Hourly solar radiation estimation from limited meteorological data to complete missing solar radiation data

Abstract—This paper presents two methods to estimate hourly solar radiation from available weather data. The two methods presented in this paper are based on atmospheric transmittance determination using available meteorological data. In the first method a decision matrix was used while in the second method regression correlation of meteorological parameter was used. The calculations results were evaluated using statistical parameter. Though the result shows both of the methods perform well, more satisfactory results were obtained from first method with Root Mean Square Error of 87.6 Watt/m², Normalized Root Mean Square Error of 8.29%, correlation coefficient of 0.95 and index of agreement of 0.97. Furthermore, the first method only need RH and ambient temperature data which commonly measured in every meteorological station.

Keywords: Hourly solar radiation; meteorological data; missing data estimation

I. INTRODUCTION

Time series solar radiation data is important for modeling and design of solar radiation related devices or systems, such as photovoltaic (PV) system. The PV systems performance is greatly affected by daily variation of the solar radiation. Hence, complete time series data is significant for performance prediction to ensure reliability of the system. However, this solar radiation data are not always available for every area, even if there is a weather station near the area, the data access often limited. Furthermore, the data available data may contain missing data for several days in the absence of measurement. Sometimes this missing data only occurred for some parameter in meteorological data set, while other parameters are complete data. This might happen due to sensor error or damage. Therefore, missing data estimation is required in order to utilize the data set for various purposes.

Meteorological parameters such as ambient temperature, sunshine duration and cloud cover has been used and evaluated to estimate solar radiation [1, 2, 3 and 4]. Satisfactory result was obtained by using atmospheric transmittance model [1] while other authors have used diffuse fraction or clearness index [5, 6]. Some studies also added measured precipitation to temperature based hourly solar radiation prediction and the methods claimed to perform well [7, 8].

Not all of hourly solar radiation estimation methods can be performed due to the non-availability of the data in the area. In this paper missing data of solar radiation was estimated using two methods, first, beam atmospheric transmission determination with measured RH and ambient temperature data and second method using RH-beam transmittance correlation through clearness index-beam transmittance correlation. Several authors proposed clearness index-beam transmittance numerical correlation based on data in certain areas [9, 10]. In this paper clearness index-beam transmittance numerical correlation was proposed using measured data in Universiti Teknologi PETRONAS, Bandar Sri Iskandar.

Statistical parameters were used to validate the estimation results which can be used to fill missing data in the data set. Then complete hourly time series of solar radiation data set can be used for any purpose such as PV system performance prediction.

II. DATA SET

In this research Ipoh city weather data was used as a study case. The data contain 5 (five) parameters that recorded hourly through year 2003. The parameters are solar radiation, ambient temperature, relative humidity, speed and direction of the wind. In the data set, there are 23 days missing solar radiation data while other parameters were complete data set.

Figure 1: One year hourly solar radiation data with 23 days missing data
G beam radiation (sun can be separated into two basic components: direct

Firstly, in this method total radiation energy from the sun can be separated into two basic components: direct beam radiation \((G_b)\) and diffuse solar radiation \((G_d)\). The sum of these two results in the total incident solar radiation \((G_{th})\) and is represented by:

\[
G_{th} = G_b + G_d
\]  

(1)

The local intensity of solar beam radiation is determined by the angle between the direction of the sun’s rays and the earth’s surface. The location of the sun is given by the angle between the sun location and the normal to the surface, referred to as the zenith angle \((\theta_z)\). Zenith angles vary temporally and geographically but are a function of the time of day, latitude, and time of year given by the following relationship [4]:

\[
\cos (\theta_z) = \sin (\phi) \sin (\delta) + \cos (\phi) \cos (\delta) \cos (\omega)
\]  

(2)

To calculate horizontal beam radiation, the following equation was used:

\[
G_b = G_{pa} \cos (\theta_i)
\]  

(3)

\(G_{pa}\) is beam radiation received on perpendicular surface to incoming radiation. The model chosen for the \(G_{pa}\) is from Liu and Jordan (1960) [12], where the beam radiation \((G_{ba})\) is given by:

\[
G_{ba} = G_{sh} \tau^m
\]  

