JOURNAL OF ENGINEERING SCIENCES

Volume 8, Issue 2 (2021)

Onyeka V. O., Nwobi-Okoye C. C., Okafor O. C., Madu K. E., Mbah O. M. (2021). Estimation of global solar radiation using empirical models. Journal of Engineering Sciences, Vol. 8(2), pp. G11-G19, doi: 10.21272/jes.2021.8(2).g2

Estimation of Global Solar Radiation Using Empirical Models

Onyeka V. O.[1], Nwobi-Okoye C. C.[1], Okafor O. C.[2*], Madu K. E.[1], Mbah O. M.[3]

[1] Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State, Nigeria;
[2] Grundtvig Polytechnic, Oba, Anambra State, Nigeria;
[3] Federal University Oye, Ekiti State, Nigeria

Article info:
Paper received: September 7, 2021
The final version of the paper received: December 10, 2021
Paper accepted online: December 15, 2021

***Corresponding email:**
okaforobiorac@gmail.com

Abstract. The dearth of solar radiation data availability has necessitated the development of several mathematical models for estimating global solar radiation (GSR) of regions using the readily available meteorological data of the region. This study was centered on estimating the GSR of the Ihiala region in Sub-Saharan Africa using empirical models. For the last ten years, meteorological data from the Nigerian Meteorological Agency (NIMET) were used. The sunshine-based equation, temperature-based equation, and multivariate polynomial equations were the empirical models employed to estimate the GSR of the region. The performance of the seven models was determined using statistical measures. From the results obtained, the seven models had their respective P-values all less than 5 % significant level for a confidence interval of 95 %. Thereby attesting their suitability for GSR estimation of the region is needed. Also, from the other statistical tools employed, the considered multivariate model had better estimation performance than the other models. Therefore, the considered multivariate model is suitable for estimating the GSR of the Ihiala region in Sub-Saharan Africa.

Keywords: renewable energy, global solar radiation, artificial neural network, statistical tests.

1 Introduction

Solar energy is a primary source of energy and is non-polluting and inexhaustible. According to [1], solar energy can be harnessed using three different methods, namely: using photocells or photovoltaic cells for converting solar energy to electrical one directly; converting solar energy into thermal energy through the application of suitable devices which may be subsequently converted into mechanical, chemical, or electrical energy; through photosynthesis in which plant trap the solar energy and it is converted to chemical energy- biomass energy.

Solar energy is ultimately harnessed using solar cells to convert the solar energy that falls on it into electricity. Research [2] argues that renewable and clean energy, such as solar energy, is needed to maintain quality of life and the environment. The use of solar energy, like any other natural resource, requires detailed information on the total number of cases of solar radiation on the earth's surface. The total solar radiation incident on the earth's surface is called global solar radiation (GSR). GSR data are necessary for various design, engineering, simulation, and performance evaluation of any project utilizing solar energy. Most recent technologies are hinged principally on solar energy applications. These technologies are seen in the invention of solar cars, solar heaters, solar pumps, solar refrigerators, solar air conditioners etc. Though, some designed solar cars are hybrid systems that still employ internal combustion (IC) engines in their operations.

Owing to the broad areas of application of solar technologies, the knowledge of the intensity of global solar radiation of a geographical location is imperative. Also, the intensity of GSR is influenced by seasonal changes, geographical location, and collector position [3]. Thereby the need for frequent measurement of GSR parameters is created. Therefore, there is a need to develop mathematical models using empirical equations to estimate the GSR of a region of interest.

Therefore, this study estimated the GSR of the Ihiala region in Sub-Saharan Africa using empirical models. The meteorological data employed in the study were obtained from the Nigerian Meteorological Agency (NIMET). The estimation performance of all the models was validated using some statistical tools, such as mean bias error (MBE), mean percentage error (MPE), and root mean square error (RMSE).

2 Literature Review

Ihiala is a city in Sub-Saharan Africa, located in the southern part of Anambra state at latitude 5.85 N, longitude 6.86 E, and altitude 146 m above sea level and has long served as the Local Government Area of the region. The Local Government Area has a population of about 87,796 persons.

The frequent measurement of GSR of a region of interest due to seasonal changes and the position of collectors [3] has thus called the attention of researchers to develop mathematical models for estimating the GSR of a geographical location. In [4], a linear regression model used in correlating the GSR data with relative sunshine duration was developed using the Angstrom type model. [5] studied the correlation between the measurement data on global solar radiation and meteorological parameters using solar radiation, average daily maximum temperature, average daily relative humidity, average daily sea level pressure, average daily vapor pressure, and many hours of bright sunshine obtained from different parts of Egypt.

Also, [6] developed two decomposition-based models that predict atmospheric transmittance using temperature and relative humidity data of Petronas city in India. [7], developed a mathematical model for estimating global solar radiation using Angstrom's formulation equation for the Himalayas region, Nepal. The statistical analysis results proved that the model was suitable for estimating global solar radiation parameters of the region. Furthermore, [8] examined and compared the results of three isotropic sky models (Liu and Jordan, Badescue and Koronakis model) with experimental data for global solar radiation estimation in eastern Nigeria. Scilab computational software tool was used in implementing the models. Statistical tools were used to determine the accuracy of the models. It was observed that Liu and Jordan's model gave the most negligible value of mean bias error and t-stat of 0.0127 $W \cdot h/m^2$ and 3.3947 respectively. In addition, [9] compared anisotropic sky models with experimental data for solar radiation estimation on tilted surfaces in the Sub-Saharan African climate. The models considered were the Perez model and Hay, Davies, Klucher, and Reindl (HDKR) models. The result showed that the Perez model recorded the least mean bias error (0.05 $W \cdot h/m^2$) and root mean square error (0.02 $W \cdot h/m^2$) than the HDKR model.

