GIS-AIDED GEOGRAPHICAL AND METEOROLOGICAL DATA OVERVIEW OF SOLAR RADIATION MAPPING FOR MALAYSIA – AN EXPLANATORY STUDY BASED ON SOLAR RADIATION PREDICTION MODELING USING NEURAL NETWORK APPROACH

Mohd Rizman Sultan Mohd*, Fazlina Ahmat Ruslan, Juliana Johari

Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM) Shah Alam, 40450, Shah Alam, Malaysia

Graphical abstract

Abstract

Solar radiation mapping has used geographical and meteorological data. To obtain geographical and meteorological data, a Geographic Information System (GIS) is required. GIS is defined as an integrated geographic resource that presents data in terms of spatial information. This data is important for Neural Networks as it will be used as input parameters for the development of solar radiation prediction models. Solar radiation prediction is one way to map the sun's rays in certain places where there are insufficient resources or space to build a complete solar radiation measurement station. Since predictions about solar radiation require meteorological and geographical data, this paper will give an overview of GIS-assisted geographical and meteorological data to be used as input parameters for solar radiation mapping which will eventually be used as input for prediction models developed for the whole country of Malaysia using Neural Networks. Based on the results, the prediction model developed managed to obtain a coefficient of determination, $R^2$ value of 0.9329.

Keywords: GIS, Spatial, Solar Radiation, Prediction Modeling, Neural Network

Abstrak

Pemetaan tenaga mahatari menggunakan data-data geografi dan meteorologi. Bagi mendapatkan data-data ini, Sistem Maklumat Geografi (GIS) diperlukan. GIS ditakrifkan sebagai sumber geografi bersepadu yang menyediakan data dan maklumat berbentuk spatial. Data ini merupakan input penting yang akan digunakan dalam proses pembangunan sistem ramalan tenaga matahari menggunakan kaedah Neural Network. Sistem ramalan tenaga matahari ini digunakan bagi mendapatkan peta potensi tenaga matahari bagi kawasan-kawasan tertentu yang tidak mempunyai sumber atau ruang yang mencukupi bagi pembinaan stesen pengukuran tenaga matahari. Kertas kajian ini akan membentangkan gambaran keseluruh data geografi dan meteorologi yang diperolehi daripada sistem GIS untuk digunakan sebagai input bagi pembangunan model ramalan tenaga matahari di Malaysia dengan menggunakan kaedah Neural Network. Model ramalan yang dihasilkan oleh kaedah Neural Network berjaya mencatat nilai $R^2$ sebanyak 0.9329.

Kata kunci: GIS, Spatial, Tenaga Matahari, Model Ramalan, Neural Network

© 2021 Penerbit UTM Press. All rights reserved

83:6 (2021) 19–33 | https://journals.utm.my/jurnalteknologi | eISSN 2180–3722 | DOI: https://doi.org/10.11113/jurnalteknologi.v83.16634
1.0 INTRODUCTION

The sun is the center of our solar system. The sun provides energy in the form of light and heat through a continuous reaction of radiation that strikes the earth at all times. With this energy, the earth can get enough energy resources to allow organisms to live on it. With the study of converting solar energy into electricity through photovoltaic systems, most countries in the world take seriously the development of infrastructure and in-depth study of the potential of this solar energy and how they can maximize this exchange process to generate more electricity for their use and even generate income through it.

Efforts to implement this solar power plant construction project are difficult due to financial constraints [1]. Therefore, the researchers have created a new strategy to conduct a preliminary study on the potential of solar energy for a specific area before the implementation of the project. Based on the results of the previous study, it is found that most researchers have implemented the development of a forecast system to find out the extent of solar energy potential that can be generated for that specific area of interest [2]. Among the systems developed, the prediction model using the Neural Network is the most widely applied. There are many new discoveries successfully obtained by this method [3-4]. Therefore, this paper will review some of the results of the past studies that have been conducted on the development of solar radiation prediction systems using Neural Network techniques.

The forecast model developed for solar radiation requires a set of input data. Many approaches have been applied to design and develop predictive models for solar radiation using various input parameters. These input parameters are basically based on meteorological and geographical data that can be retrieved from the GIS observation database [5]. GIS is a complete and integrated geographical data center that delivers data in terms of spatial information. This spatial information is used according to the type of usage [6]. Many countries are participating in developing their own GIS database including Malaysia. As a developing country, Malaysia needs more sources of electricity to run many industries at the same time, providing sufficient electricity for the needs of the people. Figure 1 shows the Electricity Consumption for Malaysia for 2014 until 2019 [7]. Based on the figure, it is shown that the electricity consumption in Malaysia is increasing with the growth of population and economic-driven industries such as commercial, industrial, and agriculture. Being able to receive solar energy throughout the year makes solar energy an excellent alternative energy source and will eventually become a major source of electricity for the whole country in the future.

Figure 1: Electricity Consumption for Malaysia

This paper will provide an overview of GIS-aided geographical data and meteorological data to be used as an input parameter for mapping solar radiation through several sections. The first part will introduce an overview of solar radiation. The second will cover the geographical and meteorological data retrieved from the GIS system to be used as an input parameter for the prediction models development for solar radiation. The third part covers literature on prediction model developed for solar radiation prediction including the Neural Network approach. Neural Network approach had been widely used to develop a prediction model for solar radiation. With the numerous paper presented in this topic, this paper will also highlighted their findings to prepare the best strategies to produce an precise prediction for solar radiation values. This paper will end with a brief discussion of the findings before a distinct closure in the conclusion section.

