# Empirical Model for Estimating Measured Monthly Average Global Solar Radiation in Lawra using TMY Data

Albert K. Sunnu, Abdul-Rahim Bawa, Adams Yunus, Emmanuel A. Sarsah, Joshua A. Akanbasiam, and Philemon K. Mensah

***Abstract* — In this study, simple and multiple regression models were developed to estimate the monthly average daily global solar radiation in Lawra, Ghana using ground measurement of global horizontal irradiance (Nov 2020–May 2022) and typical meteorological year (TMY) data (Jan 2017–Dec 2019). Various predictor variables such as sunshine ratio, minimum relative humidity and maximum relative humidity ratio, minimum and maximum temperature ratio, etc. were correlated from the TMY data. Many model equations were developed with the variables ranging from one to eight. The best model from each category was chosen and compared using statistical indices to determine the overall best model. We used the JMP statistical software's 'All Possible Models' functionality to select the best model from each category. The selected models where then compared using the adjusted R-squared, mean absolute percentage error, and the root mean square error statistical indices. The best model equation correlated with eight independent variables with adjusted R-squared of 0.99. The equation can be used to estimate monthly global solar radiation in Lawra and in locations with similar climatic conditions where ground measurement of radiation data is unavailable but have access to the National Solar Radiation Database's (NSRDB) TMY data.**

***Key words* —** Model, Regression, Solar Radiation, Statistical.

## I. Introduction

An accurate knowledge of the distribution of solar radiation in a given geographic location is essential for the design of many solar energy devices. Global solar radiation (GSR) is of economic importance as a renewable energy alternative. Global solar radiation has been studied for its importance in providing energy for the Earth's climate system.

Accurate information about solar radiation availability is very important for researchers. GSR data are required for the design of solar energy systems to evaluate their long-term performance [1]. However, the high maintenance and high cost of GSR measuring devices make it very difficult to measure GSR in most areas. The lack of adequate solar radiation data is a common problem in many countries [2].

Numerous methods for predicting GSR have been proposed over the past few decades. These methods include empirical methods [3],[4], machine learning methods and remote sensing methods [5]. These methods can be used to develop mathematical models to estimate GSR using the astronomical, physical, meteorological and geographic parameters recorded at a location. Empirical modeling, however, is an important and cost-effective tool for estimating GSR. As a result, many empirical models have been developed that can be used to estimate GSR using more readily available meteorological parameters such as sunshine hours, number of rainy days, relative humidity, ambient temperature, evaporation, cloud cover and soil temperature.

Bayrakçı *et al.* [1] compared various empirical models and developed new models for estimating GSR on the horizontal surface in the city of Muğla, Turkey. Okundamiya and Nzeako [6] developed a temperature-based empirical model for estimating monthly mean daily global solar radiation on horizontal surfaces for selected cities, representing the six geopolitical zones in Nigeria. Sarsah and Uba [7] developed an empirical model which correlated relative humidity, sunshine duration, and solar declination for estimating GSR in Wa, Ghana.

Zhao *et al.* [8] used sunshine hours and air pollution index to estimate GSR in China. In Jahani, Dinpashoh and Nafchi's work [9], the authors evaluated the accuracy and suitability of eleven models from three different categories for estimating GSR in Iran. Kirmani *et al.* Rizwan [10] used wind speed and precipitation data to estimate monthly GSR in India. The evaluation of the accuracy and applicability of thirteen empirical GSR models for warm subhumid regions that use daily meteorological data at six stations in Mexico were evaluated by Quej *et al.* [11].

In this work, we have developed an empirical model to estimate the monthly average daily GSR from available global horizontal irradiance data measured at the site of the Lawra Solar Power Plant in Ghana and Lawra Typical Meteorological Year (TMY) data available at NSRDB website [12].

Submitted on January 23, 2023.
Published on February 23, 2023.
A. K. Sunnu, Kwame Nkrumah University of Science and Technology, Ghana.
(email: albertsunnu@yahoo.com)
A.-R. Bawa, Dr. Hilla Limann Technical University, Ghana.
(e-mail: kascowemah@yahoo.com)
A. Yunus, Dr. Hilla Limann Technical University, Ghana.
(e-mail: dms_yunus@yahoo.com)
E. A. Sarsah, Dr. Hilla Limann Technical University, Ghana.
(e-mail: emmanuelsarsah@gmail.com)
J. A. Akanbasiam, Dr. Hilla Limann Technical University, Ghana.
(e-mail: ja.akanbasiam@gmail.com)
P. K. Mensah, Dr. Hilla Limann Technical University, Ghana.
(email: philmens34@gmail.com)