3 Research Methodology

3.1 Data processing

The meteorological data of the Ihiala region was retrieved from the Nigerian Meteorological Agency (NIMET), Awka, during the last ten years. The climatic data collected were: monthly mean daily sunshine duration, monthly mean daily hours of bright sunshine, monthly mean minimum and maximum temperatures, monthly mean global solar radiation, monthly mean relative humidity, and monthly mean extraterrestrial global solar radiation of the Ihiala region.

The following characteristics are used:

– monthly mean global solar radiation H_m – the monthly mean of total short-wave radiation of the sky falling on the horizontal surface of the earth. It includes both direct sunlight and scattered radiation resulting from the reflected or scattered sun's energy;

– monthly mean extraterrestrial global solar radiation H_O – defined as the monthly mean of the total short-wave radiation from the sky before reaching the atmosphere, where its reflection, absorption, scattering, and diffusion occurs;

– monthly mean minimum temperature T_{min} – the monthly average of the minimum temperature of an area under study;

– monthly mean maximum temperature T_{max} – gives the monthly average of the maximum temperatures of an area being studied;

– monthly mean sunshine duration/day length N – gives the mean of the day length during the sunshine period for a month;

– monthly mean relative humidity R_H – the average amount of water vapor present in the air as a percentage of the maximum airflow at a given temperature for a month. The relative humidity is the actual water vapor pressure ratio to the saturation vapor pressure;

– monthly average hours of bright sunshine n – the average of all the daily hours of bright sunshine for a month.

MINITAB 2020 software was used to determine the correlation coefficient between the region's measured and estimated global solar radiation values. It was also employed in deriving the mathematical models using the formulation equations proposed by other researchers. Microsoft Excel 2020 was used to simplify and compute the model terms.

The estimation of global solar radiation of the Ihiala region was done by employing Angstrom–Prescott (sunshine-based), Hargreaves–Samani (temperature-based), and various multivariate (polynomial-based) formulation equations. The estimation performance of the models was validated using some standard statistical tools, such as root mean square error (RMSE), mean bias error (MBE), mean percentage error (MPE), the coefficient of correlation (r), Nash-Sutcliffe model (NSE), R-square value, and t-statistic test. The best suitable model for estimating global solar radiation of the Ihiala region was obtained through the confirmatory results of the statistical tests. The methodical procedures employed in this study are further elucidated in the preceding sections.

3.2 Applied formulation models

According to the Angstrom-Prescott model, the first correlation proposed for estimating the monthly mean daily global solar radiation on a horizontal surface using sunshine duration [10] and clear sky radiation data was first developed by Angstrom. The correlation is shown thus:

$$\frac{H_m}{H_c} = a + b\left(\frac{n}{N}\right). \quad (1)$$

However, due to the problems usually encountered in The new model, the Angstrom-Prescott model or sunshine model. In [12, 13], it was stated that many researchers have equally employed this concept.

According to the Hargreaves-Samani model, the temperature-based model used to estimate global solar radiation is used. The model is based on the temperature difference of the experimental site and is shown thus:

$$\frac{H_m}{H_o} = K_r(T_{max} - T_{min})^{\frac{1}{2}}, \quad (2)$$

where K_r – an empirical coefficient.

Moreover, the multivariate model development involves applying multivariate regression analysis to model climatic data having two or more independent variables. The presence of many predictors in a multivariate model enhances its estimation performance. The five multivariate formulation equations are of the form indicated in the following equations [12]:

$$H = acos\emptyset + bcosn + cT_{max} + d\left(\frac{\bar{n}}{N}\right) + e\left(\frac{T_{max}}{R.H}\right) + f\left(\frac{T_{max}}{R.H}\right)^2 + gcos\emptyset.cosn + h; \quad (3)$$

$$H = acos\emptyset + bcosn + cT_{max} + d\left(\frac{\bar{n}}{N}\right) + e\left(\frac{T_{max}}{R.H}\right) + f\left(\frac{T_{max}}{R.H}\right)^4 + gcos\emptyset.cosn + h\frac{T_{max}}{cos\emptyset} + i; \quad (4)$$

$$H = acos\emptyset + bcosn + cT_{max} + d\left(\frac{\bar{n}}{N}\right) + e\left(\frac{\bar{n}}{N}\right)^3 + f\left(\frac{T_{max}}{R.H}\right) + \\ +g\left(\frac{T_{max}}{R.H}\right)^2 + h\left(\frac{T_{max}}{R.H}\right)^3 + icos\emptyset.cosn + j\left(\frac{T_{max}}{cos\emptyset}\right) + kcos^2n + l; \quad (5)$$

$$H = acos\emptyset + bcosn + cT_{max} + d\left(\frac{\bar{n}}{N}\right) + e\left(\frac{\bar{n}}{N}\right)^3 + f\left(\frac{T_{max}}{R.H}\right) + g\left(\frac{T_{max}}{R.H}\right)^2 + \\ +h\left(\frac{T_{max}}{R.H}\right)^3 + i\left(\frac{T_{max}}{R.H}\right)^4 + jcos\emptyset.cosn + k\left(\frac{T_{max}}{cos\emptyset}\right) + lcos^2n + m; \quad (6)$$

$$H = acos\emptyset + bcosn + cT_{max} + d\left(\frac{\bar{n}}{N}\right) + e\left(\frac{T_{max}}{R.H}\right) + f(R.H) + \\ +gcos\emptyset.cosn + h\left(\frac{T_{max}}{cos\emptyset}\right) + i\left(\frac{T_{max}}{R.H}\right)^2 + j\left(\frac{\bar{n}}{N}\right)^2 + kcos^2n + l, \quad (7)$$

where $\emptyset$ – location latitude, degrees; n – the monthly mean sunshine hours; N – maximum sunshine duration or day length; T_{max} – maximum monthly mean temperature, °C; n – day number in the year, R_H – monthly mean relative humidity; a–m – correlation coefficients or constants; n/N – the relative sunshine ratio; H – the monthly mean global solar irradiance value.