2.0 SOLAR RADIATION

Solar energy is described as an amount of energy that radiates from the sun. Solar radiation is the study on the propagation of solar energy from the sun to the earth. This electromagnetic energy is in the form of radiant light and heat [8]. This radiant energy is essential for the metabolism of the environment and its inhabitants. The solar radiation spectrum divides into three bands with different amounts of a composition consisting of 42.3% visible light, 49.4% infrared radiation, and 8.3% ultraviolet radiation [9]. The intensity of solar radiation depends on several factors such as latitude, season, time of the day, cloud cover and altitude [10]. Solar radiation has not totally reached the earth’s surface as it is being absorbed, reflected, or scattered in the atmosphere. The solar energy that can be absorbed by the earth’s surface is defined as direct solar energy while the solar energy that is scattered as it enters the atmosphere is defined as indirect solar radiation.
Solar radiation is measured in wavelengths and frequency. The wavelength is defined as the distance from peak to peak of the sine-wave of the radiation and measured in meters, m. Frequency is defined as the wavelength’s cycles in 1 second time and measured in Hertz, Hz. The relationship between wavelength and frequency is reciprocal, thus the shorter wavelength produced higher frequencies. A shorter wavelength is more energetic than a longer wavelength as the energy of wavelength increases with the frequency value [11].

Solar irradiance is the intensity value of solar radiation which enters the earth’s atmosphere. Solar irradiance can be defined as the amount of radiant flux on an area and is measured in watts per meter square (Wm⁻²). Solar irradiance received by a particular location depends on the elevation above the sea level, the angle of the sun, and scattering elements such as clouds [12]. The higher the elevation is, the shorter path it travels from the atmosphere. The lower the angle of the sun, the larger amount of ozone that the light has to pass through. The angle of the sun depends on the latitude, seasons, time of the year, and time of the day. During the day, the angle of the sun will be decreased from sunrise until noon and will increase until sunset. At a greater angle, solar radiation has to pass through more atmosphere which reduces its irradiance.

Cloud coverage and air pollution reduce the amount of solar radiation on the earth’s surface. Clouds and aerosols in the atmosphere can scatter and absorb all radiation bands. Because cloud cover increases in this situation, the angle of the sun becomes less important when measuring solar irradiance due to the increase of radiation scattering. Surface temperature is indirectly dependent on solar radiation. It receives heat from the earth’s surface by conduction and convection [13]. Earth’s surface absorbs infrared radiation and converts it to thermal energy, making it warmer than the surrounding atmosphere. The heat is transferred by conduction from the warmer earth’s surface to the cooler atmosphere. As the air is a poor conductor of heat, the rise and fall of warm and cool air eventually warm the rest of the atmosphere. The rising warm air is referred to as thermal. As the warmed air rises, cooler air sinks to the surface, where it continues in the convection process [14]. Since infrared radiation is being reflected back to the air, this radiation can be trapped and absorbed by the gases in the atmosphere or eventually re-radiated back to the earth [15]. This process is defined as the greenhouse effect.

Solar radiation prediction is commonly measured using global solar radiation data. Prediction on global solar radiation is made using suitable empirical correlations based on measured data at certain places using known meteorological parameters such as air temperature, sunshine hours, relative humidity, latitude and longitude of the targeted place, and others [16-17]. Geographical data retrieved from the GIS has also been used as the parameters to predict the solar radiation values such as irradiance data, solar azimuth angle, wind speed, surface pressure, and solar zenith angle.

### 3.0 GIS-BASED DATA

Malaysia is a country situated in the equatorial region. Its climate depends on the northeast and southwest monsoons throughout the year. From May to September, southwest monsoons dominate Malaysia’s climate until northeast monsoons start from October to March [18]. The intermission between these two monsoons was filled by the heavy rainfall. Malaysia is divided into two regions; Peninsular and Borneo Island, but since it is located at the same equatorial line, both areas received the same monsoons climate influences. Studies show that Malaysia had received an average of 170 times rainy days throughout the year [19]. The average ambient temperature is between 26°C to 29°C [20]. The relative humidity is constantly on average of between 80 to 88% based on its altitude above sea level [21]. The development of solar radiation prediction models is based on several geographical and meteorological input parameters which are dedicated to the weather and climate of Malaysia. Among other parameters, several parameters had been chosen based on the database provided by the GIS system. The parameter chosen for the prediction model development is based on their frequent use in the previous studies and also the significant contribution to the overall performance of the prediction model. The entire GIS data used in this study is gained from the SOLCAST Database (https://solcast.com).

Defined as a powerful database that contains data related to the earth’s surface, GIS has been an important medium for spatial data mining. GIS applies in location-based studies where there is various information that can be retrieved such as human population and its topologies, landscape information including water resources mapping and crops planting, electrical consumptions, and climate of a specific area [22-24]. There are many studies that have utilized the GIS database including solar radiation forecast, flood risk assessment, pollution mapping, and many more.

