Available online at www.sciencedirect.com
SciVerse ScienceDirect
Energy Procedia 18 (2012) 1413 – 1420
Energy Procedia

An Estimation of Solar Radiation using Robust Linear Regression Method

S. Ibrahim[b], I. Daut[a], Y. M. Irwan[a*], M. Irwanto[a], N. Gomesh[a], A.R.N. Razliana[a]

[a]Cluster Power Electronic and Machine Design, School of Electrical System Engineering, University Malaysia Perlis (UniMAP), Malaysia
[b]Institute of Mathematical Science, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia

Abstract

The air temperature data is the most important component to estimate the solar radiation in photovoltaic systems. From the Malaysia Meteorological Department, the data of air temperature and solar radiation can be found the hourly, daily, monthly and also the annually. Based on solar radiation data for the past 26 years, the average monthly solar radiation was 5009.56 Wh/m^2. It was greater than the normal solar radiation (3 kWh/m^2), which indicates that the sky in Perlis was clear and very high solar radiation intensity for the months in the past 26 years. This paper presents an investigation of a relationship between solar radiation and temperature in Perlis, Northern Malaysia for the year of 2006. The Least Trimmed Squares (LTS) robust regression model was selected to estimate the solar radiation since the robust method is do not breakdown easily and are not much influenced by outliers.

Keywords: Solar radiation; air temperature; robust method;statistical analysis;

1. Introduction

1.1 Renewable Energy

Solar radiation is the result of fusion of atoms inside the sun. Part of the energy from the fusion process heats the chromosphere, the outer layer of the sun that is much cooler than the interior of the sun, and the radiation from the chromosphere becomes the solar radiation incident on the earth. The sun acts as a blackbody with a temperature around 6000 K [1], transports its vast energy to the earth in the form of electromagnetic radiation with a wide spectrum of frequency range that varies from infrared, visible lights to ultra-violet. The amount of solar power perpendicular to the beam outside the Earth's atmosphere is known as solar constant (S), approximately 1367 W/m^2 [1] [2] [3].

When the solar radiation enters the earth's atmosphere as Fig. 1, a part of the incident energy is removed by scattering or absorption by air molecules, clouds and particulate matter usually referred to as aerosols. The radiation that is not reflected or scattered and reaches the surface directly in line from the PV module is called beam radiation. The scattered radiation which reaches the ground is called diffuse radiation. The albedo is a radiation that reaches a receiver after reflection from the ground. The total solar radiation on a horizontal surface of PV module consisting three components is called global irradiance. When the skies are clear and the sun is directly in line from the PV module, the global irradiance is about 1000 W/m^2 [2]. Although the global irradiance on the surface of the earth can be as high as 1000 W/m^2, the available radiation is usually considerably lower than this maximum value due to the rotation on the

* Corresponding author. Tel.:+60104004700; fax: +604-9798903.
E-mail address: irwanyusoff@unimap.edu.my

1876-6102 © 2012 Published by Elsevier Ltd. Selection and/or peer review under responsibility of The TerraGreen Society.
Open access under CC BY-NC-ND license. doi:10.1016/j.egypro.2012.05.157

earth and climate condition (cloud cover), as well as by the general composition of the atmosphere. For this reason, the solar radiation data is the most important component to estimate output of photovoltaic systems [3] [4] [5]. Solar radiation is greater than 3 kWh/m^2 indicates that the sky is clear, its intensity very high and very good for PV application [6].

