Contents list available at IJRED website

Int. Journal of Renewable Energy Development (IJRED)

Journal homepage: https://ijred.undip.ac.id

Research Article

A Comparative Study of Regression Models and Meteorological Parameters to Estimate the Global Solar Radiation on a Horizontal Surface for Baghdad City, Iraq

Moafaq K.S. Al-Ghezi[a]*, Bashar K. Mahmoud[b], Tamadher M.A. Alnasser[c], Miqdam Tariq Chaichan[a]

[a]*Energy and Renewable Energies Technology Center, University of Technology, Baghdad, Iraq*

[b]*Ministry of Higher Education and Scientific Research, Baghdad, Iraq*

[c]*Mechanical Eng. Dept., University of Technology, Baghdad, Iraq*

Abstract. In this study, data of the monthly average of daily global solar radiation falling on a horizontal surface, relative humidity, maximum temperature, and duration of sunshine for the city of Baghdad were collected through two sources. First, from the Iraqi Meteorological Authority (IMA) for a period extending from 1961 to 2016. The second is from NASA, for the period from 1984 to 2004. Then, four linear regression models, two single and two polynomials were formulated to calculate the values of the monthly average of daily global horizontal solar radiation (GHSR) incidents. The models calculated the monthly average of daily extraterrestrial radiation and day length, using some data provided by NASA and the IMA. To ensure the validity of the used models, a statistical test was performed for the performance of the proposed models, using the indicators mean bias Error (MBE), root mean square error (RMSE) as well as mean percentage error (MPE). The validation shows the relationship between the measured and computed values (through the analysis of the results), where a great convergence was found between the measured and calculated values. This means that the proposed models can be adapted to predict global solar radiation. The highest values of measured solar radiation were during the month of June, which were 28.555 and 27.280 MJ/m^2/day from the IMA and NASA, respectively. The same applies to the radiation calculated using the four empirical models. The month of June was the highest in terms of solar radiation values. The radiation values were 28.947, 26.315, 29.699, and 26.716 MJ/m^2/day for the first, second, third, and fourth models, respectively. The lowest values of measured and calculated radiation were during the month of December. Always, radiation measured by the IMA was greater than those of NASA, as well as the values of radiation calculated in the two IMA-based models were greater than the other two NSA-based models. In the absence of a method for measuring the diffuse and direct (beam) solar radiations, as well as the lack of such values by meteorological authorities, and its paramount importance, they were reported to mathematically calculate them in this study. The values of statistical indicators RMSE; MJ/m^2/day, MBE; MJ/m^2/day and MPE% were (0.4769, 0.0164, 0.2207), (0.8641, 0.1773, -0.9680), (0.6420, 0.3996, -1.1487), (0.9604, 0.218, -1.0225) for the first, second, third and fourth models, respectively. According to the results of the statistical test, it can be indicated that the single linear regression model, based on the IMA's data (model No.1), is the most accurate to calculate global solar radiation for Baghdad City.

Keywords: Global solar radiation, Iraqi meteorological authority, linear regression model, meteorological Data, NASA, sunshine duration.

Article History: Received: 21st May 2021; Revised: 6th August 2021; Accepted: 1st Sept 2021; Available online: 20th Sept 2021
How to Cite This Article: Al-Ghezi, M.K.S., Mahmoud, B.K., Alnasser, T.M.A., Chaichan, M.T. (2022) A Comparative Study of Regression Models and Meteorological Parameters to Estimate the Global Solar Radiation on a Horizontal Surface for Baghdad city, Iraq. *Int. J. Renew. Energy Dev*, 11(1), 71-81
https://doi.org/10.14710/ijred.2022.38493

1. Introduction

The sun is the main source of all kinds of energies on the surface of the earth. All types of energies are directly or indirectly generated by the influence of the sun. For instance, but not limited to, wind energy is generated by the difference in pressure between two geographical regions, and this pressure difference is affected by temperature difference due to the effect of the sun; this makes the winds move from high-pressure areas towards the lower pressure areas. Another example is the tidal energy that is produced by the effect of moon gravity, which is originally a reflection of the movement of the sun's mass. There are also other types of energies that are related to the sun (Trenberth *et al.*, 2009). Interest in solar energy is focussed on its promising future on the surface of the earth, due to its abundance, lack of depletion, and being environmentally friendly. On the other hand, there is the inability to rely on fossil fuels currently, which represents the first source of energy at present, as it is expected to decrease its stocks significantly in the next few decades. In addition to this, there is the onslaught of

