International Journal of Energy and Environmental Research
Vol.5, No.3, pp.19-30, November 2017
Published by European Centre for Research Training and Development UK (www.eajournals.org)

GLOBAL SOLAR RADIATION MODELING ON A HORIZONTAL SURFACE USING POLYNOMIAL FITTING

Michael Gyan* and Abraham Amankwah

University of Ghana - Department of physics

ABSTRACT: *An attempt has been made to use a polynomial fitting to model global solar radiation on a horizontal surface that was observed by using Pyranometer at University of Ghana Legon, (U.G), situated in Accra, Ghana. The observed solar radiation data was filtered by using fitting and smoothing methods. The polynomial data fitting method was tested by using different degrees of polynomial curve fittings. The root mean square error (RMSE) was used to calculate the error and the R^2 (coefficient of determination) value was also determined. The polynomial fittings were carried out for various periods (pre- harmattan, early harmattan and late harmattan period) of the year.*

KEYWORDS: Solar Radiation, Curve Fitting, Polynomial, Pyranometer, RMSE, Smoothing, Error, Coefficient of Determination

INTRODUCTION

The radiant energy from the sun incident on the earth's surface either directly or scattered radiation determines the temperature of both the surface of the earth and the lower atmosphere of the earth, and also determines the evaporation capacity and climatic features (Baroti et al, 1993).

Most living things on the surface of the earth depend on the sun's radiant energy for survival. Solar radiation is largely optical radiation within a broad region of the electromagnetic spectrum which includes ultra-violet, visible- light and infrared radiation. This radiation consists of electromagnetic radiation emitted by the sun in the spectral region (Zoltan et al, 2000). Solar radiation involves near-infrared and ultraviolet radiation emitted from the sun. The intensity of solar radiation outside the earth's atmosphere is 1367 w/m^2 and this is also called the solar constant.

The magnitude of solar radiation can be obtained by either modeling approaches or observational methods. The observational method that can be used to measure solar radiation and atmospheric parameters can be classified into two types: the surface based subsystem and space- based subsystem. The instruments that can be used for the surface based subsystem include thermograph, weather radar, Pyranometer, transmissometer, and Stevenson screen, etc.

In Karim et, al., 2011, wavelet transform was used to compress the solar radiation data and to develop a new mathematical model for solar radiation data forecasting and prediction. Their work utilized two types of wavelets namely Meyer wavelets and Symlet 6 wavelets. Wu and Chan, 2011 proposed a novel hybrid model to predict the hourly solar radiation data collected at Nanyang Technological University, Singapore. They use Autoregressive and Moving Average (ARMA) and Time Delay Neural Network (TDNN). Their method gives better prediction with higher accuracy. Genc et al., 2002, studied the use of cubic spline functions to analyze the solar radiation in Izmir, Turkey. They conclude that cubic spline regression

ISSN 2055-0197(Print), ISSN 2055-0200(Online)

provides a more accurate description of the relationship between total solar radiation data and the time of the day as compared to the linear regression model.

Motivated by the works of Wu and Chan 2011, and Karim et al. 2011, in this paper polynomial data fitting will be used to predict global the solar radiation data on a horizontal surface with good Prediction accuracy at minimal error.

Data Fitting

Data fitting or regression model is a statistical technique used in modeling and investigating the relationship between variables (Yorukoglu and Celik, 2006). Regression analysis is the most widely used statistical technique.

The regression model determination is estimated based on the procedure shown below. For given measured data (x_i, y_i), i= 1, 2, 3…N. the model is described as

$$y_i = f(x_i) + \varepsilon_i \qquad (1)$$

where f is regression function and ε is a random error. The mean of the errors is estimated to be zero and independent.

The function *f* can be approximated either by using a polynomial of degree n, exponential, Gaussian function or wavelet based procedure (Karim, 2011). Data regressions can be based on either polynomial approach or non polynomial approach.