(4)

\(G_{sh}\) is the solar constant \((1,360 \text{ W/m}^2)\), \(\tau\) is the atmospheric transmittance, and \(m\) is the optical air mass number. The optical mass number \((m)\) is found from the following relationship [1]:

\[
m = P_a/101.3 \cos (\theta_j)
\]  

(5)

\(P_a\) being the atmospheric pressure \((\text{kPa})\) at the site and the zenith angle from Equation 3.40. Average barometric pressure was estimated from the relationship[1]:

\[
P_a = 101.3 e^{a\theta/8200}
\]  

(6)

\(a\) is the elevation of the site \((\text{meters})\).

However, not all of the beam radiation reaches the earth’s surface. Radiation is reflected or absorbed by atmospheric gases, clouds, and dust particles. Some of this radiation is scattered toward earth and is referred to as diffuse radiation \((G_d)\). Campbell and Norman [1] devised an empirical relationship based on work of Liu and Jordan [12] for an estimation of diffuse radiation. This relationship is given by:

\[
G_d = 0.30 (1- \tau^m) G_{sh} \cos (\theta_j)
\]  

(7)

III. MISSSING DATA ESTIMATION APPROACH

There are various ways to estimate solar radiation on certain area on the earth. Ambient temperature based estimation is widely used since ambient temperature data are measured in many weather stations. In this study, missing data were estimated based on ambient temperature measurement and used measured RH data as atmospheric transmittance determination criteria. The procedures developed by Kurt and Spokas method [6] which estimate hourly solar radiation based on developed Campbell and Norman method [1], was adapted in this study.

Firstly, in this method total radiation energy from the sun can be separated into two basic components: direct beam radiation \((G_b)\) and diffuse solar radiation \((G_d)\). The sum of these two results in the total incident solar radiation \((G_{th})\) and is represented by:

\[
G_{th} = G_b + G_d
\]  

(1)

The local intensity of solar beam radiation is determined by the angle between the direction of the sun’s rays and the earth’s surface. The location of the sun is given by the angle between the sun location and the normal to the surface, referred to as the zenith angle \((\theta_z)\). Zenith angles vary temporally and geographically but are a function of the time of day, latitude, and time of year given by the following relationship [4]:

\[
\cos (\theta_z) = \sin (\phi) \sin (\delta) + \cos (\phi) \cos (\delta) \cos (\omega)
\]  

(2)

To calculate horizontal beam radiation, the following equation was used:

\[
G_b = G_{pa} \cos (\theta_i)
\]  

(3)

\(G_{pa}\) is beam radiation received on perpendicular surface to incoming radiation. The model chosen for the \(G_{pa}\) is from Liu and Jordan (1960) [12], where the beam radiation \((G_{ba})\) is given by:

\[
G_{ba} = G_{sh} \tau^m
\]  

(4)

\(G_{sh}\) is the solar constant \((1,360 \text{ W/m}^2)\), \(\tau\) is the atmospheric transmittance, and \(m\) is the optical air mass number. The optical mass number \((m)\) is found from the following relationship [1]:

\[
m = P_a/101.3 \cos (\theta_j)
\]  

(5)

\(P_a\) being the atmospheric pressure \((\text{kPa})\) at the site and the zenith angle from Equation 3.40. Average barometric pressure was estimated from the relationship[1]:

\[
P_a = 101.3 e^{a\theta/8200}
\]  

(6)

\(a\) is the elevation of the site \((\text{meters})\).

However, not all of the beam radiation reaches the earth’s surface. Radiation is reflected or absorbed by atmospheric gases, clouds, and dust particles. Some of this radiation is scattered toward earth and is referred to as diffuse radiation \((G_d)\). Campbell and Norman [1] devised an empirical relationship based on work of Liu and Jordan [12] for an estimation of diffuse radiation. This relationship is given by:

\[
G_d = 0.30 (1- \tau^m) G_{sh} \cos (\theta_j)
\]  

(7)

IV. ATMOSPHERIC TRANSMITTANCE ASSIGNMENT

PROCEDURE

The key for the accuracy of above method is in the determination of beam atmospheric transmittance \((\tau)\). Beam atmospheric transmittance is the percentage of the beam (direct) radiation that will penetrate the atmosphere without being scattered. Kurt and Spokas [5] used precipitation data to built decision matrix of atmospheric transmittance.