* Corresponding author: moafaq.k.shiea@uotechnology.edu.iq

global warming and negative impacts on the environment, due to the emissions of toxic gases associated with burning oil, gas, coal and other types of fuels (Chaichan *et al.*, 2018; Moafaq, 2019). Solar energy reaches the earth through solar radiation, whose intensity on the earth's surface varies from one region to another, as a result of many factors, the most important being geographic factors such as longitude, latitude, height above sea level, etc. The other factors include weather conditions such as rainfall level and wind speed. Furthermore, the amount of dust, relative humidity and temperature, length of daylight hours, clarity of the weather, etc., also affect solar radiation (El-Sebaii and Trabea, 2005; Moafaq, 2017a; Ertekin and Yaldiz, 2000).

Energy demand is expected to double by 2030, and quadruple by 2050, compared to the year 2000. There are great ambitions to contribute renewable energies, especially solar energy, to meet these needs (Chaichan and Kazem, 2018). Solar energy is used in several applications, including heating homes, generating electricity, water heating, drying crops, water distillation and others (Chaichan and Kazem, 2018). There are many technological applications that can benefit from solar energy, the most important of which are thermal solar concentrators, photovoltaic panels, solar water heaters, etc. (Rabaia *et al.,* 2021).

For projects, which depends on solar energy, sufficient information must be available about the solar radiation values for the selected region, so these values are obtained either by employing measurements or theoretical calculations (Abdelhafidi *et al.,* 2021).

Many meteorological stations worldwide use different devices to measure weather parameters such as temperature, relative humidity, solar radiation intensity, etc. (Bailek *et al.,* 2020). However, these stations do not cover the entire region of the world. Also, the measured values cannot be assured of their reliability due to several reasons, the most important of which are the accuracy of the devices used in the measurement, and the methods used to calibrate these devices.

These issues have prompted researchers from different continents of the world to find other methods for calculating weather parameters. The most important weather parameter is solar radiation, and comparing it with the measured values of these stations to ensure their reliability (Kazem *et al.*, 2016; Yousif *et al.*, 2019).

Iraq is characterised by sunny weather most days, except for some rainy and cloudy winter days. The number of sunny days reach more than 300 days during the year. In addition, the intensity of Iraq's solar radiation is relatively high, compared to many regions of the world, which provides a fertile environment and a promising future for the use of solar energy technology in various applications in the country (Chaichan *et al.*, 2018; Moafaq *et al.*, 2021).

Through the investigation of previous studies and looking at process of calculating the intensity of solar radiation, it was found that most of them depended on one type or several weather parameters (Bamehr and Sabetghadam, 2021). Some of these studies have been based on mathematical modeling on daily sunshine duration, relative humidity, the number of sunny hours, the values of latitude, the angle of inclination or maximum temperature, height above sea level, and geographical location (Sridharan, 2021; Mujabar and Venkateswara, 2021; Blal *et al.*, 2020). Additionally, other studies were based on surface albedo and turbidity of water (Liu *et al.*, 2020; Yu *et al.*, 2021). Besharat *et al.* (Besharat *et al.*, 2013) reviewed the literature related to the study of empirical models for calculating solar radiation. These models have been classified into four groups, depending on the parameters used in these models, as follows: models that are depended on temperature, models that are based on sunshine duration, models that that are based on the number of clouds and models that are based on other weather parameters. These models were applied and compared using known statistical indicators, and the measured data from the weather station of the Iranian Yazd city.

Koussa *et al.* (Koussa *et al.*, 2009) studied several statistical models to calculate the global and diffuse solar radiation for three cities in Algeria. The authors recommended the necessity of formulating models that calculate solar radiation in other Algerian cities without weather stations that provide solar radiation data.