Derivation of the Quality of Fit (R^2)

The R^2 can be derived by assuming a function: $f(x) = a_o + a_1x + a_2x^2 + ... + a_nx^n$ where n is a positive integer and the degree of the polynomial. Assumed N > n+1 then $\varepsilon_i = y_i - f(x_i)$, $\varepsilon_i = y_i - \{a_o + a_1x_i + a_2x_i^2 + ... + a_nx_i^n\}$ (2)

Squaring the error in equation (2) and taking the sum.

$$S = \sum_{i=0}^{N} \varepsilon_i^2 = \sum_{i=0}^{N} [y_i - \{a_o + a_1x_i + a_2x_i^2 + ... + a_n x_i^n\}]^2 \qquad (3)$$

The sum of errors in (3) must be minimized in order to obtain least square fitting. Hence

$$\frac{\partial s}{\partial a_i} = 0, i = 0,1,2...n \qquad (4)$$

From equation (4) the following set of equations can be obtained in matrix forms.

$$BA = C \qquad (5)$$

$$\begin{bmatrix} N & \sum x_i & \sum x_i^2 & ... & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & ... & \sum x_i^{n+1} \end{bmatrix}$$

ISSN 2055-0197(Print), ISSN 2055-0200(Online)

□

□ □ □xi 2□xi 3□xi 4...□xi n□2 □ □

where B□ □. □ ,

□ □. □ □

□. □

□ n n□1 n□2 2n □

□ □x_i □x_i □x_i ...□x_i □

□ □ □ □

□a_0 □

□ □a_1 □ □

□a_2 □

□ □

A□ □˙□ and

□. □

□ □

□. □

□ □an □ □

□ □y_i □

□ □

□ □x_i y_i □

□ $_2$ □

□ □x_i y_i □

C □□. □

□ □

□. □

□ □

□. □

□ □ □

xin yi

□□□

Equation (5) can be solved by numerical methods such as Gaussian elimination, Choslesky's method, etc. If n = 1, then the least square fitting in (2) is called linear fitting which is given by

$$y = a_0 + a_1 x \qquad (6)$$

The system of equation in (5) becomes

$$\begin{bmatrix} N & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} \qquad (7)$$

The solutions for (7) is given by

$$a_0 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \quad a_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

The R-squared (R^2) value is also known as the coefficient of determination. This is the statistics that give information about the goodness of fit of the model to a data. The R^2 can be calculated from

$$R^2 = \frac{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2} \qquad (8)$$

where $0 \le R^2 \le 1$

When $R^2 = 1$, it means a perfect fit and the variability in y is completely explained by the regression model.

However, when $R^2 = 0$, it means the model explains none of the variability of the data used.

Data Collection

Solar radiations data were observed using Kipp and Zonen CM 11 pyranometer. This type of pyranometer has a calibration factor or sensitivity of 4.88μV/Wm^{-2} and a response time of 5

ISSN 2055-0197(Print), ISSN 2055-0200(Online)

s. It has a field view of 180 degree and maximum solar irradiance of 4000 Wm^{-2} . The pyranometer is coupled to a voltmeter. The voltmeter helps in taking the actual values of the solar irradiance in volts. In order to obtain the values of solar irradiance, the pyranometer was placed on a stand at a height of about 1.7 meters. Using a stopwatch, the recordings of the total solar irradiance was done in a time interval of 10 minutes. The sensitivity on the pyranometer and the values obtained from the voltmeter helps to determine the solar irradiance at a particular time. This was done by the relation: $1\ Wm^{-2} = 4.88\mu V$. Data was observed between 17th September 2015 and 30th December 2015.

RESULT AND DISCUSSION

Polynomial fitting (Regression) was applied to the observed data starting with degree 2 through degree 7. The coefficient of the polynomial fitting is determined base on 95% confidence interval. Table 1, Table 2, and Table 3 sum up the polynomial fitting results for pre-harmattan, early harmattan and late harmattan respectively. The pre- harmattan period consists of the month of April to October. The data of October 24, 2015, considered as a representative of the pre-harmattan period for the presentation of results. The harmattan period, this also consist of the period of November to March . The data of November 18, 2015, was also regarded as a representative for early harmattan period and data of December 22, 2015, was also used for late harmattan period. The graphs for the various polynomial fit are shown below.