According to [14, 15], an approximate equation for determining the declination angle is as follows:

$$\delta = 23.45\, sin\left(360\frac{284+n}{365}\right). \quad (8)$$

According to [14], the sunset hour angle is computed using the relation:

$$\omega_s = cos^{-1}(-tan\emptyset tan\delta). \quad (9)$$

Monthly mean daily sunshine duration N is computed using:

$$N = \frac{2}{15}cos^{-1}(-tan\emptyset tan\delta) = \frac{2}{15}\omega_s. \quad (10)$$

Monthly mean daily extraterrestrial global solar radiation H_O is determined using the relation:

$$H_O = \frac{24\times3600G_{sc}}{10^6\pi}\left(1 + 0.033cos\frac{360n}{365}\right)\left(cos\emptyset cos\delta sin\omega_s + \frac{\pi\omega_s}{180}\, sin\emptyset sin\delta\right), \quad (11)$$

where G_{sc} = 1367 W/m^2 – the solar constant; $\emptyset$ – the area's latitude considered, degrees; δ – solar declination angle, degrees; ω_s – sunset hour angle, degrees.

3.3 Validation of the models

Validatrions of the models has been carried out based on the calculation of the following characteristics.

The mean bias error (MBE) is determined using the following relation [15]:

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(H_{ical} - H_{imea}). \quad (12)$$

The mean percentage error (MPE) is computed with the equation expressed by [16]:

$$MPE(\%) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{H_{ical}-H_{imeas}}{H_{imeas}}\right)\times 100. \quad (13)$$

This is the relation used in the computation of MPE. A percentage error in a range from –10 % to +10 % is sufficient.

The smaller the root mean square error (RMSE), the better the model's estimation strength and accuracy. Its computational formula is given by [17]:

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(H_{ical} - H_{imeas})^2\right]^{\frac{1}{2}}, \quad (14)$$

where H_{imeas}, H_{ical}, n – the i-th measured values and i-th calculated values of daily global solar radiation and the number of values, respectively.

According to [18], a model is more efficient only when the Nash-Sutcliffe equation (NSE) is closer to 1.0:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(H_{imea}-H_{ical})^2}{\sum_{i=1}^{n}(Hi_{mea}-\bar{H}_{mea})^2}, \quad (15)$$

where H_{imea} – the mean measured global solar radiation.

Also, Karl Pearson's method was used in this work, and the relation is given as:

$$\boldsymbol{r} = \frac{\sum \boldsymbol{XY}}{\sqrt{(\sum \boldsymbol{X^2})(\sum \boldsymbol{Y^2})}}, \quad (16)$$

where X is the difference between the measured global radiation and the mean of the measured global radiation; Y is the difference between the estimated global radiation and the mean of the estimated global radiation.

According to the t-statistic test, at a confidence interval of 95 % and significance level of 5 %, the t-statistics test was carried out to determine how small its value was based on the fact that the smaller the value, the better the performance of the model. The formula used to compute it is given by the following equation:

$$t = \left[\frac{(n-1)(MBE)^2}{(RMSE)^2-(MBE)^2}\right]^{\frac{1}{2}}, \quad (17)$$

where the smaller t, the better its estimation performance.

According to the R-squared method, models with R2 closer to one are considered the best model. The formula used to compute its value is given as follows:

$$R^2 = \frac{\sum_{i=0}^{n}\left(H_{imea}-\bar{H}_{imea}\right)^2}{\sum_{i=0}^{n}\left(H_{imea}-\bar{H}_{ical}\right)^2}, \quad (18)$$

where H_{imea} – the mean measured global solar radiation; H_{ical} – the mean of calculated global solar radiation.

4 Results and Discussion

4.1 Sunshine-based model

With the aid of the sunshine-based formulative equation, the following equation was obtained:

$$\frac{H_m}{H_o} = 0.397 + 0.421\left(\frac{\bar{n}}{N}\right), \quad (19)$$

where H_m – measured global solar radiation; H_o – extraterrestrial global solar radiation; n/N – relative sunshine or fraction of sunshine.

The significance level of each model term and the entire regression model were ascertained from the performed analysis of variance (ANOVA). The results are presented in Tables 1, 2.

Table 1 – The significance level of the model terms

Predictor	Coef	SE Coef	*T*	*P*
Constant	0.397	0.024	16.56	0.000
n/*N*	0.421	0.057	7.38	0.000

S = 0.0818436   R-Sq = 31.6 %   R-Sq(adj) = 31.0 %

From Table 1, each of the model terms was all significant as their respective p-values are far less than the chosen ∝-value of 0.05 for a confidence interval of 95 %. The values of these statistical parameters: S, R-Sq, and R-Sq(adj.), show the level of accuracy of the model in fitting data. The small values of R-Sq and R-Sq(adj.) explain that the model's predicted strength is poor. The small value of S implies that the model's data fitting ability is not too strong.

Table 2 – Analysis of variance (ANOVA)

Source	*DF*	*SS*	*MS*	*F*	*P*
Regression	1	0.365	0.365	54.48	0.000
Residual error	118	0.790	0.007	–	–
Total	119	1.155	–	–	–

The p-value of 0.000 for the regression model attests that the model is statistically significant. In other words, the model's degree of accuracy is only justified by using other statistical and analytical tools employed in this work.

4.2 Temperature-based model

This model was solely based on the temperature difference of the test site- Ihiala region. The model is given thus:

$$\frac{H_m}{H_o} = 0.236 + 0.106(T_{max} - T_{min})^{\frac{1}{2}}. \quad (20)$$

The statistical significance is presented in Table 3.

Table 3 – The significance level of the model terms

Predictor	Coef	SE Coef	*T*	*P*
Constant	0.236	0.057	4.10	0.000
n/*N*	0.106	0.018	5.79	0.000

S = 0.0873280   R-Sq = 22.1 %   R-Sq(adj) = 21.5 %

From Table 3, it is quite vivid that each of the model terms is significant as their respective p-values are less than the significant level (∝) of 0.05.

Therefore, the model's response predictors or independent variables are all essential. In addition, the smaller values of R-Sq and R-Sq(adj) imply a poor predicting strength of the temperature model since their higher values connote a better-predicting strength of models.