The data retrieved from GIS are in the form of cartographic, photographic, or tabulated values stored in a spreadsheets format. Cartographic data defined as the data which is stated in map-form including geographical information and survey-data results, while photographic data defined as visual-aided data [25]. Besides these data, other digital data also had been contributed to the GIS. This digital data can be retrieved from satellite imaginary applications and remote-sensing-based systems. This application is widely used for agricultural development in Malaysia whereas they used the data to plan for the plantation and harvesting time setting. Other practical uses of this remote-sensing-based system are the mapping for the disease mapping and health-risk assessment.
GIS had been used to display spatial relationships and linear networks. Spatial relationships display topography-based data which relates to the human’s population such as agriculture mapping and settlement. Linear or geometric networks is the indicator used in the map to show the river and connected roads and grids in GIS. It also can perform a boundary line of places such as demographic differences and other attributes. GIS gathers, align, and scaling these entire data to give a full overview to ease the data analysis process.

The application of GIS has eased the data collection of input parameters to develop solar radiation prediction. Entire data gained in the GIS module is basically from the measurement tools which are directly located at specific locations throughout the nation. Each input parameter is measured and collected before it being placed at the GIS database as cumulative data for Malaysia. There is plenty of other data available on the GIS but focusing on the solar radiation prediction, parameters are chosen based on their relation and impact towards generating a solar radiation value [26].

3.1 Irradiance Parameter

The irradiance parameter is the direct measurement for the solar radiation value. It is necessary to include the irradiance parameter for prediction model development to produce an accurate prediction based on these irradiance trend. Global Horizontal Irradiance (GHI) shows the measurement of the total terrestrial radiation received by a horizontal surface on the ground [27]. It is measured using a pyranometer; an instrument with a hemispherical angle view. Figure 2 shows an illustration of GHI measurement on the earth’s surface.

![Figure 2 Illustration of GHI measurement](image)

Based on Figure 2, it is shown that there are multiple parameters that influenced the total GHI received by the solar panel. In other words, GHI also referred to the total amount of solar radiation from Direct Normal Irradiance (DNI), Diffuse Horizontal Irradiance (DHI), and ground-reflected radiation as shown in equation (1) below.

\[ GHI = DHI + DNI \cdot \cos (\theta_Z) \]  

(1)

Where \( \theta_Z \) is an incident angle of the beam or zenith angle of the sun.

DNI is referred to as direct beam radiation, the radiation which manages to penetrate the atmospheric layer and goes directly on the surface of the earth [28]. It is measured by an instrument called a pyrheliometer; an arrangement of photosensitive elements called thermopile sensors at the base of a light-collimating tube built in a glass-type material. DNI is the most concentrated solar radiation and the most dominant solar irradiance compared to DHI [29]. DNI can be measured as shown in Figure 3.

![Figure 3 Illustration of DNI measurement](image)

The tilted surface is widely applied in photovoltaic system equipment to produce a maximum solar energy potential. This can be achieved by applying motorized equipment on the solar panel that always keeps the incident angle to 90°. The measurement of sunlight is based on the sunshine duration with the pattern as shown in Figure 4.

![Figure 4 Solar Irradiance and Sunshine Duration Pattern](image)

Based on Figure 4, it is indicated that the maximum solar energy potential that could be retrieved and generated using a photovoltaic system is at noon. That is when the sun is directly perpendicular to the surface of the earth.
DHI is defined as the amount of solar radiation received by the earth’s surface which did not directly come from the sun [30]. DHI is the solar radiation that has been scattered by particles at the atmospheric layer of the earth. It can also refer to as illumination from clouds. Ground-reflected irradiance also counts in the DHI properties.

The irradiance parameter is notably the main parameter for solar radiation prediction. But with the absence of the measuring equipment and the dependencies of other parameters had made it not be necessary to be used in the prediction modeling. It can also be predicted using natural-linked parameters that had been applied in weather forecast applications such as temperature, humidity, surface pressure, and others.

### 3.2 Solar Azimuth Angle & Solar Zenith Angle

Solar radiation measurement varies from one place to another. It is also measured when the sun is visible to that area. In other words, it only applicable during daytime. There is some measured parameter which is directly implied to the sun location as pinpoint towards the direction on the surface of the earth. These parameters are solar azimuth angle and solar zenith angle. Basically, these solar properties are directly related to the sunshine duration. Figure 5 shows the monthly average sunshine duration for Malaysia.

Based on Figure 6, the azimuth angle is measured on the horizontal surface of the earth within the equatorial line. The starting point is on the North pole which indicates the 0° angle. Solar azimuth angle, \( \theta_A \) can be calculated using equation (2).

\[
\theta_A = \left[ \sin \delta \cos L - \cos \delta \sin L \cos h \right] / \cos \alpha
\]  

(2)

Where \( \alpha \) is defined as the elevation, \( L \) is defined as the latitude, \( \delta \) is defined as the declination [32] as shown in equation (3), \( \alpha \) is the angle between sun location and the horizontal surface, and \( h \) is defined as the hour angle as shown in equation (4).

\[
\delta = 23.45 \sin \left[ \frac{360}{365} (n + 284) \right]
\]  

(3)

Where \( n \) is the day of the year.

\[
h = - \frac{t - 12}{12}
\]  

(4)

Where \( t \) is refers to the current time.