TABLE I: MONTHLY AVERAGE DAILY EXTRATERRESTRIAL INSOLATION ON HORIZONTAL SURFACE (MJ/M2 DAY) [14]

| Latitude | Jan 17 | Feb 16 | Mar 16 | Apr 15 | May 15 | Jun 11 | Jul 17 | Aug 16 | Sep 15 | Oct 15 | Nov 14 | Dec 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 60 °S | 41.1 | 31.9 | 21.2 | 10.9 | 4.4 | 2.1 | 3.1 | 7.8 | 16.7 | 28.1 | 38.4 | 43.6 |
| 55 °S | 41.7 | 33.7 | 23.8 | 13.8 | 7.1 | 4.5 | 5.6 | 10.7 | 19.5 | 30.2 | 39.4 | 43.9 |
| 50 °S | 42.4 | 35.3 | 26.3 | 16.8 | 10 | 7.2 | 8.4 | 13.6 | 22.2 | 32.1 | 40.3 | 44.2 |
| 45 °S | 42.9 | 36.8 | 28.6 | 19.6 | 12.9 | 10 | 11.2 | 16.5 | 24.7 | 33.8 | 41.1 | 44.4 |
| 40 °S | 43.1 | 37.9 | 30.7 | 22.3 | 15.8 | 12.9 | 14.1 | 19.3 | 27.1 | 35.3 | 41.6 | 44.4 |
| 35 °S | 43.2 | 38.8 | 32.5 | 24.8 | 18.6 | 15.8 | 17 | 22 | 29.2 | 36.5 | 41.9 | 44.2 |
| 30 °S | 43 | 39.5 | 34.1 | 27.2 | 21.4 | 18.7 | 19.8 | 24.5 | 31.1 | 37.5 | 41.9 | 43.7 |
| 25 °S | 42.5 | 39.9 | 35.4 | 29.4 | 24.1 | 21.5 | 22.5 | 26.9 | 32.8 | 38.1 | 41.6 | 43 |
| 20 °S | 41.5 | 39.9 | 36.5 | 31.3 | 26.6 | 24.2 | 25.1 | 29.1 | 34.2 | 38.5 | 41.1 | 42 |
| 15 °S | 40.8 | 39.7 | 37.2 | 33.1 | 28.9 | 26.8 | 27.6 | 31.1 | 35.4 | 38.7 | 40.3 | 40.8 |
| 10 °S | 39.5 | 39.3 | 37.7 | 34.6 | 31.1 | 29.2 | 29.9 | 32.8 | 36.3 | 38.5 | 39.3 | 39.3 |
| 5 °S | 38 | 38.5 | 38 | 35.8 | 33 | 31.4 | 32 | 34.4 | 36.9 | 38.1 | 37.9 | 37.6 |
| 0 | 36.2 | 37.4 | 37.9 | 36.8 | 34.8 | 33.5 | 33.9 | 35.7 | 37.2 | 37.3 | 36.4 | 35.6 |
| 5 °N | 34.2 | 36.1 | 37.5 | 37.5 | 36.3 | 35.3 | 35.6 | 36.7 | 37.3 | 36.3 | 34.5 | 33.5 |
| 10 °N | 32 | 34.6 | 36.9 | 37.9 | 37.5 | 37 | 37.1 | 37.5 | 37 | 35.1 | 32.5 | 31.1 |
| 15 °N | 29.5 | 32.7 | 35.9 | 38 | 38.5 | 38.4 | 38.3 | 38 | 36.5 | 33.5 | 30.2 | 28.5 |
| 20 °N | 26.9 | 30.7 | 34.7 | 37.9 | 39.3 | 39.5 | 39.3 | 38.2 | 35.7 | 31.8 | 27.7 | 25.7 |
| 25 °N | 24.1 | 28.4 | 33.3 | 37.5 | 39.8 | 40.4 | 40 | 38.2 | 34.7 | 29.8 | 25.1 | 22.9 |
| 30 °N | 21.3 | 26 | 31.6 | 36.8 | 40 | 41.1 | 40.4 | 37.9 | 33.4 | 27.5 | 22.3 | 19.9 |
| 35 °N | 18.3 | 23.3 | 29.6 | 35.8 | 39.9 | 41.5 | 40.6 | 37.3 | 31.8 | 25.1 | 19.4 | 16.8 |
| 40 °N | 15.2 | 20.5 | 27.4 | 34.6 | 39.7 | 41.7 | 40.6 | 36.5 | 30 | 22.5 | 16.4 | 13.7 |
| 45 °N | 12.1 | 17.6 | 25 | 33.1 | 39.2 | 41.7 | 40.4 | 35.4 | 27.9 | 19.8 | 13.4 | 10.7 |
| 50 °N | 9.1 | 14.6 | 22.5 | 31.4 | 38.4 | 41.5 | 40 | 34.1 | 25.7 | 16.9 | 10.4 | 7.7 |
| 55 °N | 6.1 | 11.6 | 19.7 | 29.5 | 37.6 | 41.3 | 39.4 | 32.7 | 23.2 | 13.9 | 7.4 | 4.8 |
| 60 °N | 3.4 | 8.5 | 16.8 | 27.4 | 36.6 | 41 | 38.8 | 31 | 20.6 | 10.9 | 4.5 | 2.3 |