Further, on the model's statistical analysis, the ANOVA result shown in Table 4 proves that the temperature model is statistically suitable for response prediction since the p-value is less than ∝-value of 0.05.

Table 4 – Analysis of variance (ANOVA)

Source	*DF*	*SS*	*MS*	*F*	*P*
Regression	1	0.255	0.255	33.50	0.000
Residual error	118	0.890	0.008	–	–
Total	119	1.155	–	–	–

4.3 Multivariate models

The following multivariate models were developed using the polynomial models given in equations (3)–(7).

The response predictors or independent variables which are not significant have been removed from all the models.

1) multivariate model 1:

$$H_m = -8.85 - 0.922cosn + 0.768T_{max} + 6.25\frac{\bar{n}}{N} + 1.9\frac{T_{max}}{R.H} - 1.6\left(\frac{T_{max}}{R.H}\right)^2. \quad (21)$$

The ANOVA analysis for multivariate model 1 is shown in Table 5.

From Table 5, the p-value of the regression model is much less than the ∝-value of 0.05 at a confidence interval of 95 %. Hence, the model is statistically significant for response prediction (global solar radiation) of the Ihiala region.

Table 5 – Analysis of variance (ANOVA) for model 1

Source	*DF*	*SS*	*MS*	*F*	*P*
Regression	5	719.72	143.94	18.89	0.000
Residual error	114	868.72	7.62	–	–
Total	119	1588.44	–	–	–

S = 2.76050 R-Sq = 45.3 % R-Sq(adj) = 42.9 %

2) multivariate model 2:

$$H_m = -7.65 - 0.910cosn + 0.716T_{max} + 6.19\frac{\bar{n}}{N} + 2.51\frac{T_{max}}{R.H} - 1.94\left(\frac{T_{max}}{R.H}\right)^4 \quad (22)$$

The ANOVA analysis for multivariate model 2 is shown in Table 6.

Table 6 reveals that the model is statistically significant as the p-value of the model is far less than the ∝-value of 0.05 for a confidence interval of 95 %.

Table 6 – Analysis of variance (ANOVA) for model 2

Source	*DF*	*SS*	*MS*	*F*	*P*
Regression	5	721.23	144.25	18.96	0.000
Residual error	114	867.21	7.61	–	–
Total	119	1588.44	–	–	–

S = 2.75810 R-Sq = 45.4 % R-Sq(adj) = 43.0 %

3) multivariate model 3:

$$H_m = 2.07 - 0.670cosn + 0.691T_{max} + 6.41\frac{\bar{n}}{N} + 1.28\left(\frac{\bar{n}}{N}\right)^3 - -50.2\frac{T_{max}}{R.H} + 88.0\left(\frac{T_{max}}{R.H}\right)^2 - 46.9\left(\frac{T_{max}}{R.H}\right)^3 + 1.58cos^2n. \quad (23)$$

The ANOVA analysis for multivariate model 3 is shown in Table 7.

The p-value of the multivariate regression model 3 is less than the 0.05 value of the significant level (∝). This proves that the model is suitable for response prediction.

Table 7 – Analysis of variance (ANOVA) for model 3

Source	*DF*	*SS*	*MS*	*F*	*P*
Regression	8	775.45	96.93	13.23	0.000
Residual error	111	812.98	7.32	–	–
Total	119	1588.44	–	–	–

S = 2.70632 R-Sq = 48.8 % R-Sq(adj) = 45.1 %

4) multivariate model 4:

$$H_m = -7.4 - 0.646cosn + 0.686T_{max} + 6.20\frac{\bar{n}}{N} + 1.76\left(\frac{\bar{n}}{N}\right)^3 + 18\frac{T_{max}}{R.H} - 85\left(\frac{T_{max}}{R.H}\right)^2 + 139\left(\frac{T_{max}}{R.H}\right)^3 - 71\left(\frac{T_{max}}{R.H}\right)^4 + 1.58cos^2n. \quad (24)$$

The ANOVA analysis for multivariate model 4 is shown in Table 8.

The p-value of the multivariate regression model 4 is less than the 0.05 value of the significant level (∝). This proves that the model is suitable for response prediction.

Table 8 – Analysis of variance (ANOVA) for model 4

Source	*DF*	*SS*	*MS*	*F*	*P*
Regression	9	776.28	86.25	11.68	0.000
Residual error	110	812.16	7.38	–	–
Total	119	1588.44	–	–	–

S = 2.71721 R-Sq = 48.9 % R-Sq(adj) = 44.7 %

5) multivariate model 5:

$$H_m = -1.0 - 0.631cosn + 0.633T_{max} + 1.3\frac{\bar{n}}{N} + 6.1\left(\frac{\bar{n}}{N}\right)^2 + +18\frac{T_{max}}{R.H} - 85\left(\frac{T_{max}}{R.H}\right)^2 + 139\left(\frac{T_{max}}{R.H}\right)^3 - 71\left(\frac{T_{max}}{R.H}\right)^4 + 1.58cos^2n. \quad (25)$$

The ANOVA analysis for multivariate model 5 is shown in Table 9.

Table 9 reveals that the model is statistically significant as the p-value of the model is far less than the ∝-value of 0.05 for a confidence interval of 95 %.

Table 9 – Analysis of variance (ANOVA) for model 5

Source	*DF*	*SS*	*MS*	*F*	*P*
Regression	8	768.59	96.07	13.01	0.000
Residual error	111	819.85	7.39	–	–
Total	119	1588.44		–	–

S = 2.71772 R-Sq = 48.4 % R-Sq(adj) = 44.7 %

From the ANOVA results, it is evident that the values of R-Sq. and R-Sq. (adj.) for all the multivariate models were more significant than sunshine and temperature-based models. This implies that the multivariate models have better estimation performance than the sunshine and temperature models for the region under study.

4.4 Graphical plots of the models

To visually compare and contrast the measured and estimated global solar radiation values of the Ihiala region, the following figures show the respective graphical plots for each model.