Solar zenith angle is defined as the angle between vertical line pin-pointed at the earth’s surface and the line of sight from the point to the sun [33] and can be formulated as equation (5).

\[
\theta_z = \cos^{-1} (\sin \delta \sin \varphi + \cos \delta \cos \varphi \cos h)
\]  

(5)

Where \( \varphi \) is defined as the latitude of the location.

The value of angle depends on the position of the sun, the value of the solar zenith angle varies with the local time and date [34]. Therefore, the solar zenith angle is also dependent on the location parameter consisting of latitudes and longitudes values.

Solar azimuth angle and solar zenith angle are basically the reference angle to determine how much is the time taken for the sun to appear on the particular surface. It is also used to predict the sunshine duration which is widely used as one of the key parameters in solar radiation prediction.

### 3.3 Wind Speed

The wind is caused by the flow of air pressure differences as air tends to move from the high-pressure area to the low-pressure area [35]. The anemometer shown in Figure 7 is equipment used to measure the wind speed. It is also used to detect the direction of the wind with a revolving cup at the top of it. This revolving cup generates electricity equivalent to the wind speed.
Wind speed is measured in distance from one point to another at a specific time. In order to further classify wind speed and its impact on the earth’s surface; land and sea, it is denoted by Knots. The wind speed measured by anemometer will be converted from kilometers per hour to miles per hour \[ WS_{mph} = \frac{WS_{km/h}}{1.609344} \] using equation (6): 

After retrieved the wind speed in miles per hour, \( WS_{mph} \) it further converted into feet per hour and then divided by the number of feet in a knot to finally gain the wind speed values in knots [36] as shown in equation (7):

\[ WS_{Knots} = WS_{mph} \times 0.8689762 \] (7)

As for meteorological data used in the weather forecast, each wind speed categorized based on its value. This is used to further describe the wind speed characteristics to be more understandable. The wind speed also counts as the input parameter for solar radiation prediction because it gives an impact on the air and surface temperature on the measured location. It is also widely used as a parameter for weather prediction along with relative humidity measurement.

### 3.4 Air Temperature

There are two temperature properties that influenced the solar radiation value; air temperature and surface temperature. Air temperature is one of the main input parameters for solar radiation prediction. It contains the information of temperature which is directly proportional to the solar radiation value on the particular places on the earth’s surface. As solar radiation supplies heat and energy toward the earth, air temperature is the product of heat generated in atmosphere due to the impact of solar radiation energy from the sun.

Air temperature referred to the ambient temperature and measured in Celsius (°C) or Fahrenheit (°F). Air temperature is an important factor that influences the growth and reproduction of living organisms on earth [37]. It also affected the evaporation rate, humidity, and being used as a prescript pattern for weather forecasts [38].

Air consists of gas molecules. This gas molecule travels at high speeds and is constantly moving. The energy of the molecules is directly related to the value of the air temperature. The thermometer is the tool used to measure air temperature. Common thermometer having a glass tube with a liquid that is directly influenced by the heat. As the heat increases, the liquid in the tube of the thermometer will expand thus indicates the temperature level.

Surface temperature is the measured temperature directly on the solar panel. It is also considered as one of the parameters in determining the lifespan of the solar panel. The heat generates can reduce the efficiency of the solar panel [39].

### 3.5 Relative Humidity

Besides air temperature, another parameter related to heat is humidity. Humidity is defined as the amount of moisture in the air. Relative humidity refers to the ratio between current absolute humidity and the maximum absolute humidity based on current air temperature [40]. This relative humidity is also used in weather forecasts to predict rain. If the percentage of relative humidity is 100%, it is shown that the air is saturated with water vapor thus leading to possibilities of rain [41].

Relative humidity gives information on the volume of water vapor existing in the air and expressed in a ratio form. It is constantly changing and relies on other factors such as temperature, dew point, and air saturation. It also shows the percentage of humidity saturation and generally calculated using equation (8):

\[ \text{Relative Humidity} = \frac{D_{actual}}{D_{saturation}} \times 100\% \] (8)

Where \( D_{actual} \) referred to actual vapor density and \( D_{saturation} \) referred to saturation vapor density measured in g/m³.

The relative humidity is measured by an instrument called a hygrometer. Hygrometers measure the moisture in the air by relying on measurement of other gear such as temperature detector, pressure detector, and mass detector. These three components will be calibrated together to determine the measurement of the humidity. There are many types of hygrometers that had been developed and applied in humidity measurements such as psychrometer, sling psychrometer, chilled mirror dew point hygrometer, and modern hygrometer such as capacitive, resistive, thermal, gravimetric, and optical hygrometer.
The relative humidity is frequently used as an input parameter for solar radiation prediction model development. It is because it can be retrieved from the meteorological station which is used the readings for a weather forecast.

3.6 Surface Pressure

Surface pressure is defined as the atmospheric pressure at a specific location on the surface of the earth [42]. Surface pressure is directly proportional to the mass of the air on that specific area as shown in equation (9).

$$P = \frac{F}{A}$$  \hspace{1cm} (9)

Where $P$ is denotes as the surface pressure, $F$ is the force being applied to the dedicated surface and $A$ is the area. The force is measured in Newton, N thus equivalent to the mass of the air being pressed to the surface of the earth within a specific range of area.

