 3 shows the sea surface temperature of both daytime and night-time in year 2016, 2017 and 2018, respectively.

Sea Surface Temperature via NASA GIOVANNI

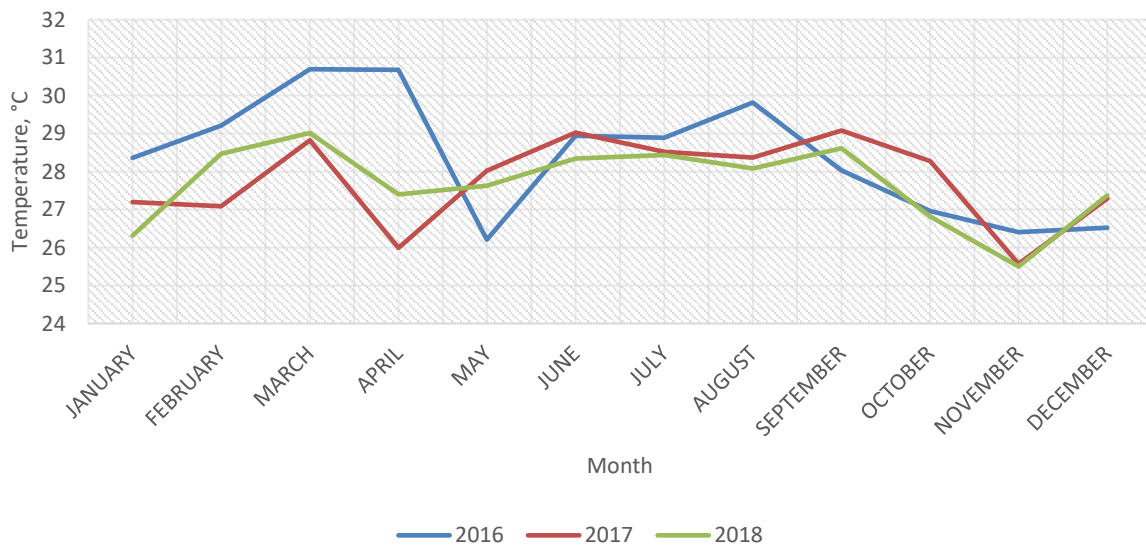


Figure 3 : Sea Surface Temperature via NASA GIOVANNI

3.2.2 Land Surface Temperature (LST)

Land Surface Temperature (LST) is one of the main parameters for this study. As for the definition, LST can be defined as the radiative skin temperature of a land derived from the solar radiation. In other words, it is basically how hot the surface of the Earth would feel to the touch in a particular location. There are several factors that were found to fundamentally influence the derivation of this LST parameter, such as [7],[12]:

- Temperature variations with angles
- Sub-pixel in-homogeneities in temperature and cover
- Surface spectral emissivity at the channel wavelengths
- Atmospheric temperature and humidity variations
- Clouds and large aerosol particles such as dust

For this study of analyzing the temperature profile for year 2016 to 2018, it was done by using MODIS Aqua as the selected satellites of collecting the temperature data. In general, the existing purposes of monitoring of analyzing LST data is to monitor the increasing rate of an atmospheric greenhouse gas that may affect the LST and also how rising land surface temperature affect glaciers, ice sheets, permafrost and the vegetation in Earth's ecosystems. For the data collected, Figure 4 is obtained regarding the LST in Penang Island, Malaysia for 2016, 2017 and 2018.