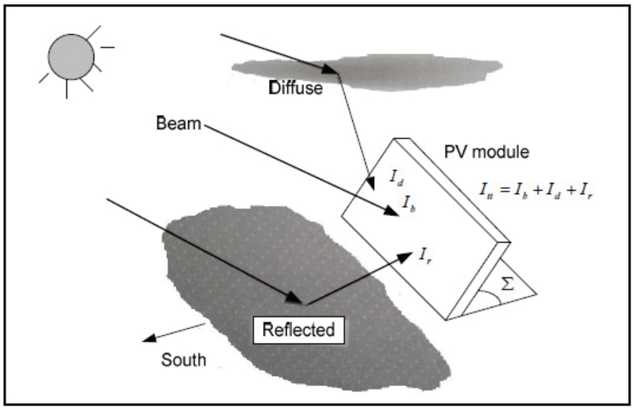

Fig. 1 Solar radiation in the earth's atmosphere

Temperature is an important consideration in the operation of PV module system [7]. At lower temperatures, PV module systems produce more power. For higher temperature, optimum operation requires modification of electrical load and removal of excess heat. At high temperatures, two predominating effects can cause efficiency to drop. As thermal energy increases, (1) lattice vibrations interface with the free passing of charge carries and (2) the junction begins to loss its power to separate charges. The efficiency losses for PV systems can be minimized in the presence of temperature variations. In most cases, good solutions are a temperature-dependent charge controller or a maximum power tracker. Both devices improve the overall system efficiency at higher temperature where the performance is poor [7].

1.2 Statistical Analysis

The statistical technique used in investigating and modelling the relationship between variables average solar radiation (y) and average air temperature (x) [8]. One of the most popular is using robust method. Robust fitting is commonly used when the data contain outliers. In the presence of outliers, least squares estimation is inefficient and can be affected by inaccuracy. This is due to the shifting of the least squares estimates towards the outliers and to the corresponding changed increase of the estimates variance.

2. Data and Methods

2.1 Climate and Solar Radiation in Perlis

Malaysia naturally has abundant sunshine and thus solar radiation. However, it is extremely rare to have a full day with completely clear sky even in periods of severe drought. The cloud cover cuts off a substantial amount of sunshine and thus solar radiation. On the average, Malaysia receives about 6 hours of sunshine per day. Solar radiation data can be got from Meteorological Station, Chuping Perlis. Unit of solar radiation is Wh/m^2 or J/m^2, to convert a quantity given in Wh/m^2 to J/m^2, it should be multiplied by 3600 [3]. PSHs is the length of time in hours at a radiation level of 1000 W/m^2 needed to produce energy equivalent to the total energy in one day or it is ratio of solar radiation (Wh/m^2) to solar radiation level of 1000 W/m^2 [9]. The solar radiation and PSHs are needed to calculate a minimum size of PV module.

Based on Meteorological Station in Chuping Perlis, Perlis ($6^0\ 29^{'}$ N , $100^0\ 16^{'}$ E) has about 795 square kilometers land area, 0.24% of the total land area of Malaysia, with a population about 204450 people [10]. Perlis's climate is tropical monsoon. Its temperature is relatively uniform within the range of 21°C to 32°C throughout the year. During the months of January to April, the weather is generally dry and warm. Humidity is consistently high on the lowlands ranging between 82% to 86% per annum. The average rainfall per year is 2,032 mm to 2,540 mm and the wettest months are from May to December. In Fig. 2 as shown below, indicate that Peninsular Malaysia recorded 14 to 17.0 MJ/m^2 of daily solar radiation and Perlis has solar radiation above 12 MJ/m^2.

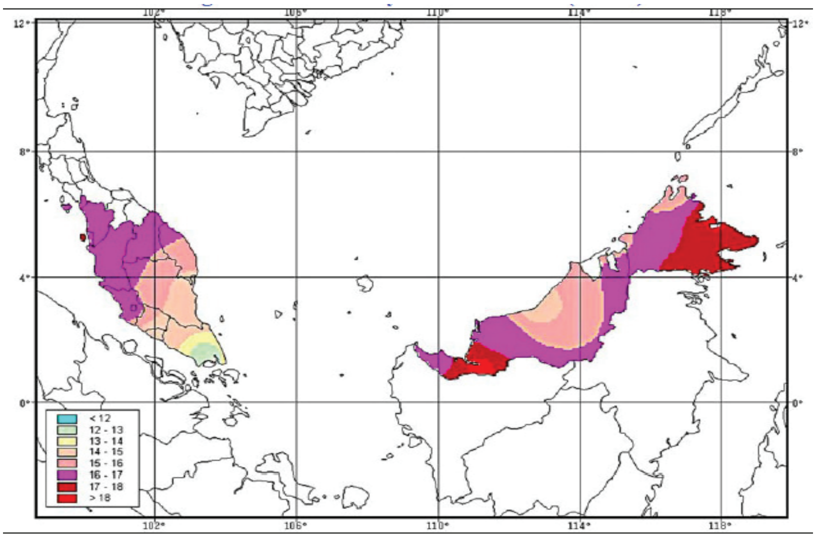

Fig. 2 Mean daily solar radiation (MJ/m^2)