Specifically, we developed models with multiple variables ranging from one to eight variables. The best model from each category was chosen and compared using statistical indices to determine the overall best model.

Using the TMY data has the advantage that the model can be used to estimate GSR in locations of similar climatology, latitude and altitude as Lawra that do not have access to ground measurement data. TMY is a compilation of selected weather data for a specific location, listing values of solar radiation and meteorological elements for a period of one year [13]. The values are generated from a database that contain at least 12 years data. It is specifically chosen to reflect the range of weather phenomena for each location and still provide annual averages that are consistent with the long-term averages of each location.

## II. METHODOLOGY

Ground measurements of global horizontal irradiance (GHI) in W/m$^2$ at 5-minute intervals were obtained from the Volta River Authority's Lawra Solar Power Plant (LSPP) in Ghana. 15-minute interval TMY data for Lawra was obtained from the NSRDB website. The LSPP data was from November 2020 to May 2022, and the TMY data was from January 2017 to December 2019. The geographical location of Lawra is 10.6°N, 2.8°W.

Using the *tabulate* feature in the JMP software, we obtained the monthly daily averages of GHI, temperature and relative humidity. The GHI values were multiplied by 0.08 MJ/m$^2$ to get the monthly global solar radiation, $\bar{H}$ in MJ/(m$^2$ day) since 1 W/m$^2$ = 0.086 MJ/m$^2$. Monthly mean daily extraterrestrial radiation, $\bar{H}_0$ (MJ/m$^2$ day) data were obtained from Table I. We performed linear interpolation using the latitude of Lawra (10.6°N) between 10°N and 15°N in Table I.

Monthly mean daily declination angle, $\delta$ using the average day of each month was calculated using (1).

$$\delta = 23.45 \sin\left[\frac{360}{365}(284 + d)\right] \quad (1)$$

Where $d$ is the day of the year obtained from Table II with Table I as reference. Monthly mean daily hours of sunshine ratio, $\bar{n}/\bar{N}$ data were obtained at Wa from the work of Sarsah and Uba [7]. The geographical location of Wa is 10.06°N, 2.5°W and has similar climatic conditions as Lawra.

Monthly mean daily maximum temperature $\bar{T}_{max}$ (°C), minimum temperature $\bar{T}_{min}$ (°C), maximum relative humidity $\overline{RH}_{max}$ (%), and minimum relative humidity $\overline{RH}_{min}$ (%), were also obtained using the JMP software *tabulate* feature via the maximum and minimum statistical measure of the feature.

TABLE II: DETERMINATION OF D

| Month | $d$ |
|---|---|
| January | 17 |
| February | 31 + 16 |
| March | 59 + 16 |
| April | 90 + 15 |
| May | 120 + 15 |
| June | 151 + 11 |
| July | 181 + 17 |
| August | 212 + 16 |
| September | 243 + 15 |
| October | 273 + 15 |
| November | 304 + 14 |
| December | 334 + 10 |

The following model terms were then formulated: $\overline{\Delta T} = \bar{T}_{max} - \bar{T}_{min}$, $\bar{n}/\bar{N}$, $(\bar{n}/\bar{N})^2$, $(\bar{n}/\bar{N})^3$, $\log(\bar{n}/\bar{N})$, $\exp(\bar{n}/\bar{N})$, $\ln(\bar{n}/\bar{N})$, $\bar{\delta}/\bar{N}$, $(\overline{\Delta T})^2$, $(\overline{\Delta T})^3$, $\ln(\bar{T})$, $\log(\overline{\Delta T}/\bar{N})$, $\bar{\theta} = \bar{T}_{min}/\bar{T}_{max}$, $\bar{T}_{max}/\cos(L)$, $\cos(L) + (\bar{n}/\bar{N})$, $\bar{n}/(\bar{N} \times \bar{\theta})$, $(\overline{RH})^2$, $\overline{RH}_{min}/\overline{RH}_{max}$, $\bar{\theta} \times \overline{RH}$, $\overline{RH} \times \bar{T}$, $(\bar{n}/\bar{N}) - \overline{RH} - \bar{\theta}$, $\bar{T}_{max}/\overline{RH}$, $\sin(\bar{\delta})$, $\cos(L)\cos(\bar{\delta})$, and $\sin(L)\sin(\bar{\delta})$.