Sen (Sen, 2007) developed the Angstrom equation to formulate a non-linear model based on the sunshine duration to calculate global solar radiation. This model includes three variables, all of which depend on the duration of the sunshine, and he concluded with the non-linear relationship between sunshine duration and global solar radiation. Benghanem and Joraid (Benghanem and Joraid, 2007) calculated the global and diffuse solar radiations for Medina city in Saudi Arabia, using multiple proposed correlations to find a relationship between the diffuse and global solar radiation. These empirical correlations were of the first, second and third-order for sunshine duration. Sabbagh *et al.* (Sabbagh *et al.,* 1977) proposed empirical models for calculating solar radiation for several countries in the Middle East, such as Egypt, Saudi Arabia, Lebanon, Kuwait, and Sudan.

The current work aims to provide reliable values of solar radiation for the city of Baghdad, which is the capital of Iraq; Baghdad is located in the middle of Iraq at a latitude of 33.20°, longitude 44.26° and 34 m height above sea level, by developing four linear correlation models to calculate the intensity of solar radiation throughout the year. All models are compared employing the measured values obtained from the Baghdad Meteorological Station and satellite data provided by NASA.

Investigating previous studies did not yield any reliable information for calculating the solar radiation for Baghdad city. This study is considered new in collecting and using meteorological data from two different sources: the Iraqi Meteorological Authority (IMA) and NASA. These data were used to formulate four new empirical models for calculating global solar radiation falling on horizontal surfaces. The empirical models were two singular and two polynomial models, depending on the data source used to formulate each model. The global solar radiation calculated by each of the four empirical models was compared with the global solar radiation values measured by the IMA and NASA, and most often showed good compatibility. In order to increase the reliability of the results obtained by the four models, statistical error indicators such as mean bias error (MBE), root mean square error (RMSE) and mean percentage error (MPE) were calculated. The results were excellent for all four empirical models, as any of these models can be employed to estimate solar radiation reliably. Moreover, the absence of data about beam and diffuse solar radiation from weather stations, and the impossibility of measuring them were calculated using mathematical equations, due to their importance when constructing projects based on solar energy.

2. Data and Methodology

Weather stations provide long-term data for several years to calculate instantaneous solar radiation values, using different mathematical models. The metrological data for Baghdad was obtained from the IMA, extending from 1961 to 2016 (Moafaq, 2017b). Since 1980, another way to measure solar radiation appeared in the form of satellites deployed in space. The meteorological satellites provide coverage for most of the continents on the surface of the earth. Sophisticated devices are installed on satellites that measure both the solar radiation emitted by the sun, and the solar radiation reflected from the earth's surface towards outer space. The difference between these two values is the solar radiation that the earth receives. These radiation values can be obtained through the NASA site, adopted in this study (NASA site). NASA assumes that the measurement of solar radiation divides each latitude on the earth's surface into several points, at which the measurement is made. Each point is 111 km away from the other, and evaluates the solar radiation as the monthly mean, which is an average of three hours per month (NASA site). The radiation values for each month, for over 22 years, from 1983 to 2005, have been included (NASA site). NASA states that the monthly values it provides are not 100% accurate, as it estimates that the RMSE of these values is 13 to 16%, and that the absolute mean bias error (AMBE) is around -2 ± 0.7% (NASA site).

The global solar irradiance is calculated using the linear regression model, proposed by Angstrom (Angstrom,1924), later modified by Page (Page, 1961). This model includes a method that involves using some of the measured parameters in a linear equation to calculate the total solar irradiance. This model used the clarity index with the sunshine hours (Ali *et al.*, 2020).

To construct the mathematical model, we must first calculate the sun's zenith angle, which can be defined as the angle that describes the path of the sun. In this way, the solar radiation falling on the horizontal surface, at any moment is inclined with the vertical line on the earth's surface, from sunrise to sunset. The cosine of the sun's zenith angle (z) can be calculated using Eq. 1. (Moafaq, 2017b):

$$\cos z = \cos\delta \cos L \cos h + \sin\delta \sin L \quad (1)$$

Where, δ is the solar declination angle, L is latitude, and h the solar hour angle.