The barometer is the specific device used to measure atmospheric pressure in Pascal, Pa unit. It is recorded that the average value of surface pressure on earth is approximately 98,500 Pa. The mean sea level pressure involves the extrapolation of pressure on the sea level towards the location above or below sea level in which approximately at 101,325 Pa or equivalent to 1 atmosphere (atm).

Surface pressure is not often used as an input parameter for solar radiation because it is directly related to the atmospheric with the gravitational impact of earth. But since it gives an insight view of the atmospheric condition, it also can be used to determine the scattered radiation which happens in the atmosphere layer of earth.

3.7 Cloud Opacity

Cloud opacity measures how impenetrable the clouds are. This opaque level of the cloud determines the percentage of solar radiation that can penetrate the surface of the earth [43]. To ease the understanding of the cloud coverage on the sky, classification has been made to determine its opacity which are the clear sky, optically thin clouds, and optically thick clouds [44].

A clear sky or transparent sky shows that there is a thin cloud that allows the solar radiation to pass without any major losses. Optically thin clouds or translucent sky shows that there are medium-thickness clouds in the sky which still allows the solar radiation to pass but with a minor loss. Optically thick clouds or opaque sky shows the thick clouds in the sky which block the solar radiation which made the radiation diffuse and scattered above these clouds. During this condition, the sun’s location is difficult to detect.

Satellite imagery had been applied to observe the properties of the cloud such as optical depth of the cloud and the fraction of the cloud [45]. There is also a method to detect the cloud opacity level using ground-based sky imagery [46]. The similarity between these two imagery techniques is to determine the cloud cover by using a mapping technique. The cloud opacity is determined by the cloud appearances in the single window frame as shown in Figure 8.

Figure 8 Cloud opacity classification based on small-scale approximation of (a) clear sky model, (b) optically thin cloud model, and (c) optically thick cloud model.

To determine the cloud cover index, an algorithm was developed to distinguish the cloud cover pattern using the image processing method. This approach eases the determination of the cloud cover by simply comparing the detection image with the image based on the three models; clear sky, optically thin cloud, and optically thick cloud. Because of cloud opacity related to the total solar radiation received on the earth’s surface, it is also one of the important parameters to be considered in solar radiation prediction model development.
4.0 NEURAL NETWORK APPROACH ON SOLAR RADIATION PREDICTION MODEL DEVELOPMENT

Neural Network is a method which used a mathematical function to solve a various problem which inspired by the functional analogy of human brain [47] as shown in Figure 8. It was first introduced by a neurobiologist, McCulloch in 1943 [48] which illustrates the concept of neurons and the brain in the decision-making process based on the input signals. It is suggested that the interaction towards finding the solution is based on a familiarization process called training. Later in 1958, psychologist, Frank Rosenblatt presents the first artificial neural network called perceptron [49].

![Figure 9 Illustration of neurons in biological brain](image)

Based on the neuron architecture in Figure 9, the inputs receive the signal in terms of information through synapses. The dendrites sum up all information received from the synapses and send this as pulses through the axon to the next process. The architecture of Neural Network mimics the process by using artificial neurons in interconnected order using weight connectors [50]. These artificial neurons calculate the weighted sum of the inputs and perform calculations for the output using the activation function. Figure 10 shows the block diagram of the artificial neuron.

![Figure 10 Block diagram of artificial neurons](image)

Based on Figure 10, there are four basic components of a Neural Network consists of synapses, adder, activation function, and bias. Synapses are the combination of the input signal, $x_i$ with the product of neuron connection of $k$ multiply by synaptic weight, $w_{ki}$. Adder functions to perform the additional calculation for the weighted inputs. The activation function produces the output of the neuron in the range of [0, 1]. The bias, $b_k$ is responsible in increased or decreased the net input of the activation function to keep the input in the range stated. The output of the neuron $k$ can be described as equation (10):

$$y_k = \varphi \left( \sum_{i=1}^{m} x_i \cdot w_{ki} + b_k \right) \quad (10)$$

Where $x_i$ is the input signal with the maximum value of $m$, $w_{ki}$ is the respective weight of neuron to the input signal, $b_k$ is the bias, and $\varphi$ is the activation function. The effect of the bias can be seen in the two-stage process where the first stage includes the weighted inputs and the sum denoted as $S_k$ as shown in equation (11) and the output of adder denoted as $v_k$ as shown in equation (12).

$$S_k = \sum_{i=1}^{m} x_i \cdot w_{ki} \quad (11)$$

$$v_k = S_k + b_k \quad (12)$$

By using equation (11) and (12), equation (10) can be simplified as equation (13):

$$y_k = \varphi(v_k) \quad (13)$$

Bias also can be considered as an input signal, $x_0$ fixed at +1 with synaptic weight equal to the bias, $b_k$ as shown in Figure 11.

![Figure 11 Neurons structure with bias](image)

There is two structure of Neural Network which are single-layer feedforward and multi-layer feedforward [51]. Single-layer feedforward neural network is the simplest structure which consists of input nodes connected directly with the neurons and the output based on activation function as shown in Figure 12 and presented in equation (14).

$$\begin{bmatrix} y_1, y_2, ..., y_k \end{bmatrix} = [x_1, x_2, ..., x_l] \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1k} \\ w_{21} & w_{22} & \cdots & w_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ w_{l1} & w_{l2} & \cdots & w_{lk} \end{bmatrix} \quad (14)$$
With the inputs presented as vectors with $1 \times i$ dimension, weights presented as a matrix with $i \times k$ dimension, and outputs presented as a vector with $1 \times k$ dimension.