To find out which of the model terms exhibit a high level of multicollinearity, we used the *multivariate* functionality in the JMP software, and evaluated the correlation between

the terms. Since our primary goal was to develop a model for making predictions, we only removed some of the terms that had correlations of above 0.99.

The *all-possible model's* method via the *stepwise functionality* in the JMP software was then used to select the best model in each variable category (from 1 to 8) using the ratio $\overline{H}/\overline{H}_0$ as the dependent variable and the model terms as the independent variables. The model with the highest coefficient of determination, $R^2$ for each number of parameters was selected as the best in that category. The $R^2$ values were not used in the overall comparison of the selected models. The statistical measures adjusted R-squared $R_a^2$, mean absolute percentage error (MAPE), and root mean square error (RMSE) were used to compare the selected models.

$$R_2^a = 1 - \left[\frac{(1-R^2)\times(n-1)}{(n-k-1)}\right] \quad (2)$$

Where $n$ is the number of data points, and $k$ is the number of independent variables. Generally, the model with a higher $R_a^2$ fits the data well.

$$MAPE(\%) = \frac{100}{n}\sum_{i=1}^{n}\left(\left|\frac{\overline{H}_{i,c}-\overline{H}_{i,m}}{\overline{H}_{i,m}}\right|\right) \quad (3)$$

A MAPE less than 5% is considered as an indication that the estimated value is acceptably accurate.

TABLE II: SUMMARY

| Month | $\overline{H}$ (MJ/m² day) | $\overline{H}_o$ (MJ/m² day) | $\overline{H}/\overline{H}_o$ | $\overline{T}_{max}$ (°C) | $\overline{T}_{min}$ (°C) | $\overline{T}$ | $\overline{RH}_{max}$ (%) | $\overline{RH}_{min}$ (%) | $\overline{RH}$ (%) | $\overline{n}/\overline{N}$ | $\overline{\delta}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| January | 20.00 | 31.70 | 0.63 | 37.37 | 14.17 | 25.26 | 74.19 | 7.30 | 31.80 | 0.48 | -21.83 |
| February | 21.95 | 34.37 | 0.64 | 39.90 | 17.33 | 28.39 | 89.54 | 5.73 | 30.80 | 0.48 | -21.80 |
| March | 20.99 | 36.78 | 0.57 | 41.03 | 18.97 | 30.45 | 94.48 | 6.95 | 45.22 | 0.51 | -21.76 |
| April | 21.09 | 37.91 | 0.56 | 40.00 | 21.87 | 30.14 | 96.78 | 13.69 | 61.57 | 0.51 | -21.71 |
| May | 20.42 | 37.62 | 0.54 | 38.70 | 22.27 | 28.55 | 100.00 | 32.64 | 73.07 | 0.49 | -21.68 |
| June | 18.54 | 37.17 | 0.50 | 36.07 | 21.17 | 26.91 | 100.00 | 41.46 | 80.37 | 0.44 | -21.66 |
| July | 16.60 | 37.24 | 0.45 | 31.93 | 20.33 | 25.72 | 100.00 | 56.16 | 84.71 | 0.41 | -21.67 |
| August | 14.38 | 37.56 | 0.38 | 31.40 | 20.37 | 25.16 | 100.00 | 59.01 | 87.05 | 0.37 | -21.70 |
| September | 16.88 | 36.94 | 0.46 | 31.93 | 20.50 | 25.59 | 100.00 | 57.05 | 86.70 | 0.4 | -21.74 |
| October | 19.94 | 34.91 | 0.57 | 34.07 | 20.37 | 26.19 | 100.00 | 44.50 | 81.79 | 0.5 | -21.79 |
| November | 20.79 | 32.22 | 0.65 | 34.87 | 16.57 | 26.55 | 97.87 | 23.87 | 65.61 | 0.43 | -21.82 |
| December | 17.96 | 30.79 | 0.58 | 34.67 | 13.47 | 24.11 | 85.07 | 15.65 | 48.21 | 0.46 | -21.84 |

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\overline{H}_{i,c}-\overline{H}_{i,m}\right)^2} \quad (4)$$

The smaller the value of RMSE, the better is the model's performance. The model equations were then used to estimate $\overline{H}$ for each month and compared to the measured values using the statistical tests.