The proposed mathematical model to be used can be developed by finding the monthly rate of daily radiation on a horizontal surface outside the planet (extraterrestrial radiation), as shown using Eq. 2. (Moafaq, 2017b; Garba *et al.,* 2020):

$$H_o = (I)(\cos z) \quad (2)$$

For any day of the year, and at a point on the surface of the earth, where the distance from the sun is equal to the average distance between the sun and the earth, the radiation outside the earth is measured on the horizontal surface, which can be defined as described by (I), as shown using Eq. 3. (Moafaq, 2017b; Garba *et al.,* 2020):

$$I = (I_o)(E_o) \quad (3)$$

The equation (2) becomes:

$$H_o = (I_o)(E_o)(\cos z) \quad (4)$$

Io is known as the solar constant, which was first introduced by the French physicist Claude Pouillet in 1837 (Iqbal, 1983), and NASA determined its value as 1353 W/m^2, but later, studies found that its value was approximately 1366.1 W/m^2. The correction factor (Eo) of aberration for the Earth's orbit at any day (n) of the year is calculated using Eq. 5 (Duffie and Beckman, 1991; Bakirci, 2009):

$$E_o = 1 + 0.033 \cos\left(\frac{360\times n}{365}\right) \quad (5)$$

Substituting equation (1) and equation (3) into equation (2) yields:

$$H_o = (I_o)(E_o)(\sin\delta \sin L + \cos\delta \cos L \cos h) \quad (6)$$

The daily solar radiation outside the earth (extraterrestrial) for horizontal surfaces can be obtained as shown by Eq. 7:

$$\overline{H}_o = \frac{24}{\pi}(I_o)(E_o)(\frac{\pi h_s}{180} \sin\delta \sin L + \cos\delta \cos L \sin h_s) \quad (7)$$

The last equation was the result of performing the mathematical integration of Eq. 6, and for the period from sunrise to sunset, which represents the limits of the integration.

Since δ is the solar declination angle, the average sunrise angle h_s can be calculated as shown by Eq. 8. and Eq. 9:

$$\delta = 23{,}45^o \sin\left(\frac{360(284+n)}{365}\right) \quad (8)$$

$$h_s = \cos^{-1}(-\tan L \tan\delta) \quad (9)$$

In the typical case of the incoming solar radiation, the zenith angle z can be considered equal to 90^o, and the angle of sunset can be considered equal to the hour angle; thus, the length of the day can be calculated as shown by Eq. 10:

$$S_{max} = \frac{2}{15} h_s = \frac{2}{15} \cos^{-1}(-\tan L \tan\delta) \quad (10)$$

It is possible to obtain a coefficient called the clarity coefficient (K_t), by calculating the ratio between the monthly average of daily global solar radiation ($\overline{H}$) and the monthly average of daily extraterrestrial radiation ($\overline{H}_o$). The relative duration of sunlight represents the ratio between the monthly average of the daily sunshine duration (S) and the monthly average of daily length (S_{max}).

The linear regression model proposed by Angstrom is most widely employed for calculating the ratio between the global solar radiation monthly rate on a horizontal surface, and the monthly average of daily calculated horizontal extraterrestrial radiation. Since $\overline{H}_o$ and S_{max} can be calculated, therefore calculating $\overline{H}$ can be obtained in this study, which will be compared with the values obtained from the IMA and NASA. A relationship between the clarity coefficient and the relative duration of sunlight will be extrapolated by modeling these equations in FORTRAN language. The programme inputs were feed from calculated and measured factors, which will be used to find the global horizontal solar radiation (GHSR) values.

The monthly average of daily calculated GHSR can be obtained by the linear regression equation (model No.1 and model No.2, depending on the data source), as shown by Eq. 11 (Angstrom,1924; Page, 1961):

$$\bar{H} = \bar{H}_o \left[\mathrm{a} + b \left(\frac{S}{S_{max}} \right) \right] \quad (11)$$

More than one value is obtained for each of the equation constants a and b, by entering different data from both IMA and NASA, thus obtaining two different empirical models (models No.1 and No.2).