3.3 Cross Validation via Linear Regression

As mentioned in the previous sub-section, all the temperature data from both AccuWeather and NASA GIOVANNI system are obtained at the desired location which is in Penang Island, Malaysia for year 2016, 2017 and 2018. The temperature profile has been analyzed regarding the trend and the flow of the graph. Hence, an effective comparison and cross validation can be made by using linear regression as the most suitable and easiest method of defining the correlations between two sets of different data. For this study, AccuWeather will be the reference for NASA GIOVANNI system in order to determine the accuracy and consistency in term of data collecting system as it is fundamentally known that AccuWeather has been established and operated since 1962 [11]. Therefore, data from NASA GIOVANNI need to be as close and as accurate as the data from AccuWeather in order to determine their reliability and validity as a way of achieving the third objective of this study. In order to visualize into making comparison, both data need to be cross validated. Hence, a linear regression graphs were generated for each year. Figure 5, Figure 6 and Figure 7 are the linear regression between NASA GIOVANNI and AccuWeather in 2016, 2017 and 2018, respectively.

Land Surface Temperature via NASA GIOVANNI

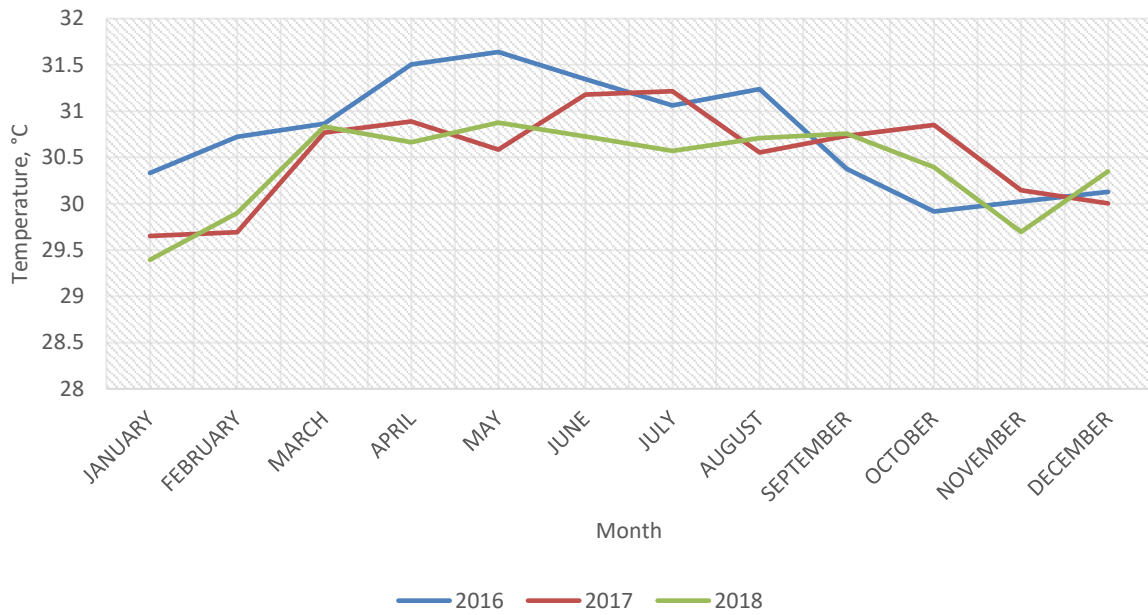


Figure 4: Land Surface Temperature via NASA GIOVANNI

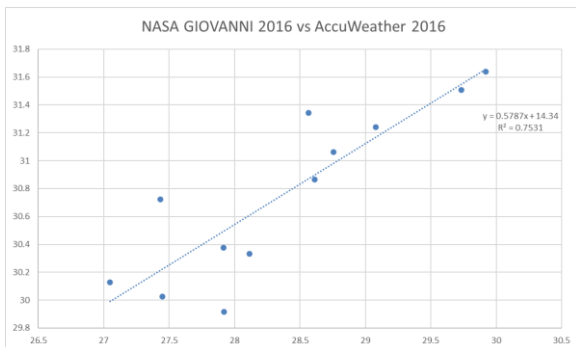


Figure 5: Linear Regression in 2016

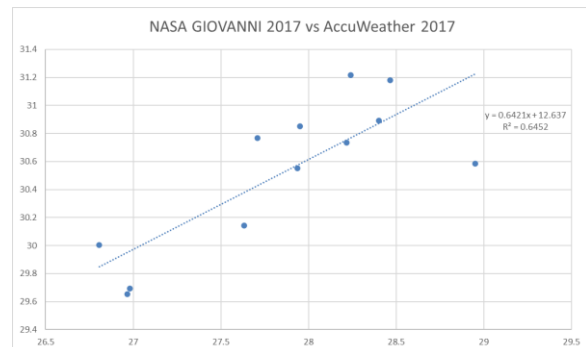


Figure 6: Linear Regression in 2017

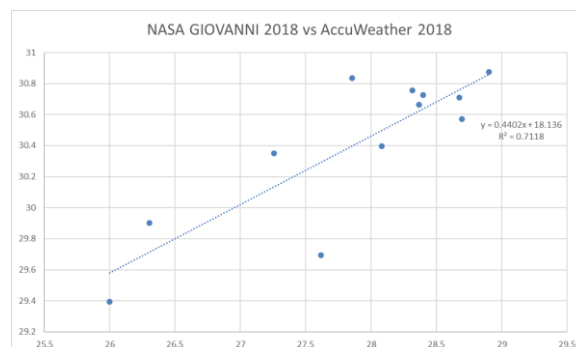


Figure 7: Linear regression in 2018

3.3.1 R-Squares (R^2) Value

Linear regression graphs are generated from both data for each year. A linear regression graph is a very simple and trusted method of determining the correlations between two sets of different data. From the linear regression graph generated, the R- Squared (R^2) values are calculated and interpreted. An interpretation of R^2 is the most common method of calculating and interpreting this parameter [11]. Fundamentally, the calculation of R^2 involves several parameters and values such as the regression and error value. The value of R^2 is normally written in percentage. For example, if the data from two different sources have a correlation for R^2 value of 0.5, it shows that the 50% or half of the data are fit or correlate with each other [13]. In other words, half of the data are accurate, while the other half are inaccurate. Equation 3-8 are used to calculate the R^2 value from the linear regression graph [8].