## III. RESULTS AND DISCUSSION

In Table II are the results of the monthly daily averages from the *tabulate* feature. The highest $\overline{H}$ was recorded in February whereas August recorded the lowest. The results indicate that for the period of data obtained, rainfall in Lawra is at its peak during the month of August when the sky is cloudy and solar radiation is fairly low, whereas solar radiation is at its peak in February.

It was found from the correlation analysis that some terms were highly correlated and were removed leaving only one of such terms for the final analysis. The correlated terms and the ones removed are in Table III. In Table IV are the best model equations from each category of variables.

TABLE III: HIGHLY CORRELATED TERMS

| Term | Correlated terms that were removed |
|---|---|
| $\overline{\Delta T}$ | $(\overline{\Delta T})^2$, $\log(\overline{\Delta T}/\overline{N})$, $\bar{\theta}\times\overline{RH}$. |
| $\overline{n}/\overline{N}$ | $(\overline{n}/\overline{N})^2$, $(\overline{n}/\overline{N})^3$, $\log(\overline{n}/\overline{N})$, $\exp(\overline{n}/\overline{N})$, $\ln(\overline{n}/\overline{N})$, $\cos(L)+(\overline{n}/\overline{N})$. |
| $\bar{\delta}/\overline{N}$ | $\bar{\delta}$, $\sin(\bar{\delta})$, $\cos(L)\cos(\bar{\delta})$, $\sin(L)\sin(\bar{\delta})$. |
| $(\overline{\Delta T})^3$ | $(\overline{RH})^2$, $(\overline{n}/\overline{N})-\overline{RH}-\bar{\theta}$. |

The actual by predicted plots of $\overline{H}/\overline{H}_0$ for the various models are in Fig. 1-8. $\overline{H}/\overline{H}_0$ is labelled *"Y"* in the figures. For a good fit, the data points should be close to the main diagonal line. It is seen from the figures that (8) is a very good fit for the response variable. All the models are also statistically significant (at the 0.05 significance level) as can be seen from the *p*-values in the figures.

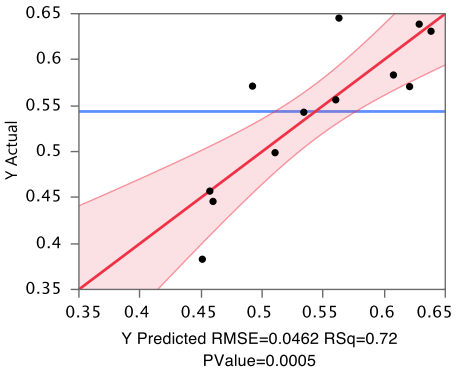


Fig. 1. Actual by predicted $\overline{H}/\overline{H}_0$ for model 1.

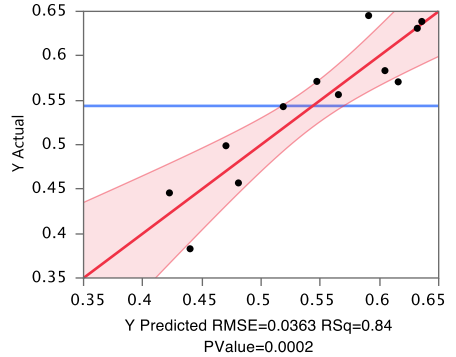


Fig. 2. Actual by predicted $\overline{H}/\overline{H}_0$ for model 2.

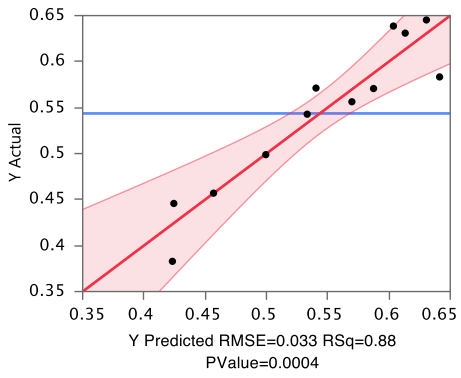


Fig. 3. Actual by predicted $\overline{H}/\overline{H}_0$ for model 3.

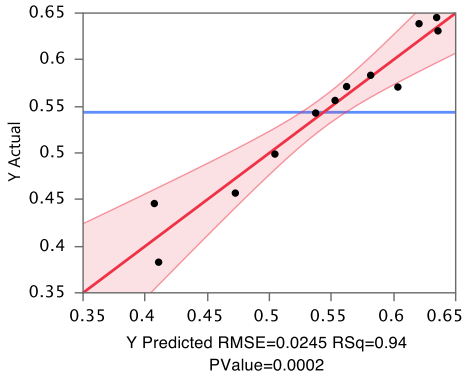


Fig. 4. Actual by predicted $\overline{H}/\overline{H}_0$ for model 4.