2.2 *Least Squares Method*

Suppose that for any observation, $(x_1, y_1),(x_2, y_2),...,(x_n,y_n)$ be a pair of random variables. To predict y, the parameters β_0 and β_1 must be estimated, so that the sum of the square of the differences between the observations y_i and the straight line is minimum. The interpolating straight line as

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i\,, \qquad i=1, 2, ..., n \tag{1}$$

and the coefficients that minimize the square of the distance between the line end the points are given by:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \qquad (2)$$

$$\beta_1 = \frac{\sum_{i=1}^{n} y_i\, x_i - \frac{\left(\sum_{i=1}^{n} y_i\right)\left(\sum_{i=1}^{n} x_i\right)}{n}}{\sum_{i=1}^{n} x_i^{\,2} - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}$$

where

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \text{and} \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

are the averages of y_i and x_i, respectively. Therefore, β_0 and β_1 are the least squares estimators of the intercept and slope. The residuals ε are the differences between the observed and the predicted values $y_i - \hat{y}_i$, i=1, 2, ..., n.

The fitted simple linear regression model is given by

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (3)$$

2.3 Least Trimmed Squares Robust Regression

In regression analysis, the subject occurrence of outliers will be affecting the estimated coefficients, fitted values, residuals and covariance matrix of linear regression models [11][12]. Therefore, the robust statistic is a best way to estimate the solar radiation which do not breakdown easily and are not much influenced by outliers. The robust approach to statistical modeling and data analysis is produce reliable parameter estimates and associate tests [13]. One of the robust method is *Least trimmed (sum of) squares* (LTS) regression. This method is proposed by Rousseeuw (1984). It is a highly robust method for fitting a linear regression model. In this method, the estimation of β is obtained from $\sum_{i=1}^{h} \hat{\varepsilon}_i^2$ is minimized, where $\hat{\varepsilon}_i$ is the i^{th} ordered residual. h may depend on a trimming proportion of α, suggested choosing $h = [n(1-\alpha)] + 1$. Thus, LTS is equivalent to ordering the residuals from a least squares fit, trimming the observations that correspond to the largest residuals, and then computing a least squares regression model for the remaining observations.

4. Result and Discussion

4.1 Daily Solar Radiation

Fig. 3 describes the daily average solar radiation for the whole year of 2006. The highest total daily average solar radiation of 7238.89 Wh/m^2 was recorded on 7th March, and the lowest of 1130.56 Wh/m^2 was recorded on 20th May (as shown in Fig. 4). Daily average solar radiation values were high during the period of January to April. Average daily solar radiation for the whole year was 5031.45 Wh/m^2 per day and the annual total solar radiation in Perlis was 1831.45 kWh/m^2 per year.