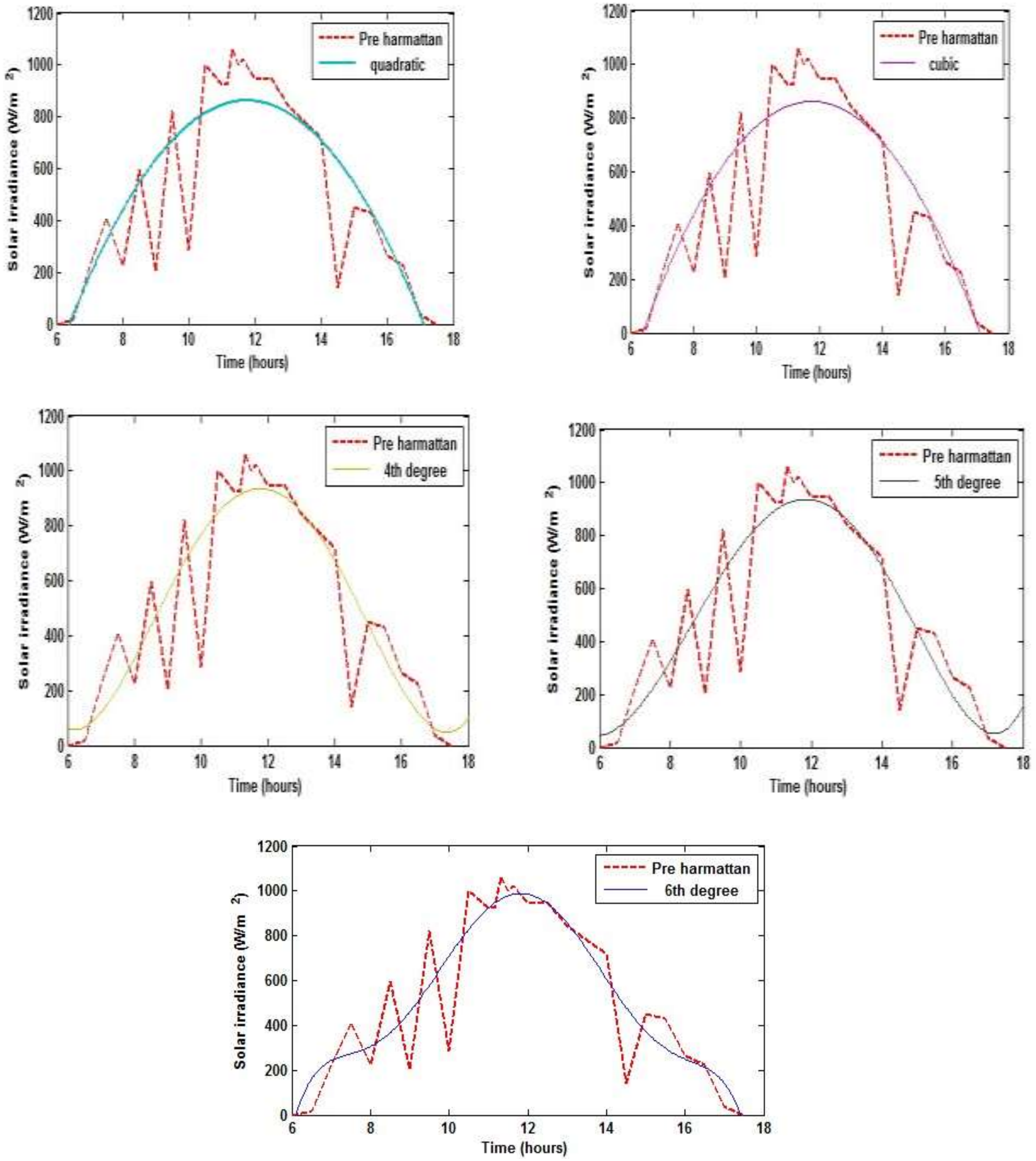
Table 1 shows various polynomial fitting results for observed data during pre-harmattan period.