The weight update in Neural Network using a backpropagation algorithm [54]. This algorithm determines the error at the output and then propagates the error back into the network. This process continues to execute to minimize the error using a computing gradient.

In time series prediction, the values for input and output both will be as a function of time. Neural Network can perform future value prediction in such ways of:

i. Future value prediction based on past values of the same time series as shown in equation (15):

\[ y'(t) = F(y(t), y(t-1), y(t-2), ...) \]  \hspace{1cm} (15)

ii. Future value prediction based on value of relevant time series as shown in equation (16):

\[ y'(t) = F(y(t), x(t), x(t-1), x(t-2), ...) \]  \hspace{1cm} (16)

iii. Future value prediction based on both past value and value of relevant time series as shown in equation (17):

\[ y'(t) = F(y(t), x(t), x(t-1), x(t-2), ..., y(t-1), y(t-2), ...) \]  \hspace{1cm} (17)

Based on a previous study, it is shown that most of the studies conducted on solar radiation prediction based on several parameters such as irradiation data [55], sunshine duration [56], temperature [57], and weather record data [58].

To achieved high-performance prediction on solar radiation, all parameter inputs that may affect the solar radiation readings should be taken into account when executing the forecast.

The prediction approach on solar radiation involves a climatological parameter to develop and estimate the global solar radiation and diffuse solar radiation. Common findings on prediction model developed to estimate global solar radiation on a horizontal surface by applying various input parameters such as temperature, relative humidity, sunshine duration, and targeted location; latitude and longitude. There are many studies that carried out the Neural Network approach to predict global solar radiation based on horizontal surfaces [59].

Angstrom was the first to develop a prediction model of global solar radiation based on sunshine duration. Sunshine duration is based on the timeline of sunrise until sunset [60]. It is based on the location because sunshine duration differs in each country. This prediction model based on sunshine duration had been applied by Prescott to calculate a monthly average based on daily global solar radiation on a horizontal surface [61]. The estimation of solar radiation is achieved by using a quadratic form of relationship between daily global solar radiation and actual sunshine duration [62].

Prediction modeling for solar radiation using Neural Network had always in form of a combination of several methods to improve their results and to reduce training, only input data is provided and the predicted output is considered as the actual output.
the complexity of the prediction model. The are several methods of Neural Network used for solar radiation prediction based on the paper from R. Meenal and A.J. Selvakumar where they used Artificial Neural Network in predicting global solar radiation for India [63]. The input parameter for this model is based on monthly average radiation such as sunshine duration, temperature, wind speed, atmospheric pressure, relative humidity, and other meteorological parameters for the specific location. S.M. Mousavi and et al. [64] used a hybrid neural network together with a simulated annealing method to develop a prediction model for solar radiation. Ojo and Adeyemi [65] applied Neural Network Autoregressive Model with Exogenous Input (NNARX) to estimate solar radiation based on various climate parameters such as minimum and maximum air temperature, wind speed, and relative humidity.

A.E. Gurel et al. [66] develop a solar radiation prediction model using surface pressure, relative humidity, wind speed, ambient temperature, and sunshine duration for four Turkey Province of Ankara, Karaman, Killis, and Sirnak. They use the data from 2008 to 2018 to predict the data for the monthly average daily global solar radiation for 2018. In the end, the prediction model developed using these parameters manages to obtain a coefficient of determination, $R^2$ value between 0.952 and 0.993. B. Belmahdi et al. [67] develop a solar radiation prediction model using solar irradiation, temperature, precipitation, humidity, and wind speed for the duration of November 2018 to March 2019 in Morocco. They manage to produce output with a low Mean Bias Error (MBE) of -0.0886.

M. Blal et al. [68] develop a solar radiation prediction model using ambient temperature with different weather conditions for Adrar, Algeria. The data used for this model is based on temperature sensor readings from 2009 to 2012 obtained from National Aeronautical and Space Administration (NASA). In addition to the weather condition, they also add in the impact of the dust in the desert on solar radiation values. By the end of the research, the prediction model developed using these parameters manage to obtain an $R^2$ value of 0.87.

J. Fan et al. [69] develop a daily solar radiation prediction model using daily global solar radiation, sunshine duration, temperature, and relative humidity using the data from 1996 to 2015 from 15 weather stations across the various climatic regions of China. Based on the results, it is shown that more input parameters given more accurate results. Another reference from J. Fan et al. [69] also indicates that the air pollutant had given significant effect on the proportion of diffuse and global solar radiation. The prediction model mainly used sunshine duration as an input parameter and was also followed by Y. Feng et al. [70] that using sunshine duration and temperature value from 96 weather stations of China in the duration of 1994 to 2016.

J. Lee et al. [71] develop a solar irradiance prediction model using global horizontal irradiance, total sky cover, dry-bulb temperature, dew-point temperature, relative humidity, wind speed, and visibility based on the distance to discernable remote objects. This one-year range of data obtained from the National Climatic Data Center for the United States of America. The prediction model develops manage to obtain the best $R^2$ value of 0.996. J. Chakchak and N.S. Cetin [72] investigate the impact of various meteorological parameters such as cloud cover, sunshine duration, clearness index, and diffuse fraction on solar radiation prediction results. They concentrate on the cloud cover condition as the main parameter for the prediction model development. Based from the results, it is shown that more input parameter develops more precise prediction results.