The total solar radiation can be calculated using other models by establishing a relationship between the duration of the sun's brightness and the maximum temperature with relative humidity. The equation was formulated in this mathematical model, based on the model proposed by Gopinathan (Gopinathan, 1988). The two other empirical models which have been used in this study can be described as shown by Eq. 12 (model No.3 and model No.4, depending on the data source) (Gopinathan, 1988; Kacem and Yahia, 2013):

$$\frac{\bar{H}}{\bar{H}_o} = a + b \left(\frac{S}{S_{max}} \right) + c(T_{max}) + d(RH) \quad (12)$$

Where, T_{max} is the maximum air temperature, RH relative humidity, and a, b, c and d are the constant coefficients of the empirical models.

Regression coefficients are constants that can be found for the city of Baghdad, using a FORTRAN language programme that was employed to solve the equation. Using two different data sources, IMA and NASA, two values for each of the equation constants were obtained, and thus two different models were formulated.

In the current study, four regression models were formulated to calculate the monthly average of daily GHSR for the city of Baghdad, which is summarised as follows:

1- The first model (model No. 1) is a single linear regression model, derived from the Angstrom equation (Eq. 11), and its variables were obtained using the IMA data as input to the computer programme.

2- The second model (model No. 2) is a single linear regression model that was formulated using the Angstrom equation (Eq. 11), and its variables were obtained using NASA data as input to the computer software.

3- The third model (model No. 3) is a polynomial linear regression model, derived from the Gopinathan equation (Eq. 12), and its variables were obtained using IMA data as input to the computer programme.

4- The fourth model (model No. 4) is a polynomial linear regression model, derived from the Gopinathan equation (Eq. 12), and its variables were obtained using NASA data as input to the computer software.

Models No. 1 and No. 2 depend on sunshine duration, while models No. 3 and No. 4 depend on sunshine duration, maximum temperature and relative humidity.

Two other essential types of solar radiation are diffused and beam radiation; however, as these types of radiation cannot be measured directly, mathematical methods are used to calculate them. Meteorological stations usually lack data on these types of solar radiation because it is impossible to measure. In this study, the diffuse solar radiation was calculated, and by using the value of the diffuse and global solar radiation, it was possible to calculate the beam solar radiation. Liu formulated an empirical equation for calculating monthly-average daily diffused solar radiation. This equation was used in the current study as shown by Eq. 13. (Liu, 1996):

$$\frac{\bar{D}}{\bar{H}} = 1.39 - 4.02K_t + 5.53K_t^2 - 3.108K_t^3 \quad (13)$$

The extraterrestrial solar radiation enters the atmosphere and reaches the surface of the earth; it is exposed to reflection as a result of water vapor and dust particles suspended in the air, and under the influence of other factors. This part of solar radiation is called diffuse radiation. The radiation reaching the earth's surface is called beam radiation, where beam radiation represents the difference between global and diffuse radiation, and it can be calculated as shown by Eq. 14:

$$\bar{B} = \bar{H} - \bar{D} \quad (14)$$

3. Statistical Test and Models Performance

To verify the reliability of any arithmetic model, the calculated values must be compared with measured values, and in this study, four values of solar radiation were calculated, one value by each model, depending on the inputs representing some weather parameters, which were obtained by the two methods: NASA and the IMA. The calculated values adopted in its inputs from the NASA data were compared with the measured values of GHSR provided by NASA, and the same was done with the calculated values that used inputs from the IMA. It is common to use statistical tests to show the reliability of computational models; some well-known and widely used indicators are MBE, RMSE as well as MPE.

MBE and MPE are given information about the performance of the mathematical model for a long-term period, and their values should be low, which indicates a convergence between the calculated values and the measured values of GHSR. Positive values of MBE indicators mean that there is an exaggeration in the calculated values compared to the measured values, and vice versa. As for the indicator MPE, it behaves opposite to the indicator MBE – if its value is positive, this indicates that the calculated values have a lower value compared to the measured values, and if its value is negative, then there is an overstatement in the calculation relative to the measured values. RMSE is in the ideal case equal to zero, which means that the calculated and measured values are equal, and its value is always positive. Whenever there is a small value, it means that the values of the proposed arithmetic model is closer to the measured values; this indicator reflects the information about the short-term correlation performance. The three indicators can be formulated as shown by Eq. 15, Eq. 16 and Eq. 17.