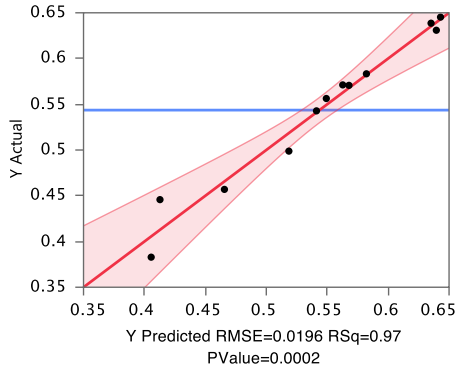


Fig. 5. Actual by predicted $\overline{H}/\overline{H}_0$ for model 5.

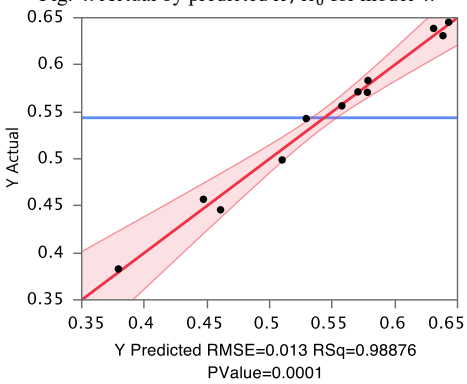


Fig. 6. Actual by predicted $\overline{H}/\overline{H}_0$ for model 6.

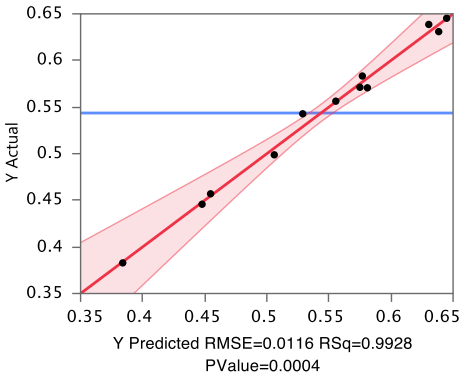


Fig. 7. Actual by predicted $\overline{H}/\overline{H}_0$ for model 7.

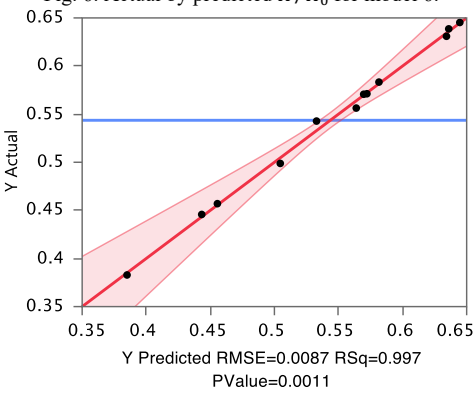


Fig. 8. Actual by predicted $\overline{H}/\overline{H}_0$ for model 8.

Estimated values of $\overline{H}$ using the regression models compared to the measured values are in Table V whereas the statistical error tests of the models are in Table VI.

The MAPE of all models were less than 5% except for (1) which was a little greater than 5%. Model (8) had the best MAPE score of 0.6%. For the RMSE test, (8) had the lowest value of 0.009 whereas (1) had the highest value of 0.046. The $R_a^2$ value of (8) was the highest whereas (1) was the lowest. From the results of the statistical measures, it can be deduced that (8) passed all tests and is therefore selected as the best model. The plot of measured $\overline{H}$ and (8) is as in Fig. 9.

TABLE IV: BEST MODELS FROM VARIOUS CATEGORIES

| Model | Equation | No. of terms |
|---|---|---|
| 1 | $\frac{\overline{H}}{\overline{H}_0} = 0.28 + 0.015\overline{\Delta T}$ | 1 |
| 2 | $\frac{\overline{H}}{\overline{H}_0} = -1.32 - 0.76\frac{\bar{\delta}}{\overline{N}} + 0.013\frac{\overline{T}_{max}}{\cos(L)}$ | 2 |
| 3 | $\frac{\overline{H}}{\overline{H}_0} = -0.87 + 0.04\overline{\Delta T} - 0.489\frac{\bar{\delta}}{\overline{N}} - 0.0000373(\Delta T)^3$ | 3 |
| 4 | $\frac{\overline{H}}{\overline{H}_0} = -1.22 - 0.75\frac{\bar{\delta}}{\overline{N}} - 1.52\bar{\theta} + 0.00051(\overline{RH} \times \overline{T}) + 0.51\frac{\overline{T}_{max}}{\overline{RH}}$ | 4 |