Gueymard C. [13] stated that if precipitation data were not available and the value estimated from ground data of temperature and humidity, separate tests have revealed that the instantaneous error in precipitation may will be >100%, depending on atmospheric conditions and on the ‘universality’ of the empirical model used. Therefore, although we do not have precipitation data, in this study we do not intend to predict precipitation value by using available temperature and humidity data. In this study two methods of beam atmospheric transmittance assignment procedure were proposed as follows:

A. Method 1

The first method is using the same approach of Kurt and Spokas by built a decision matrix. The difference is the beam transmittance assignment in this matrix is controlled by Relative humidity value. Better estimation was obtained with the concept that water vapor by mean of RH reduce the incoming radiation. Kurt and Spokas (2007) suggest the value of \(\tau\) was modified if \(\Delta T<10^\circ\text{C}\) by the following relationship assumed that the site was not near the poles as described by [5]:

\[
\tau' = \tau / (11- \Delta T)
\]  

(8)
In first method limitation of \( \Delta T \) was adjusted, above equation was used for \( \Delta T<8^\circ \text{C} \). Table 1 shows criteria for the decision of \( \tau \) value. \( \tau \) value of 0.6-0.7 are commonly used for clear sky atmospheric transmittance coefficient value. In this study \( \tau \) value of 0.69 was used for clear sky, assumed that the clear sky condition occurred when RH<40% and ambient temperature more than 8\(^{\circ}\)C. Calculation algorithm was built based on decision matrix and the \( \tau \) value was locally determined using the training of data set to get minimum error.

**Table 1. Atmospheric transmittance coefficient determination using measured RH**

<table>
<thead>
<tr>
<th>No</th>
<th>RH condition (%)</th>
<th>( \tau ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RH&lt;40</td>
<td>0.69</td>
</tr>
<tr>
<td>2</td>
<td>40&lt;RH&lt;45</td>
<td>0.67</td>
</tr>
<tr>
<td>3</td>
<td>45&lt;RH&lt;55</td>
<td>0.57</td>
</tr>
<tr>
<td>4</td>
<td>55&lt;RH&lt;65</td>
<td>0.47</td>
</tr>
<tr>
<td>5</td>
<td>65&lt;RH&lt;75</td>
<td>0.41</td>
</tr>
<tr>
<td>6</td>
<td>75&lt;RH&lt;80</td>
<td>0.3</td>
</tr>
<tr>
<td>7</td>
<td>RH&gt;80</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**B. Method 2**

The second method used in this study is by finding the correlation between RH, clearness index and beam transmittance. The data used to find correlation between beam transmittance and clearness index is measured data from new radiometer set that was installed in 2010 on the rooftop of Block P, Universiti Teknologi Petronas, which located about 30 km from Ipoh city. About 1 month, 5 minutes time step data of global, beam and diffuse radiation from June to July 2010 was used. Before find the correlation of beam transmittance and clearness index, RH-clearness index correlation was obtained from Ipoh city available data as can be seen in Figure 3.

Then beam transmittance-clearness index correlation can be obtained by scatter plot as can be seen in Figure 4. To plot Figure 4 some data were rejected due to obvious error that can be analyzed from measurement results, and the basic concept of terrestrial solar radiation characteristics.