Polynomial fit	Statistics	
	R^2	RMSE
$f(x)=\sum_{i=0}^{2} a_i x^i$ a_0 = -30.03 a_1 = 705.5, a_2 = -3281	0.7542	194.7
$f(x)=\sum_{i=0}^{3} a_i x^i$ a_0 = -0.0281, a_1 = -29.05, a_2 = 694.5 a_3 = -3243	0.7542	198.7
$f(x)=\sum_{i=0}^{4} a_i x^i$ a_0 = 0.9089 , a_1 = -42.8 , a_2 = 699.2 a_3 = -4599, a_4 = 1.055×10^4	0.7754	179.0

5 f(x)= □$a_i x^i$ i□0 a_0 □0.032 a_1 = -0.961, a_2 = - 0.183, a_3 = - 228.5, a_4 = - 2087, a_5 = 5376	0.8092	182.8 ..
6 f(x) = □$a_i x^i$ i□0 a_0 = - 0.083 , a_1= 5.855 , a_2 = - 163.300 , a_3 = 250.200, a_4 = - 2.027□10^4 ,a_5 = 8.501□10^4 , a_6 = - 1.444□10^5 ,	0.8416	170.5
7 F(x) = □$a_i x^i$ i□0 a_0 = - 0.004 , a_1= 0.243 ,a_2 = - 5.395 ,a_3 = 43.040, a_4 = 171.900, a_5 = - 5206, a_6 = 3.219□10^4 , a_7 = - 6.696□10^4 ,	0.8422	174.4

ISSN 2055-0197(Print), ISSN 2055-0200(Online)

International Journal of Energy and Environmental Research
Vol.5, No.3, pp.19-30, November 2017
Published by European Centre for Research Training and Development UK (www.eajournals.org)

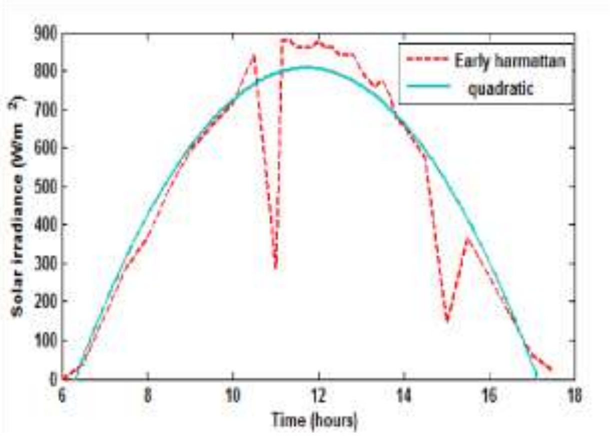


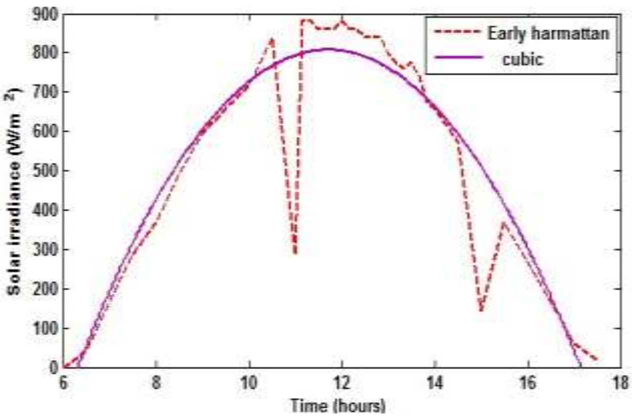
From Table 1 above, the polynomial fitting with 6th degree gives better R^2 (0.8416) and RMSE (170.5) during the pre-harmattan period. This model gives better results without any wiggle at both end points of the graph as shown above. Statistically, the fitting with n=6 gives 84.16 % indication to the variance of the observed data

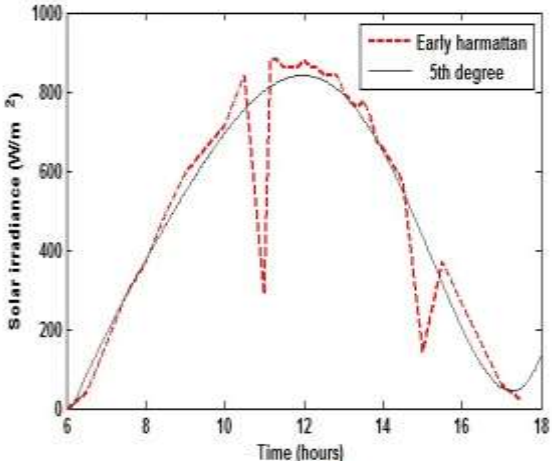
Table 2 shows various polynomial fitting results for observed data during early harmattan period