5.0 METHODOLOGY

Solar radiation prediction requires a lot of input data to be properly adjusted and aligned to generate precise and accurate values. Since there is a need for geographical and meteorological data, the application of GIS-aided data will ease the data collection and acquisition. The list of input parameters used in the prediction is shown in Table 1. The data used for prediction modeling is from 2017-2019. The result shows a prediction of the solar radiation value of Malaysia for 2019.

Table 1 List of Input Parameter

| Parameters                   | Unit | Source   |
|------------------------------|------|----------|
| GHI                          | Wm^-2| SOLCAST  |
| DNI                          | Wm^-2| SOLCAST  |
| DHI                          | Wm^-2| SOLCAST  |
| Solar Azimuth Angle          | °    | SOLCAST  |
| Solar Zenith Angle           | °    | SOLCAST  |
| Wind Speed                   | m/s^1| SOLCAST  |
| Air Temperature              | °C   | SOLCAST  |
| Relative Humidity            | %    | SOLCAST  |
| Surface Pressure             | Pa   | SOLCAST  |
| Cloud Opacity                | %    | SOLCAST  |

Based on a previous study, solar radiation prediction had widely applied geographical and meteorological data as input variables for the neural network prediction model [73-75]. The transfer function helps to denormalized the input data before it proceeds with the training phase for the neural network. There are two basic transfer functions which are the step function and a logistic sigmoid transfer function. Since solar radiation involves non-linear data, the logistic sigmoid transfer function had been widely used. The input parameter is divided into three parts consists of the training, validation, and testing phase using the 40:40:20 division. This will indicate 40% of the data for training, 40% of the data for validation, and remaining 20% of the data for testing. Neural
Network is all about the training. More training will help in better prediction results but it requires a lot of data. Based on a previous study, they use up to 5 years of data input for the prediction model development. The common flowchart for the Neural Network process is shown in Figure 14.

![Flowchart](image)

**Figure 14** Methodology flowchart for the Neural Network

### 6.0 RESULTS & DISCUSSIONS

The prediction model of NNARX is developed using MATLAB R2020a software with Neural Network toolbox using the process as shown in Figure 14. The results of the prediction are shown in Figure 15.

![Graph](image)

**Figure 15** Results of solar radiation prediction using NNARX

Figure 15 shows the solar radiation value based on the weekly average value. The accuracy of the NNARX is calculated using the coefficient of determination, \( R^2 \) as shown in equation (18) below:

\[
R^2 = \left( \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \right)^2
\]  

(18)

Where \( n \) is the total number of data, \( x \) is the actual output value, and \( y \) is the predicted output value. Based on the output value, the NNARX manages to obtain an \( R^2 \) value of 0.9329.

The solar radiation mapping had been done using the entire related parameters to develop a comprehensive view of the solar energy available on that specific location based on satellite-based data as shown in Figure 16 and Figure 17.

![Map](image)

**Figure 16** Solar radiation mapping for Peninsular Malaysia

![Map](image)

**Figure 17** Solar radiation mapping for East Malaysia

Based on Figure 16 and Figure 17, the dark-coloured contour represents the most potential solar radiation values measured giving a high potential for solar energy harvesting compared with the bright-coloured contour present on the map. To show the results on the actual map, there is a need for a guide chart. This will translate the finding in measured value into coloured-tone legend as shown in Figure 18.
Solar energy is discovered as one of the reliable and safest renewable energy sources, the study had been made to uncover its potential to bring a maximum energy output from the photovoltaic system. Since there is an advancement had been achieved in prediction modeling development using a various methods including Neural Network, a conclusive solar radiation modeling can be achieved based on place of interest.

The GIS-aided data simplifies the data collection and measurement for a better data acquisition process to mapping a solar radiation map for Malaysia. The GIS-aided data also reduces the correlation error between these parameters. It is also identified that the height of the places from sea level had given an impact on the solar energy collection. The high-altitude places tend to give more solar energy as compared with the low-altitude places.

Global solar irradiance is the combination of direct normal irradiance diffuse horizontal irradiance, and ground reflected irradiance. Since there is a huge difference between horizontal and surface and the tilted surface of this solar irradiance parameter, it is best to identify this irradiance scenario before delivering the solar radiation measurement. The prediction of solar radiation is normally done by using the data from other places and using the same training model to develop a prediction on the specific location with their input parameter.

Since Malaysia is a tropical climate country, it receives more than half of the day in one year. So, the cloud opacity for Malaysia had given a big impact on the total solar energy received throughout the year. By identifying the maximum potential for solar energy harvesting, this climate issue will be still in control. This goes the same with the relative humidity. But the relative humidity will not affect the solar energy potential in a big number since the average humidity for the country is at par and balanced with the air temperature for the surrounding country.

The study based on a prediction of solar radiation using the Neural Network had brought a discovery on the new method and implementation of several enhancements had been applied to further enhance the prediction model accuracy. This prediction model is needed to enlarge the scale of the solar radiation values in specific locations. These data are needed to maximize the potential solar energy that can be retrieved from that area. Since the prediction is based on geographical and meteorological parameters, the GIS-aided data will ease the system as it gives the cumulative response from all the sensors at one data center.