$$R^2 = 1 - \frac{SS_{Regression}}{SS_{Total}} \quad (3)$$

From Figure 4, Figure 5 and Figure 6, the R^2 are interpreted from the linear regression graph for both data from AccuWeather and NASA GIOVANNI. As calculated, the R^2 values that have been calculated for both data are 0.7531, 0.6452 and 0.7118 for year 2016, 2017 and 2018 respectively. According to the obtained R^2 value, generally the data from both AccuWeather and NASA GIOVANNI have a strong and positive what? with each other as the R^2 values are exceeding 0.5 or 50% for each year. It shows that the percentage from 60 % - 70 % of the data from NASA GIOVANNI are as accurate as data from AccuWeather. An R^2 value from 70 % - 100 % indicates that the given data are closely tracked while a score between 0 % - 40 % indicates a very low correlation [13],[8].

4. CONCLUSION

To summarize it all, there are basically three objectives of this study. The first objective is to obtain a reliable information and data from NASA GIOVANNI regarding the temperature trend over Penang Island for research purposes. Next is to compare the reading of temperature data via NASA GIOVANNI system with the real time actual reading that can be obtained from AccuWeather by using linear regression. Lastly is to verify the reliability, accuracy and consistency of the NASA GIOVANNI system. The study of temperature reading is located in Penang Island, Malaysia for three years, 2016, 2017 and 2018.

There are several organizations and associations that are involved in remote sensing technology. Basically, it is an advanced technology of collecting real time actual data via satellites. There are several options that were shortlisted which are NOAA, OPeNDAP and NASA which each of them has their own capability. However, NOAA seems to face scaling issues regarding terrestrial surface and difficult for in-situ measurement. Meanwhile, OPeNDAP seems to face issues as it only discovers the data from within the file and also tend to face automated script failure. Therefore, NASA is chosen to be the most suitable organization dealing with remote sensing technology. On the other hand, MODIS Aqua will be the selected platform or the responsible satellites for the desired location and selected period of time.