Figure 1 shows the plot of measured global solar radiation and global solar radiation values estimated using the sunshine-based model.

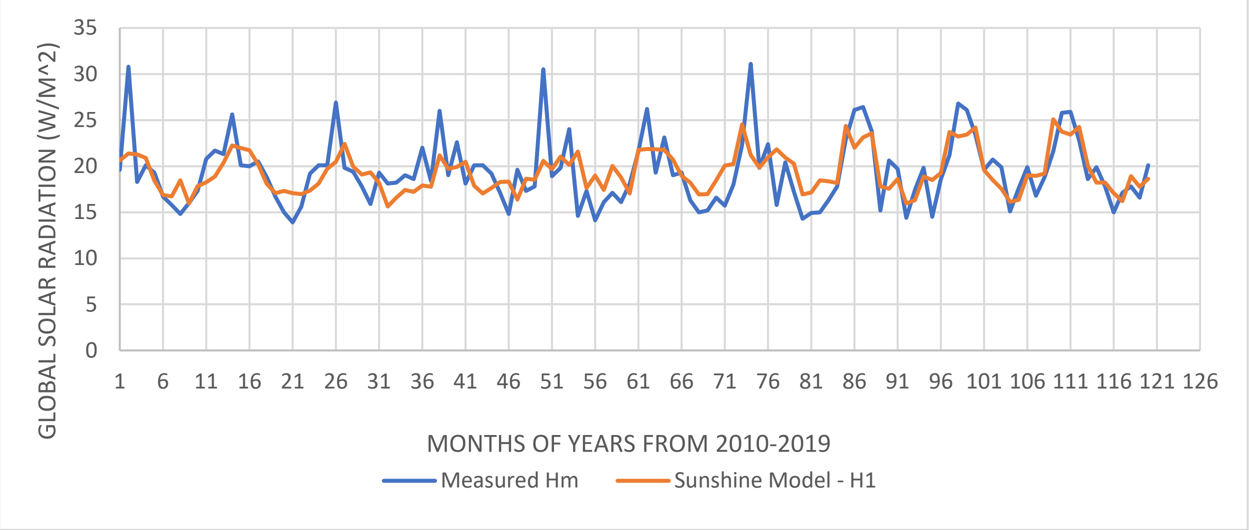

Figure 1 – Comparison of measured and estimated values of global solar radiation using sunshine-based model

From Figure 1, it could vividly be seen that the sunshine-based model did not reasonably estimate the global solar radiation (GSR) data. This is justified because estimated data are dispersed from the measured GSR, as observed in Figure 1.

Figure 2 depicts the graphical plot of the measured and estimated values of GSR using the temperature-based model.

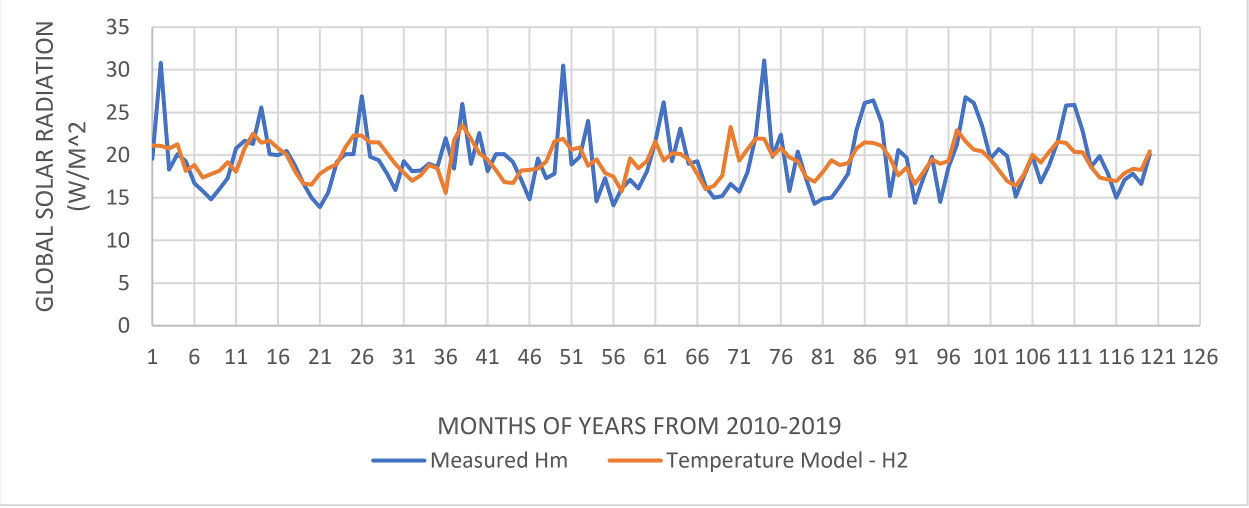

Figure 2 – Comparison of measured and estimated values of global solar radiation using temperature-based model

From Figure 2, the model failed to fit the GSR data well, and as such, the correlative associativity of both data points is reasonably good. The sunshine and temperature models performed poorly in estimating GSR values because they only considered a few regressors/response predictors/independent variables in their respective model formulations. The regressors considered in the Angstrom-Prescot model were only the sunshine index and the extraterrestrial global solar radiation. While in the temperature-based model, the global extraterrestrial solar radiation and the temperature difference were the model's regressors.

Figures 3–7 show the graphical plots of the measured GSR and estimated GSR using the multivariate models (MVM 1, MVM 2, MVM 3, MVM 4, and MVM 5, respectively).