Following constraint were used as data rejection criteria:
- Reject night data
- Reject data if clearness index >1
- Reject data if beam transmittance >1
- Reject data if beam transmittance > clearness index
- Reject data if clearness index >0.6 and beam transmittance <0.1
- Reject data if clearness index <0.2 and beam transmittance>0.15

Correlation between RH and beam transmittance was obtained from above correlation and plotted in Figure 5. Balaras et al studied the relationship between beam transmittance and clearness index in Athens, Greece [10], the results of the study was adopted to carry out second method in this study. Regression results were presented as follows:

Linear : \( \tau = -0.019 \text{RH}+1.576 \)

Quadratic : \( \tau = 0.00075 \text{RH}^2-0.12676 \text{RH}+5.56 \)

Cubic : \( \tau = 0.0000072\text{RH}^3-0.008\text{RH}^2-0.016 \text{RH}+2.7 \)
Results of the new model then compared with the existing temperature-based solar radiation prediction model as follow:

a. H-S model

Hargreaves and Samani [14] conducted an initial study on using \( T_{\text{max}} \) and \( T_{\text{min}} \) to estimate solar radiation by the following equation:

\[
G_{\text{Th}} = K_r (T_{\text{max}} - T_{\text{min}})^{0.5} \quad \text{(9)}
\]

\( K_r \) is an empirical coefficient, which was recommended to be 0.16 for interior regions and 0.19 for coastal regions. In this study \( K_r \) was locally determined using training data set.

b. H-S-A model

Annandale et al [15] modified H-S model by introducing correction factor as follow:

\[
G_{\text{Th}} = K_r (1 + 2.7 \times 10^{-7} Z) (T_{\text{max}} - T_{\text{min}})^{0.5} \quad \text{(10)}
\]

\( Z \) is elevation in m and \( K_r \) was locally determined.

V. Statistical Analysis for Model Validation

Estimation results validated using statistical parameters. Pearson correlation coefficient was calculated as routine correlation indicator. Residual error was calculated using RMSE (Root Mean Square Error) and also presented in NRMSE (Normalized Root Mean Square Error) as follows:

\[
\text{RMSE} = \left[ \sum (Y_c - Y_0)^2 / n \right]^{0.5} \quad \text{(11)}
\]

\[
\text{NRMSE} = \text{RMSE} / y_{\text{max}} - y_{\text{min}} \quad \text{(12)}
\]

where, \( Y_c \) is predicted variable \( Y_0 \) is measured variable, \( n \) is number of data, \( y_{\text{max}} \) is maximum measured data \( y_{\text{min}} \) is minimum measured data.

As an addition, index of agreement was calculated using equation below:

\[
d = 1 - \left[ \sum (x_i - y_i)^2 / \sum (x_i - \bar{x}_i)^2 + (y_i - \bar{y}_i)^2 \right]^{0.5} \quad \text{(13)}
\]

where, \( x_i \) is predicted variable \( y_i \) is measured variable, \( \bar{x}_i \) is averaged predicted variable and \( \bar{y}_i \) is averaged measured variable.

VI. Results and Discussions

Calculation have been carried out using methods 1 and 2, statistical calculation analysis also has been performed. Table 2 shows statistical analysis results of both methods and results of existing method (H-S and H-S-A method).

The most satisfactory results were obtained using method 1. Figure 3 shows scatter plot of predicted and measured data for first method. The minimum RMSE value of 87.6 Watt/m\(^2\) was obtained with 0.95 correlation coefficient and 0.97 index of agreement value. Previous method which use precipitation data obtained averaged index of agreement of 0.95, thus the model presented in this study also performed well.
The prediction of hourly solar radiation data was carried out in this study based on two methods. First method is by using decision matrix from measured RH and ambient temperature data. The second method is by using RH-clearness index, clearness index-beam atmospheric transmission and beam atmospheric transmission-RH correlation. The result shows that both methods perform well. Method 1 provided better results with minimum correlation coefficient of 0.95, RMSE of 87.6 Watt/m², NRSME of 8.29% and index of agreement of 0.97. The prediction was intended to fill missing data in solar radiation data set to get complete time series data. However, in this study only one year of one area data have been used. Validation using sufficient large amount of data is required for wider application of the method.

ACKNOWLEDGMENT

The authors would like to thank Universiti Teknologi PETRONAS for providing facilities and grant under STIRF schemes No. 63/08.09 for the research.

REFERENCES