Data from NASA GIOVANNI and AccuWeather are obtained and the temperature profile for each year are analyzed. Both data need to be cross validated with an effective and unbiased

comparison. Therefore, simple linear regression method is chosen as is the easy and result are trusted and unbiased. From the line graph, the R^2 value are interpreted and the correlations for both data can be concluded. The R^2 value are 0.7531, 0.6452 and 0.7118 for year 2016, 2017 and 2018 respectively. The value for each year exceeds 0.5 which shows that the data obtained from NASA GIOVANNI have a positive and strong correlation with AccuWeather for each year.

As mentioned earlier, there are three objectives of this study. It is shown that all of the objectives are successfully achieved. The temperature data via NASA GIOVANNI is possible to be obtained in Penang Island, Malaysia for year 2016, 2017 and 2018. Next, an effective comparison is done for both data from AccuWeather and NASA GIOVANNI via linear regression and hence obtaining the R^2 value. From the R^2 value, it shows that most of the data from NASA GIOVANNI are as accurate as AccuWeather and therefore achieving the third objective of this study which is to verify its validity and reliability.

ACKNOWLEDGEMENT

The author are grateful to the Universiti Teknologi MARA Cawangan Pulau Pinang, Faculty of Chemical Engineering for the physical and technical supported in this work.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

- [1] J. Munshi, "Trend profiles of atmospheric temperature time series," pp. 1–12, 2016.
- [2] K. Ho and D. Tang, "Science of the Total Environment Climate change in Malaysia : Trends , contributors , impacts , mitigation and adaptations," *Sci. Total Environ.*, vol. 650, pp. 1858–1871, 2019.
- [3] A. T. Abeles, L. C. Howe, J. A. Krosnick, and B. Macinnis, "Perception of public opinion on global warming and the role of opinion deviance," vol. 63, no. April, pp. 118–129, 2019.
- [4] N. Ahmad *et al.*, "The Adaptation Towards Climate Change Impacts Among Islanders In Malaysia," 2019.
- [5] T. Hujan and M. Barat, "Trends in Peninsular Malaysia Rainfall Data During the Southwest Monsoon and Northeast Monsoon Seasons : 1975 – 2004," vol. 39, no. 4, pp. 533–542, 2010.
- [6] "Malaysia - Country Profile - Nations Online Project." .
- [7] Z. Q. Huang, Y. C. Chen, and C. Y. Wen, "Real-time weather monitoring and prediction using city buses and machine learning," *Sensors (Switzerland)*, vol. 20, no. 18, 2020.
- [8] A. Rohani, M. Taki, and M. Abdollahpour, "A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I)," *Renew. Energy*, vol. 115, pp. 411–422, 2018.
- [9] S. Chung, Y. W. Park, and T. Cheong, "A Mathematical Programming Approach for Integrated Multiple Linear Regression Subset Selection and Validation," 2017.
- [10] G. K. Shrepnek, "Regression methods in the empiric analysis of health care data.," *J. Manag. Care Pharm.*, vol. 11, no. 3, pp. 240–251, 2005.
- [11] A. Gelman, B. Goodrich, J. Gabry, and A. Vehtari, "R-squared for Bayesian Regression Models," *Am. Stat.*, vol. 73, no. 3, pp. 307–309, 2019.
- [12] A. Prados, G. Leptoukh, and J. Johnson, "Visualization and exploration of NASA air quality remote sensing data via Giovanni," *Air Waste Manag. Assoc. Symp. Air Qual. Meas. Methods*

- Technol. 2008*, pp. 453–457, 2008.
- [13] B. Waske, S. van der Linden, C. Oldenburg, B. Jakimow, A. Rabe, and P. Hostert, “ImageRF - A user-oriented implementation for remote sensing image analysis with Random Forests,” *Environ. Model. Softw.*, vol. 35, pp. 192–193, 2012.