*Cont. of Table IV*

| Model | Equation | No. of terms |
|---|---|---|
| 5 | $\frac{\bar{H}}{\bar{H}_0} = -0.597 - 0.623\frac{\bar{\delta}}{\bar{N}} - 2.598(\overline{\Delta T})^3 - 2.324\bar{\theta} + 0.000548(\overline{RH} \times \bar{T}) + 0.693\frac{\bar{T}_{max}}{\overline{RH}}$ | 5 |
| 6 | $\frac{\bar{H}}{\bar{H}_0} = 0.357 - 0.141\overline{\Delta T} + 0.878\frac{\bar{n}}{\bar{N}} + 1.35\ln(\bar{T}) - 7.444\bar{\theta} + 0.00064(\overline{RH} \times \bar{T}) + 0.924\frac{\bar{T}_{max}}{\overline{RH}}$ | 6 |
| 7 | $\frac{\bar{H}}{\bar{H}_0} = -0.0988 - 0.1156\overline{\Delta T} + 0.719\frac{\bar{n}}{\bar{N}} - 0.209\frac{\bar{\delta}}{\bar{N}} + 1.1009\ln(\bar{T}) - 6.379\bar{\theta} + 0.00064(\overline{RH} \times \bar{T}) + 0.8742\frac{\bar{T}_{max}}{\overline{RH}}$ | 7 |
| 8 | $\frac{\bar{H}}{\bar{H}_0} = 0.546 - 0.1656\overline{\Delta T} + 0.5766\frac{\bar{n}}{\bar{N}} - 0.288\frac{\bar{\delta}}{\bar{N}} + 1.1398\ln(\bar{T}) - 7.757\bar{\theta} + 0.0247\frac{\bar{T}_{max}}{\cos(L)} + 0.00055(\overline{RH} \times \bar{T}) + 0.834\frac{\bar{T}_{max}}{\overline{RH}}$ | 8 |

TABLE V: COMPARISON OF MEASURED AND ESTIMATED $\bar{H}$

| Month | $\bar{H}_{measured}$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|---|
| January | 20.00 | 20.24 | 20.03 | 19.44 | 20.15 | 20.27 | 20.24 | 20.24 | 20.11 |
| February | 21.95 | 21.61 | 21.85 | 20.75 | 21.33 | 21.83 | 21.68 | 21.67 | 21.86 |
| March | 20.99 | 22.83 | 22.65 | 21.60 | 22.20 | 20.89 | 21.26 | 21.37 | 20.95 |
| April | 21.09 | 21.24 | 21.43 | 21.60 | 20.96 | 20.82 | 21.14 | 21.06 | 21.38 |
| May | 20.42 | 20.09 | 19.51 | 20.06 | 20.21 | 20.36 | 19.91 | 19.90 | 20.05 |
| June | 18.54 | 18.98 | 17.48 | 18.56 | 18.75 | 19.27 | 18.96 | 18.81 | 18.76 |
| July | 16.60 | 17.12 | 15.74 | 15.80 | 15.17 | 15.37 | 17.16 | 16.68 | 16.51 |
| August | 14.38 | 16.94 | 16.53 | 15.90 | 15.43 | 15.23 | 14.24 | 14.42 | 14.46 |
| September | 16.88 | 16.89 | 17.76 | 16.87 | 17.46 | 17.20 | 16.52 | 16.80 | 16.83 |
| October | 19.94 | 17.18 | 19.10 | 18.87 | 19.63 | 19.65 | 19.92 | 20.07 | 19.97 |
| November | 20.79 | 18.14 | 19.04 | 20.32 | 20.45 | 20.72 | 20.72 | 20.78 | 20.77 |
| December | 17.96 | 18.71 | 18.62 | 19.75 | 17.91 | 17.93 | 17.81 | 17.76 | 17.91 |

TABLE VI: STATISTICAL MEASURES OF THE MODEL EQUATIONS

| Model | MAPE (%) | RMSE (MJ/m$^2$ day) | $R_a^2$ |
|---|---|---|---|
| 1 | 5.7 | 0.046 | 0.69 |
| 2 | 5.2 | 0.036 | 0.81 |
| 3 | 4.0 | 0.033 | 0.84 |
| 4 | 2.9 | 0.025 | 0.91 |
| 5 | 2.1 | 0.020 | 0.94 |
| 6 | 1.4 | 0.013 | 0.98 |
| 7 | 1.0 | 0.012 | 0.98 |
| 8 | 0.6 | 0.009 | 0.99 |

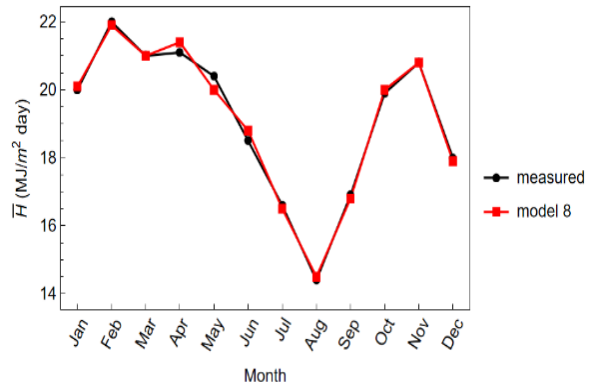


Fig. 9. Comparison of measured and estimated $\bar{H}$.