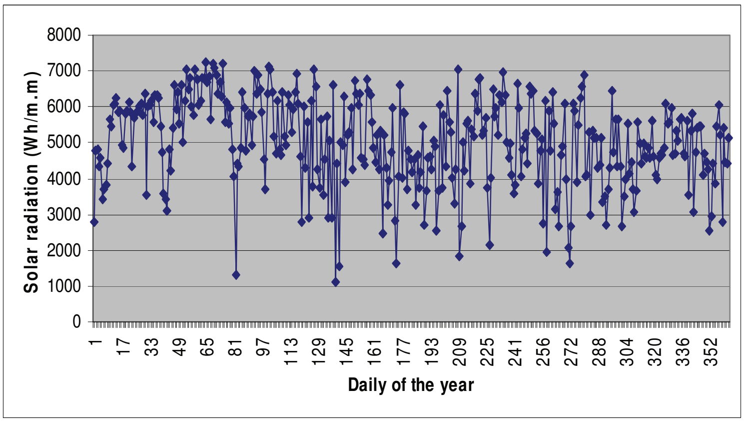

Fig. 3 Daily average of solar radiation throughout the year of 2006

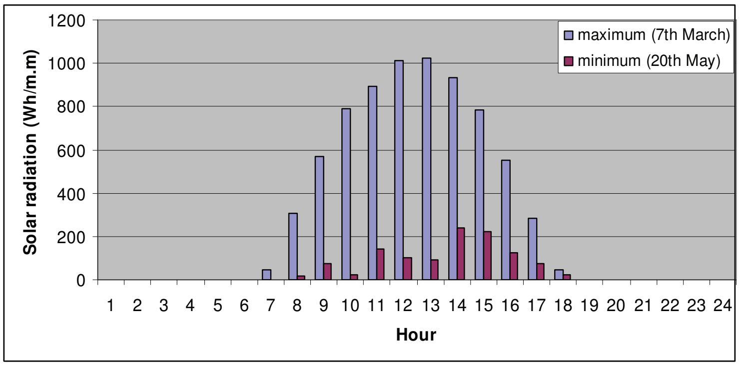

Fig. 4 Maximum and minimum hourly solar radiation in Perlis for the year of 2006

Fig. 5 shows monthly minimum, average and maximum solar radiation in Perlis for the year of 2006. Month of May had the lowest monthly minimum solar radiation of 1130.56 Wh/m^2 and month of March showed the highest monthly maximum solar radiation of 7238.89 Wh/m^2. Month of March had the highest monthly average solar radiation of 5929.12 Wh/m^2.

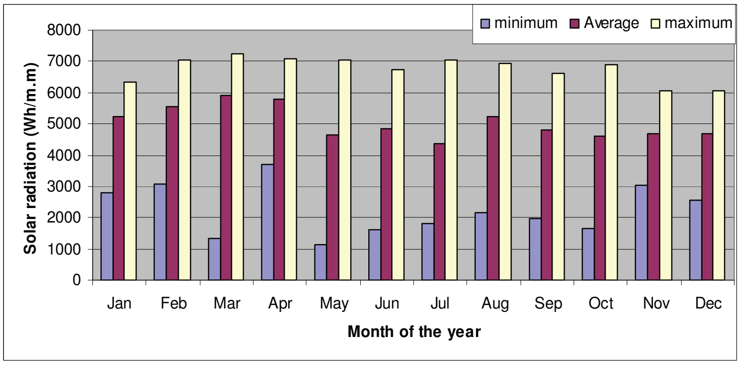

Fig. 5 Monthly minimum, average and maximum solar radiation in Perlis for the year of 2006

Fig. 6 shows the distribution of the average daily solar radiation. The number of days with average annual daily solar radiation greater than 3 kWh/m^2 was 336 days, which indicates that the sky in Perlis was clear for 92.1 % of the days in a year and the solar radiation intensity was very high for these days. Solar radiation between 1 to 3 kWh/m^2 was 29 days or 7.9 % of the days in a year, still suitable for PV application. No solar radiation lower than 1 kWh/m^2, which indicates that no days in a year were absolutely cloudy days which were not good for PV application. Above analysis result shows that solar radiation in Perlis gives potential for PV power generation.

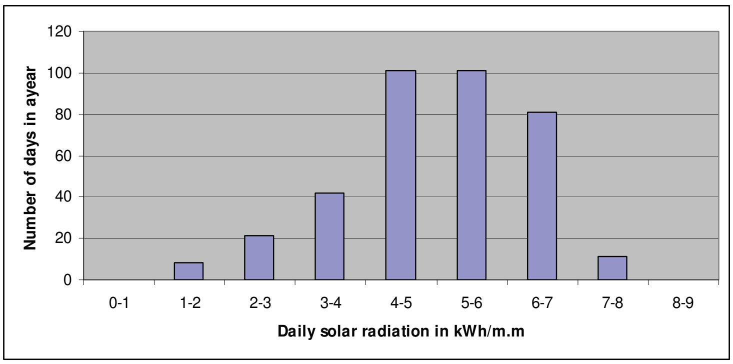

Fig. 6 Distribution of the average annual daily solar radiation in Perlis for the year of 2006

The analysis part is to obtain a simple relation between average solar radiation (y) and average air temperature (x). The scatter plot in Fig 7 shows that a strong relationship between these two variables, suggests the possibility to obtain such data by linear regressions. The three outstanding data points (observations no. 1, 22 and 198) are well apart from the remainder of the data. These data points are is identified as outliers.