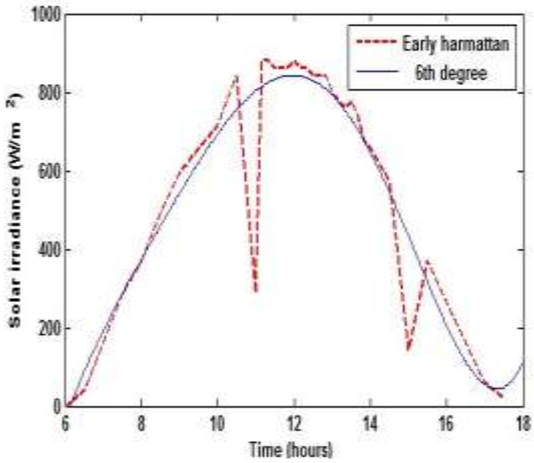
ISSN 2055-0197(Print), ISSN 2055-0200(Online)

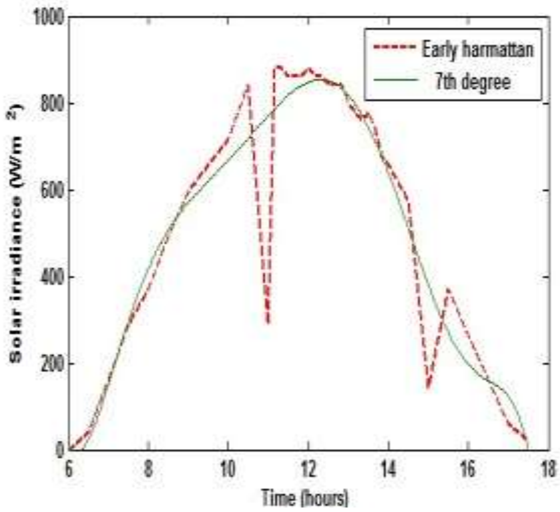
Polynomial fit	Statistics	
	R^2	RMSE
$f(x)=\sum_{i=0}^{2} a_i x^i$ $a_0 = -27.53$, $a_1 = 645.8$, $a_2 = -2980$	0.8512	121.5
$f(x)=\sum_{i=0}^{3} a_i x^i$ $a_0 = 0.04228$, $a_1 = -29.04$, $a_2 = 662.7$ $a_3 = -3040$	0.8512	123.3
$f(x)=\sum_{i=0}^{4} a_i x^i$ $a_0 = 0.532$, $a_1 = -24.96$, $a_2 = 396.3$, $a_3 =$ -2429 $a_4 = 5024$	0.8746	115.1
$f(x)=\sum_{i=0}^{5} a_i x^i$ $a_0 = 0.0813$, $a_1 = -4.252$, $a_2 = -84.430$, a_3 $= -816.800$, $a_4 = -$ 4078, $a_5 = -8455$,	0.8793	114.7
$f(x) = \sum_{i=0}^{7} a_i x^i$ $a_0 = -0.015$, $a_1 = 1.241$, $a_2 = -42.69$, $a_3 =$ 798, $a_4 = -8750$, $a_5 = 5.624 \times 10^4$, $a_6 = -$ $.959 \times 10^5$, $a_7 = 2.848 \times 10^5$,	0.8634	113.0











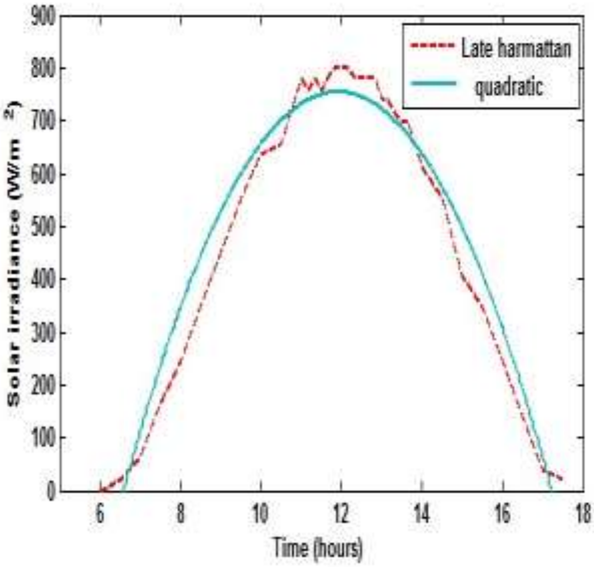
For the early harmattan period, the polynomial with degree 5 (quintic) seems to give a better result as compare with the other polynomial fitting model. From the Table 2, the polynomial fitting with 7^{th} - degree gives higher R^2 but wiggles at the end point of the graph as shown