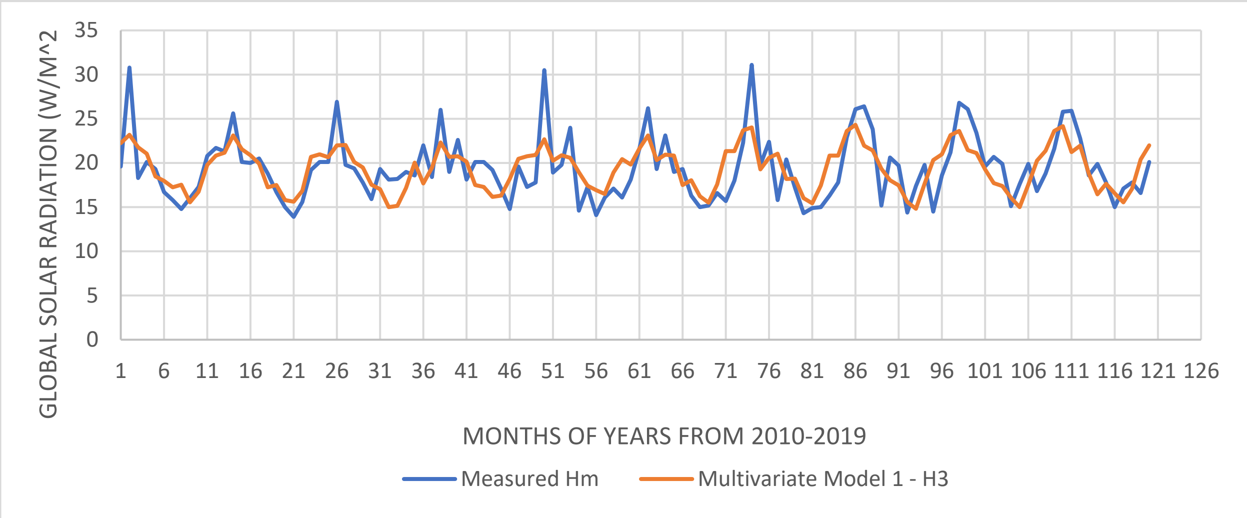

Figure 3 – Comparison of measured and estimated values of global solar radiation using multivariate model 1

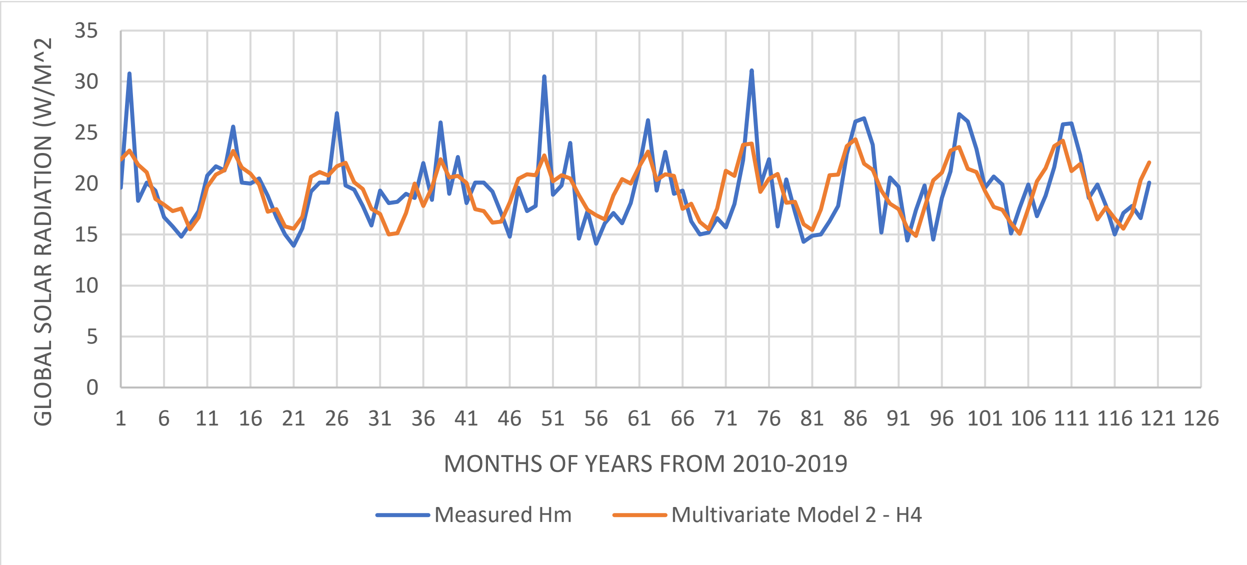

Figure 4 – Comparison of measured and estimated values of global solar radiation using multivariate model 2

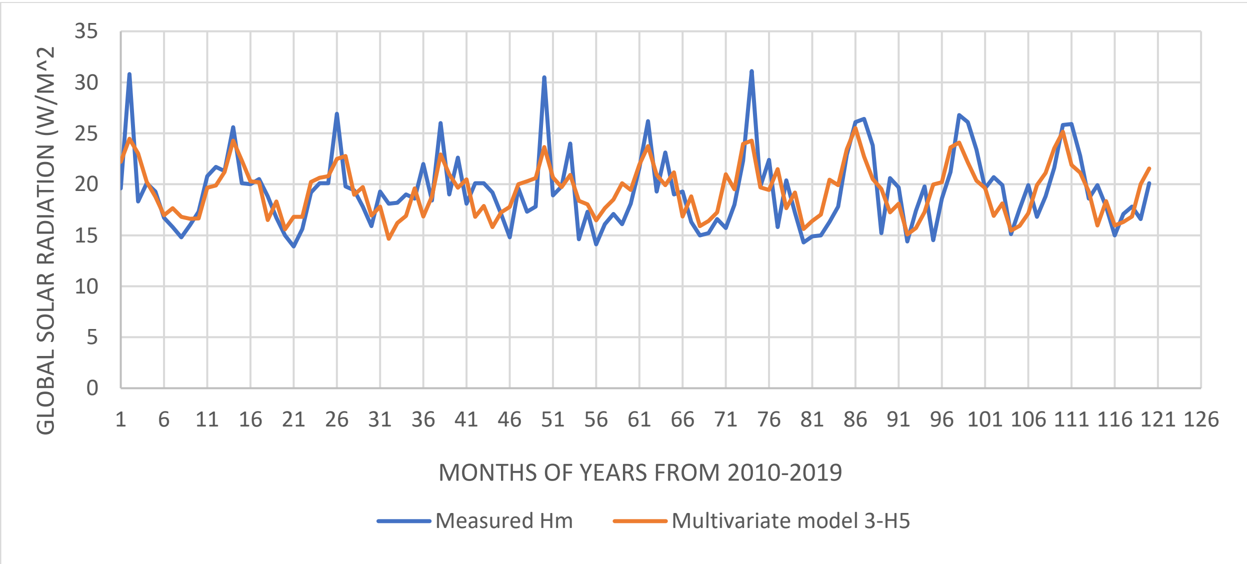

Figure 5 – Comparison of measured and estimated values of global solar radiation using multivariate model 3