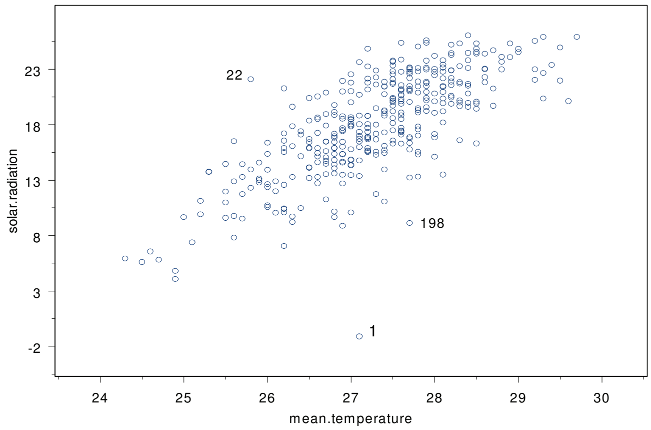

Fig. 7 The scatter plot of solar radiation versus air temperature

By using LTS method, the coefficients of $\hat{\beta}_0$ and $\hat{\beta}_1$ are -87.1081 and 3.8594 respectively. The least squares regression model fit to the solar radiation data is

$$\hat{y} = -87.1081 + 3.8594\,x$$

where $\hat{y}$ is the estimated value of solar radiation corresponding to the air temperature of x cases. The LTS fitted equation is plotted in Fig 8.

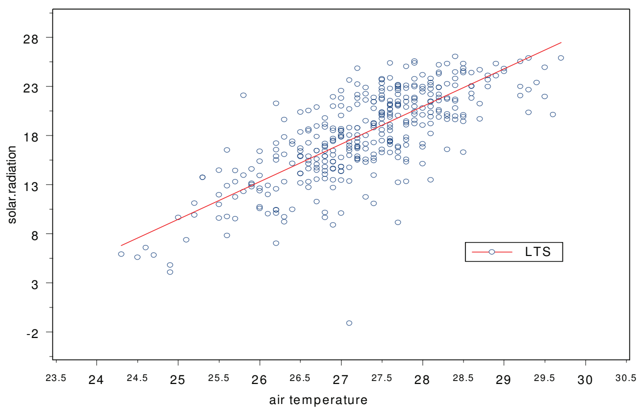

Fig. 8 The scatter plot with the fitted LST line of air temperature versus solar radiation

The robust multiple R-Squared, $R^2 = 0.5993$; means that, about 60 percent of the variability in temperature is accounted for by the straight-line fit to solar radiation. Then, a normal probability plot of residuals, e_i shows in Fig 9. It is almost the straight line with three observations is slightly apart from the others.

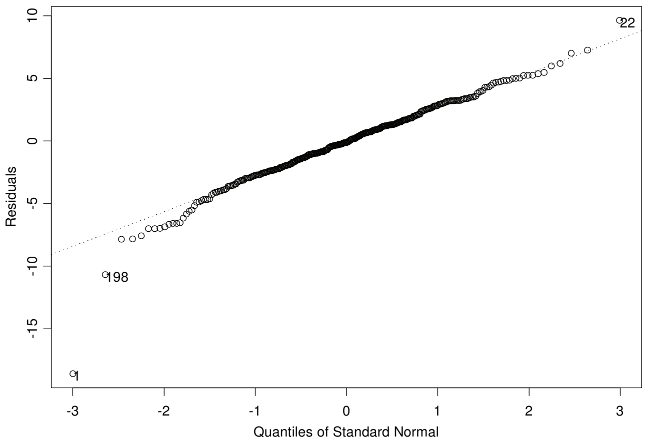

Fig. 9 The graph of the normal probability plot of residuals, *e*

A plot of residuals e_i versus the corresponding fitted values $\hat{y}_i$ is useful for detecting several common types of the model inadequacies. Plot of standardized LTS residuals, e_i versus $\hat{y}_i$ is shown in Fig 10. Based on this Fig. , the patterns show the variance of the errors is constant and stable. There is only three observations are apart from the boundaries.