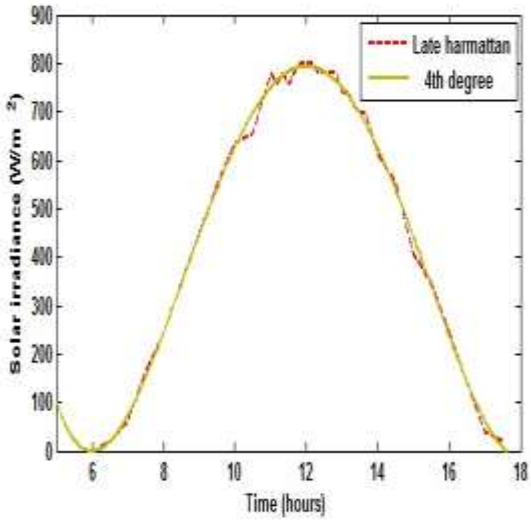
ISSN 2055-0197(Print), ISSN 2055-0200(Online)

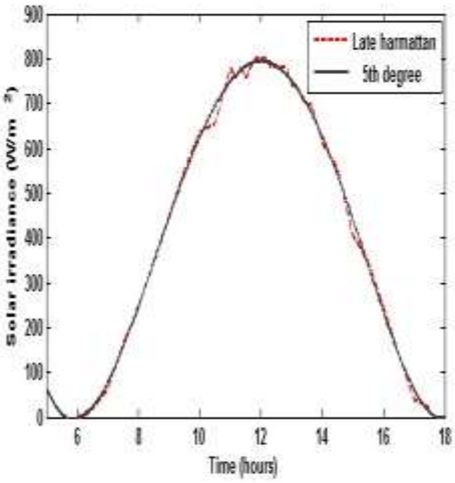
above. By careful observation to the fitting, the best fitting for early harmattan data is 5th - degree polynomial as shown above with R^2 of 0.8793 and RMSE of 114.7

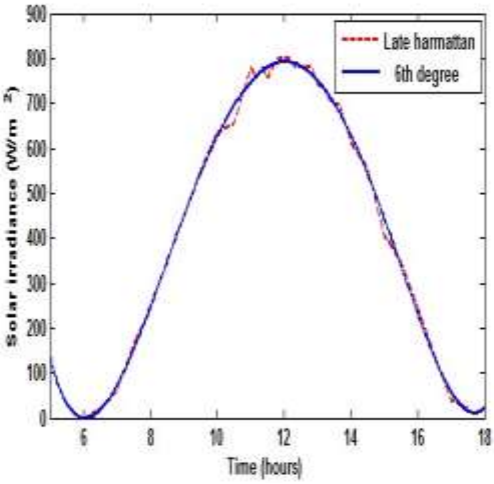
Table 3 shows various polynomial fitting results for observed data during late harmattan period.