## IV. CONCLUSION

The objective of this study was to use TMY data to develop an empirical model for estimating measured monthly mean daily solar radiation in Lawra using simple and multiple regression analysis. Many model terms were formulated, however, eleven were selected for the final analysis after multicollinearity test, and the best equation correlated eight of these terms. The statistical software package JMP was used to find all possible models using the terms formulated, from which the best models were selected from each category using the $R^2$ statistical measure for further analysis. All models were statistically significant, and the final model correlated with eight terms and had an adjusted $R^2$ of 0.99. The model can therefore be used to estimate monthly mean daily solar radiation at Lawra in the Upper West Region of Ghana using TMY data, and elsewhere with similar climatic conditions where ground measurements of radiation data are not available. When more data is available, the methodology used in this study can be used to find other models.

## ACKNOWLEDGMENT

The authors would like to thank the Deputy Chief Executive Officer of VRA for providing the data used in this study.

## CONFLICT OF INTEREST

The authors declare that the there is no conflict of interests regarding the publication of this paper.

## REFERENCES

[1] Bayrakçı HC, Demircan C, Keçebaş A. The development of empirical models for estimating global solar radiation on horizontal surface: A case study. *Renewable and Sustainable Energy Reviews*, 2018 Jan;81:2771–82.

[2] Mecibah MS, Boukelia TE, Tahtah R, Gairaa K. Introducing the best model for estimation the monthly mean daily global solar radiation on a horizontal surface (Case study: Algeria). *Renewable and Sustainable Energy Reviews*. 2014 Aug;36:194–202.

[3] Fan J, Wu L, Zhang F, Cai H, Ma X, Bai H. Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of China. *Renewable and Sustainable Energy Reviews*. 2019 May;105:168–86.

[4] Zhang Q, Cui N, Feng Y, Jia Y, Li Z, Gong D. Comparative Analysis of Global Solar Radiation Models in Different Regions of China. *Advances in Meteorology*, 2018;2018:1–21.

[5] Feng Y, Gong D, Zhang Q, Jiang S, Zhao L, Cui N. Evaluation of temperature-based machine learning and empirical models for predicting daily global solar radiation. *Energy Conversion and Management,* 2019; 198, 111780. https://doi.org/10.1016/j.enconman.2019.111780.

[6] Okundamiya MS, Nzeako AN. Empirical Model for Estimating Global Solar Radiation on Horizontal Surfaces for Selected Cities in the Six Geopolitical Zones in Nigeria. *Journal of Control Science and Engineering*, 2011; 2011: 1–7.

[7] Sarsah EA, Uba FA. Empirical correlations for the estimation of global solar radiation using meteorological data in Wa, Ghana. *Pelagia Research Library*. 2013;4(4):63–71.

[8] Zhao N, Zeng X, Han S. Solar radiation estimation using sunshine hour and air pollution index in China. *Energy Conversion and Management*, 2013;76:846–51.

[9] Jahani B, Dinpashoh Y, Raisi Nafchi A. Evaluation and development of empirical models for estimating daily solar radiation. *Renewable and Sustainable Energy Reviews*. 2017;73:878–91.

[10] Kirmani S, Jamil M, Rizwan M. Empirical correlation of estimating global solar radiation using meteorological parameters. *International Journal of Sustainable Energy*. 2013;34(5):327–39.

[11] Quej VH, Almorox J, Ibrakhimov M, Saito L. Empirical models for estimating daily global solar radiation in Yucatán Peninsula, Mexico. *Energy Conversion and Management*. 2016; 110: 448–56.

[12] NSRDB [Internet]. nsrdb.nrel.gov. [cited 2023 Jan 20]. Available from: https://nsrdb.nrel.gov/data-viewer/.

[13] Typical meteorological year [Internet]. Wikipedia. 2023 [cited 2023 Jan 20]. Available from: https://en.wikipedia.org/wiki/Typical_meteorological_year.

[14] Kalogirou SA. *Solar Energy Engineering: Processes and Systems*. Amsterdam; Boston; Heidelberg Etc: Academic Press; 2014.