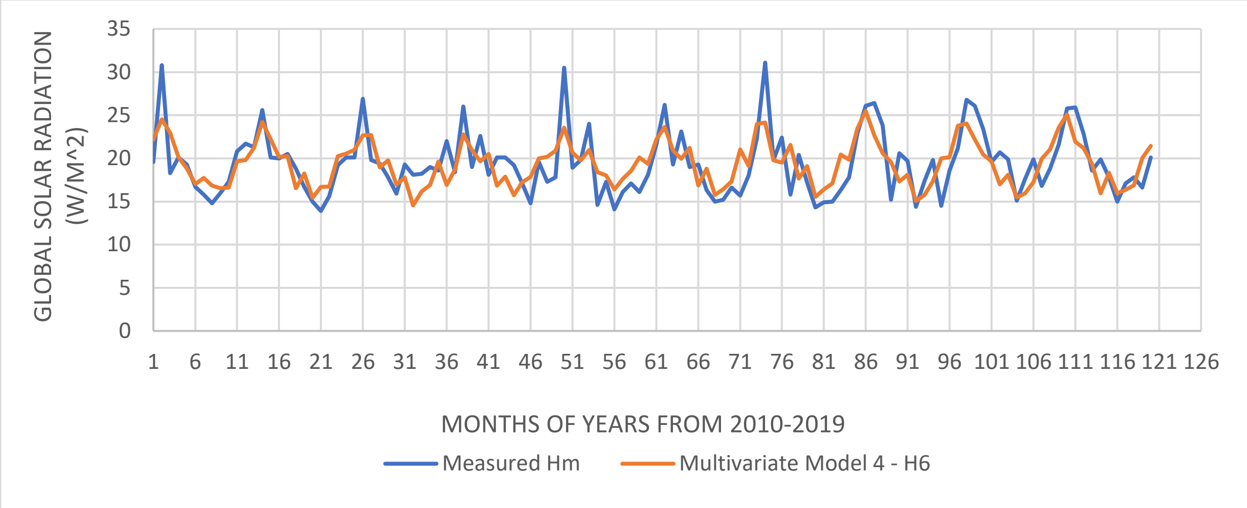

Figure 6 – Comparison of measured and estimated values of global solar radiation using multivariate model 4

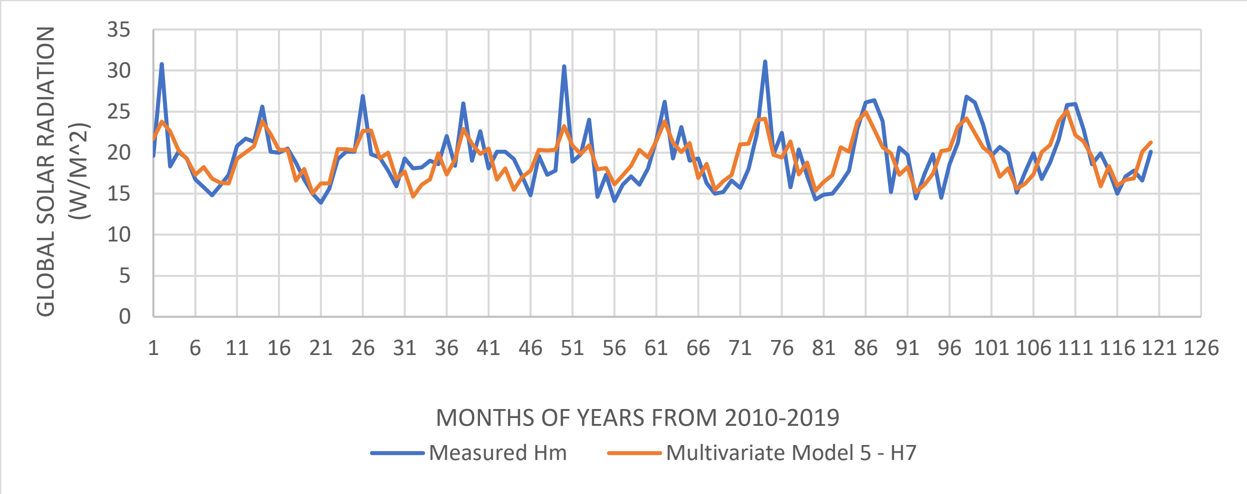

Figure 7 – Comparison of measured and estimated values of global solar radiation using multivariate model 5

From Figures 3–7, it is evident that the estimated and measured values of GSR are more correlated than those of Figures 1, 2. This could be observed from the sequential rise and fall of the data points in the figures.

Figure 6 depicts a greater degree of correlation between the measured and estimated values of GSR gotten using the multivariate model 4. This was because multivariate model 4 had more regressors/ response predictors than the other models.

Table 10 shows the result of the employed statistical validation tools that aided in deciding the best model out of the seven applied empirical equations for estimating the GSR of the Ihiala region.