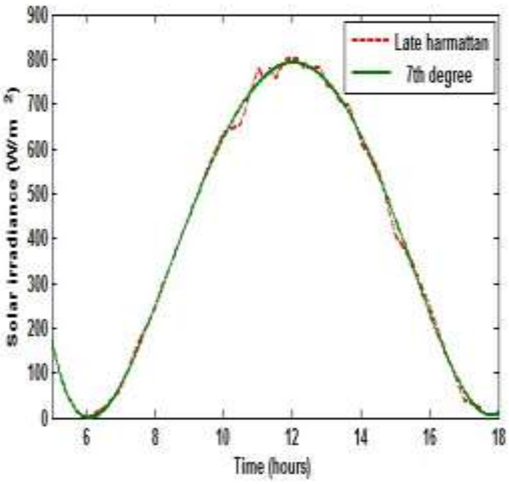
Polynomial fit	Statistics	
	R^2	RMSE
$f(x)=\sum_{i=0}^{2} a_i x^i$ $a_0 = -26.64$, $a_1 = 634.3$, $a_2 = -3023$	0.957	60.08
$f(x)=\sum_{i=0}^{3} a_i x^i$ $a_0 = -0.6347$, $a_1 = -4.103$ $a_2 = 380.8$ $a_3 = -2130$	0.9613	57.92
$f(x)=\sum_{i=0}^{4} a_i x^i$ $a_0 = 0.6097$, $a_1 = -29.29$ $a_2 = 483.4$ 7) $a_3 = -3163$ $a_4 = 7111$	0.9974	15.25
$f(x)=\sum_{i=0}^{5} a_i x^i$ $a_0 = 0.008$, $a_1 = 0.119$, $a_2 = -18.080$, $a_3 = 359$, $a_4 = -2496$, $a_5 = 5730$,	0.9975	15.32
$f(x)=\sum_{i=0}^{6} a_i x^i$ $a_0 = 0.003$, $a_1 = -0.2159$, $a_2 = 6.576$, $a_3 = -4.900$, $a_4 = 1154$, $a_5 =$	0.9975	15.36
-5.882, $a_6 = 1.157 \times 10^4$,		
$F(x)=\sum_{i=0}^{7} a_i x^i$ $a_0 = -0.0003$, $a_1 = 0.036$, $a_2 = -1.343$, $a_3 = 27.770$, $a_4 = -348.900$, $a_5 = 2670$, $a_6 = -1.121 \times 10^4$, $a_7 = 1.940 \times 10^4$,	0.9975	15.61


Late harmattan
quadratic
Solar irradiance (W/m 2)
Time (hours)


Late harmattan
4th degree
Solar irradiance (W/m 2)
Time (hours)


Late harmattan
5th degree
Solar irradiance (W/m 2)
Time (hours)


Late harmattan
6th degree
Solar irradiance (W/m 2)
Time (hours)


Late harmattan
7th degree
Solar irradiance (W/m 2)
Time (hours)

Among the entire fitting for late harmattan, the polynomial with degree 4 gives better results as compared with other polynomial fitting model. The R^2 and RMSE values for various fitting can be seen in table 3. The best polynomial fitting model for the late harmattan data is quartic (n= 4) as shown in Fig above with R^2 of 0.9974 and RMSE of

15.25.

CONCLUSION

Solar radiation data fitting by using polynomial fit method have been studied for three different periods (pre- harmattan, early harmattan, and late harmattan) in this work. The model for solar radiation can be used to forecast the amount of solar radiation received at the University of Ghana within a certain period of the year. The polynomial fit model can be used to determine the optimum system sizing for solar electricity generation. From the results, the fitting model with 4^{th} - degree order, 5^{th} - degree order and 6^{th} - degree order gives a better result for late harmattan period, early harmattan period, and pre- harmattan respectively without wiggle at the ends and the value of RMSE is 15.25, 114.7 and 170.5 respectively. Also, R^2 was found to be 0.9974 (late harmattan), 0.8793 (early harmattan) and 0.846 (pre-harmattan).

REFERENCES

Wu, J. and Chan, C.K. Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN. Solar Energy 85:808-817, 2011.

Karim, S.A.A., Singh, B.S.M., Razali, R. and Yahya, N. Data Compression Technique for Modeling of Global

Solar Radiation. In Proceeding of 2011 IEEE International Conference on Control System, Computing and Engineering (ICCSCE) 25-27 November 2011, Holiday Inn, Penang, pp. 448-35, 2011a.

Genc, A., Kinaci, I., Oturanc, G., Kurnaz, A., Bilir, S. and Ozbalta, N. Statistical Analysis of Solar Radiation Data Using Cubic Spline Functions. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects. 24:12 1131-1138, 2002.

Yorukoglu, M., and Celik, A.N., A critical review on the estimation of daily global solar radiation from sunshine duration, Energy Conversion and Management 47: 2441-2450, 2006.

Baróti, István Energiafelhasználói Kézikönyv, Manual of energy utilization. Környezettechnika Szolgáltató Kft.

Budapest, 1993.

ISSN 2055-0197(Print), ISSN 2055-0200(Online)