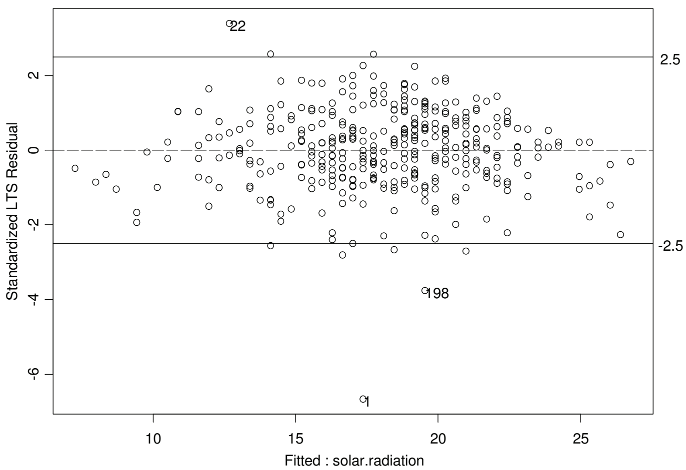

Fig. 10 The graph of the standardized LTS residual plot

Conclusion

According to result can be concluded that the LTS robust regression model can be used to estimate solar radiation for the year of 2006 in Perlis, Northern Malaysia. The relationship between average air temperature and average solar radiation is linear and the value of R^2 is 0.5593 which is quite good to present variability in average air temperature by average solar radiation.

Acknowledgements

The authors wish to thank School of Electrical System Engineering, University of Malaysia Perlis (UniMAP) for the technical and Fundamental Research Grant Scheme 2011 for financial support as well.

References

[1] I. Daut, M. Irwanto , Y.M. Irwan, N. Gomesh, N.S. Ahmad "Combination of Hargreaves method and linear regression as a new method to estimate solar radiation in Perlis, Northern Malaysia" Journal of Solar Energy 85 (2011) 2871–2880

[2] F. Jiang, "Investigation of Solar Energy for Photovoltaic Application in Singapore", the 8th International Power Engineering Conference (IPEC 2007), pp.86 – 89.

[3] Tomas Markvart, " Solar Electricity", John Wiley & Sons, LTD., New York, pp.5-19, 1994.

[4] A. Itagaki, H. Okamura, M. Yamada, " Preparation of Meteorological Data Set Throughout Japan For Suitable Design of PV Systems", 3rd World Conference on Photovoltaic Energy Conversion, 2003, Japan., pp. 2074 – 2077.

[5] A. Mellit, S.A. Kalogirou, S. Shaari, H. Salhi, A.H. Arab, " Methodology for Predictiong Sequences of Mean Monthly Clearness Index and Daily Solar radiation Data in Remote Areas: Application for Sizing a Stand-alone PV System", Renewable Energy, Science Direct, 2007, pp. 1570 – 1590.

[6] Y. Shijun, Y. Hongxing, "The Potential Electricity Generating Capacity of BIP In Hong Kong", IEEE Xplorer, pp. 1345 –1348.

[7] M.J. Wu, E.J. Timpson and S.E. Watkins, "Temperature Consideration in Solar Arrays", IEEE Explorer, 2004.

[8] Montgomery, D. C. and Peck, E. A "Introduction to linear regression analysis" 2nd ed. Wiley. New York, 1992.

[9] S. Weixiang, A.S.K. Bin and O.K Seng, "A study on Standalone Photovoltaic System with Real Meteorological Data at Malaysia", IEEE Xplorer, pp. 937 – 941.

[10] M.D. Kadderi, " Geoasset Assessment of North Perlis Natural Resources for Tourism Development", Proceedings of the Regional Symposium on Environment and Natural Resources (Vol. 1) 10-11th April 2002.

[11] Barnett, V. and Lewis, T, "Outliers in Statistical Data" Wiley and son, New York, 1994.

[12] Belsley, D. A., Kuh, E. and Welsch, R. E. "Regression Diagnostics: Identifying influential data and sources of collinearity" John Wiley & Sons, New York, 1980.

[13] Rousseeuw . P. J. "Least Median of Squares Regression", Journal of the American Statistical Association, Vol. 79, No. 388, 1984.