Table 10 – Statistical validation of the models

Models	MBE	MPE, %	RMSE	NSE	*r*	t-statistictest	R-Square, %
SM	–0.007	1.910	2.819	0.400	0.633	0.026	0.316
TM	–0.021	2.162	2.998	0.321	0.572	0.075	0.221
MVM1	$-3.553\cdot10^{-16}$	1.731	2.691	0.453	0.673	$1.44\cdot10^{-15}$	0.453
MVM2	$8.333\cdot10^{-7}$	1.726	2.688	0.454	0.674	$3.381\cdot10^{-6}$	0.454
MVM3	$4.167\cdot10^{-6}$	1.649	2.603	0.488	0.697	$1.746\cdot10^{-5}$	0.488
MVM4	$3.333\cdot10^{-6}$	1.644	2.602	0.489	0.699	$1.398\cdot10^{-6}$	0.489
MVM5	$-4.293\cdot10^{-16}$	1.647	2.614	0.484	0.696	$1.792\cdot10^{-15}$	0.484

The acronyms used in Table 10 to represent the model names are expressed thus: SM – sunshine-based model; TM – temperature-based model; MVM1–MVM5 – the multivariate models.

Table 10 attests that the multivariate model 4 (MVM 4) performed better than the other models as it best satisfied five (5) out of the seven (7) stated standard statistical tools for model validation:

1) MBE value is lower than those of the other models except MVM2;
2) MPE percentage falls within ±10, and it is the smallest;

3) RMSE is smaller than those of the other models;
4) NSE value is closer to one;
5) coefficient of correlation is higher than those of the other models;
6) T-statistics test is smaller than those of the other models except MVM5 and MVM1;
7) R-Square value is higher than those of the other models.

5 Conclusions

Under using empirical models, the study focused on estimating GSR of the Ihiala region, Anambra state in Sub-Saharan Africa. A total of seven mathematical models were derived using the formulated equations of sunshine, temperature, and multivariate polynomial equations. The obtained regression models were all statistically significant as attested by their respective p-values that were less than the chosen significance level of 5 % for a confidence interval of 95 %.

Despite the models being statistically significant, the multivariate model 4 had a better estimation performance than the other models based on the statistical error indices employed, MBE, MPE, RMSE, NSE. This implies that the multivariate model 4 is suitable for the estimation/prediction of global solar radiation of Ihiala in Sub-Saharan Africa.

6 Acknowledgments

The authors appreciate the Nigerian Meteorological Agency (NIMET), Awka, for providing all the meteorological data employed in this study despite the hurdles encountered during the process. Special thanks to Prof. C. H. Achebe for his invaluable insights that contributed significantly to the success of this scientific article.

References

1. Appelbaum, J. (2001). *Photovoltaics: Present and Future, a Seminar Series*. Katholieke Universiteit, Leuven, Belgium.
2. Mandalia, H. C., Jain, V. K., Pattanaik, B. N. (2012). Application of Super-molecules in solar energy conversion: A review. *Research Journal of Chemical Science*, Vol. 2(1), pp. 89-102.
3. Abdulrahim, A. T., Diso, I. S., El-Jummah, A.M. (2011). Solar concentrators' developments in Nigeria: A review. *Continental Journal of Engineering Sciences*, Vol. 6(3), pp. 30-37.
4. Page, J. K., (1964). The estimation of monthly mean values of daily total short-wave radiation on vertical and inclined surfaces from sunshine records. *Proceeding of the UN Conference on New Sources of Energy*, p. 98.
5. Trabea, A. A., Shaltout M. A., (2000). Correlation of global solar-radiation with meteorological parameters over Egypt. *Renewable Energy*, Vol. 21, pp. 297-308.
6. Dimas, F. A. R., Syed, I. U. H. G., Mohd, S. A. (2011). Hourly solar radiation estimation using ambient temperature and relative humidity data. *International Journal of Environment Science and Development*, Vol. 2(3), pp. 188-193.
7. Khem, P. N., Bhattarai, B. K., Sapkota, B., Kjeldstad, B. (2012). Estimation of global solar radiation using sunshine duration in Himalaya region. *Research Journal of Chemical Sciences*, Vol. 2(11), pp. 20-25.
8. Mbah, O. M., Ozor,P., Mgbemene, C., Enibe, S. O., Mbohwa, C. (2018). Comparison of experimental data and isotropic sky models for global solar radiation estimation in Eastern Nigeria. *World Congress on Engineering, WCE 2018*, paper no. ICME_107.
9. Mbah, O. M., Ozor, P., Mgbemene, C., Enibe, S. O., Mbohwa, C. (2018). Comparative analysis of anisotropic sky models and experimental data in estimating solar radiation on tilted surface in Sub-Saharan African climate. *IEOM Conference. IEOM 2018*, pp. 592-598.
10. Aruna, R. K., Janarthanan, B. (2014). Study of clearness and cloudiness index at tropical locations. *International Journal of Scientific and Engineering Research*, Vol. 5(2), 161.
11. Prescott, J. A. (1940). Evaporation from a water surface in relation to solar radiation. *Journal of Energy Technology*, Vol. 64, pp. 114-148.
12. Page, J. K. (1961). The estimation of monthly mean values of daily total short wave radiation on vertical and inclined surface from sunshine records for latitude 40N-40S. *Proceedings of UN Conference on New Sources of Energy*, Vol. 4(598), pp. 378-390.
13. Cohen, J., Cohen, P., West, S. G., Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, New Jersey, USA, pp. 168-174.
14. Cooper, P. I., (1969). The absorption of radiation in solar stills. *Solar Energy*, Vol. 12(3), pp. 333-346.
15. Duffie, J. A., Beckman, W. A. (2013). *Solar Engineering of Thermal Processes*. John Wiley and Sons, New York, USA.
16. Robaa, S. M. (2008). Evaluation of sunshine duration from cloud data in Egypt. *Energy*, Vol. 33, pp. 785-795.
17. Glover, J., McCulloch, J. S. (1958). The empirical relation between solar radiation and hours of sunshine. *Journal of Royal Meteorological Society*, Vol. 84, pp. 172-175.
18. Chen, R., Ersi, K., Yang, J., Lu, S., Zhao, W. (2004). Validation of five global radiation models with measured daily data in China. *Energy Conversion and Management*, Vol. 45, pp. 1759-1769.