7.0 CONCLUSION

This paper had presented the GIS-based data which had been used as an input parameter for solar radiation prediction model development. GIS consist of both geographical and meteorological data which is the influential parameter to determine the solar radiation value on the surface of the earth. This solar radiation value is used to estimate the solar energy potential that can be produced at a specific location in Malaysia. Since the data is driven by satellite equipment and gives real-time data for the overall solar energy potential based on a large-scale, prediction model development will help in narrowing the specific location to develop a solar radiation prediction model for the solar energy potential which is exclusive on a specific area. Neural Network is one of the Machine Learning methods that is widely used to develop a prediction model based on dynamic non-linear time-series data. Based on the review, it is found that Neural Network had the ability to produce precise prediction results with the consideration of the entire parameter highlighted in this paper. Based on the prediction results using NNARX, it is shown that Neural Network manages to obtain prediction performance with the R² value of 0.9329. For a full-scale solar radiation mapping, there is a need for a detailed and large number of data that can be retrieved from the GIS database. The findings in this paper are beneficial for a further recommendation on developing a solar radiation prediction model using Neural Network.

Acknowledgement

The authors would like to thanks and acknowledge the Research Management Institute (RMI), Innovative Electromobility Research Laboratory (ITEM), and School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA Shah Alam for their support.

References

[1] W. S. Wan Abdullah, M. Osman, M. Z. A. Ab Kadir and R. Verayiah. 2019. The Potential and Status of Renewable Energy Development in Malaysia. Energies. 12. DOI: https://doi.org/10.3390/en12122437.

[2] W. Linhua and S. Jiarong. 2021. A Comprehensive Application of Machine Learning Techniques for Short-Term Solar Radiation Prediction. Applied Sciences. 11(13), DOI: https://doi.org/10.3390/app11135808.

[3] D. Gujo-Rubio, A. M. Duran-Rosales, P. A. Gutierrez, A. M. Gomez-Orellana, C. Casanova-Mateo, J. Sanz-Justo, S. Salcedo-Sanz, C. Hervas-Martinez. 2020. Evolutionary Artificial Neural Networks for Accurate Solar Radiation Prediction. Energy. 210. DOI: https://doi.org/10.1016/j.energy.2020.118374.

[4] J. M. Barrera, A. Reina, A. Mate and J. C. Trujillo. 2020. Solar Energy Prediction Model Based on Artificial Neural Networks and Open Data. Sustainability. 12. DOI: https://doi.org/10.3390/su12176915.
[67] B. Belmahdi, M. Louzazni and A. E. Bouardi. 2020. One Month-ahead Forecasting of Mean Daily Global Solar Radiation using Time Series Models. Optik. 219. DOI: https://doi.org/10.1016/j.ijleo.2020.165207.

[68] J. Fan, et al. 2020. Predicting Daily Diffuse Horizontal Solar Radiation in Various Climatic Regions of China using Support Vector Machine and Tree-based Soft Computing Models with Local and Extrinsic Climatic Data. Journal of Cleaner Production. 248. DOI: https://doi.org/10.1016/j.jclepro.2019.119264.

[69] J. Fan, et al. 2020. Hybrid Support Vector Machines with Heuristic Algorithms for Prediction of Daily Diffuse Solar Radiation in Air-Polluted Regions. Renewable Energy. 145: 2034-2045. DOI: https://doi.org/10.1016/j.renene.2019.07.104.

[70] Y. Feng, et al. 2020. National-scale Development and Calibration of Empirical Models for Predicting Daily Global Solar Radiation in China. Energy Conversion and Management. 203. DOI: https://doi.org/10.1016/j.enconman.2019.112236.

[71] J. Lee, W. Wang, F. Harrou and Y. Sun. 2020. Reliable Solar Irradiance Prediction using Ensemble Learning-based Models: A Comparative Study. Energy Conversion and Management. 208. DOI: https://doi.org/10.1016/j.enconman.2020.112582.

[72] J. Chakchak and N. S. Cetin. 2021. Investigating the Impact of Weather Parameters Selection on the Prediction of Solar Radiation Under Different Genera of Cloud Cover: A Case Study in A Subtropical Location. Measurement. 176. DOI: https://doi.org/10.1016/j.measurement.2021.109159.

[73] M. R. S. Mohd, J. Johari and F. A. Ruslan. 2020. Application of NNARX in Modeling a Solar Radiation Prediction. 2020 IEEE 8th Conference on Systems, Process and Control (ICSPC), Melaka, Malaysia.

[74] E. A. Enríquez-Velásquez, V. H. Benitez, S. G. Obukhov, L. C. Félix-Herrán, and J. de-J. Lozoya-Santos. 2020. Estimation of Solar Resource Based on Meteorological and Geographical Data: Sonora State in Northwestern Territory of Mexico as Case Study. Energies. 13. DOI: https://doi.org/10.3390/en13246501.

[75] Q. Zhang, N. Cui, Y. Feng, Y. Jia, Z. Li, and D. Gong. 2018. Comparative Analysis of Global Solar Radiation Models in Different Regions of China. Advances in Meteorology. 2018. DOI: https://doi.org/10.1155/2018/3894831.