$$\boldsymbol{MBE} = \left(\frac{1}{n}\right)\sum_{i=1}^{n}\left(\overline{\boldsymbol{H}}_{i,c} - \overline{\boldsymbol{H}}_{i,m}\right) \quad (15)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\overline{H}_{i,c} - \overline{H}_{i,m}\right)^2} \quad (16)$$

$$MPE = \left[\sum_{i=1}^{n}\left(\frac{\overline{H}_{i,m}-\overline{H}_{i,c}}{\overline{H}_{i,m}} \times 100\right)\right]/n \quad (17)$$

Where, $\overline{H}_{i,c}$ is the ith calculated and $\overline{H}_{i,m}$ is the ith measured monthly average daily *GHSR*, and n is the number of observations. The negative values give an indication that the average measured values are greater than the mean values calculated mathematically.

4. Results and Discussions

As has been shown earlier, four models were formulated to calculate daily GHSR for the city of Baghdad, where two models were obtained from each equation, depending on the programme's inputs. These inputs represent some of the weather parameters. In the first and second models, the sunlight duration (S) was used, while more parameters were added in the third and fourth models. The maximum temperature (T_{max}), relative humidity (RH%) in addition to the sunlight duration (S) were used as the programme's inputs. These parameters were obtained from two different sources. The first source of the programme's inputs was the IMA, which provided meteorological data for the period 1961 to 2016. The second source was NASA data for the period of 1984 to 2004. In the current study and for the city of Baghdad, the location of the study, the solar radiation was calculated using single and polynomial linear regression models (Eq 18-21).

For single linear regression model employing IMA inputs, Eq. 11 can be formulated as shown by Eq. 18 (model No. 1).

$$\frac{\overline{H}}{\overline{H}_o} = 0.390 + 0.354\left(\frac{S}{S_{max}}\right) \quad (18)$$

When NASA inputs are used, Eq. 11 can be formulated as shown by Eq. 19 (model No. 2).

$$\frac{\overline{H}}{\overline{H}_o} = 0.331 + 0.349\left(\frac{S}{S_{max}}\right) \quad (19)$$

For the polynomial linear regression models, Eq. 12 can be formulated as shown by Eq. 20 and Eq. 21.

- By using IMA inputs (model No. 3):

$$\frac{\overline{H}}{\overline{H}_o} = 0.603 + 0.018\left(\frac{S}{S_{max}}\right) + 0.0027(T_{max}) - 0.0008(RH) \quad (20)$$

- By using NASA inputs (model No. 4).

$$\frac{\overline{H}}{\overline{H}_o} = 0.527 + 0.016\left(\frac{S}{S_{max}}\right) + 0.0025(T_{max}) - 0.0006(RH) \quad (21)$$

The coefficients a and b for the first and second models, as well as the coefficients a, b, c, and d for the third and fourth models, are usually constants. These constants depend on the periodic climate changes, so they have variable values, according to the change of the source of the climate input. The values of the constants (a, b) for Baghdad city for the first and second models were 0.390, 0.354 and 0.331, 0.349, respectively. The values of (a, b, c, d) for the third and fourth models were 0.603, 0.018, 0.0027, -0.0008 and 0.527, 0.016, 0.0025, -0.0006, respectively. These resulting values are summarized in Table 1. The constants in all models have greater values in the case of relying on the IMA than NASA data. In general, regardless of the atmospheric parameters, the measured radiation values provided by NASA are lower than the values provided by the IMA. This is due to the change in constants for all models. Moreover, the period used by the IMA to collect data extended to more than 55 years, while NASA collected data for 20 years only. The period used data was reflected in all measured and recorded values, in both sources, leading to the emergence of a difference in the constants of values of the four models.

To evaluate the proposed models, RMSE, MBE, and MPE% indicators were used. It can be noted from Table 1 that the lowest value of the RMSE, MBE, and MPE% by using a single linear regression model were related to the values provided by the IMA (model No. 1). The RMSE, MBE and MPE% values were 0.4769, 0.0164 and 0.2207, respectively. Prior studies recommend that whenever these values are small, the calculated GHSR values are close to the measured values (Mohamed *et al.*, 2014). Also, the value of MPE% should be close to zero.

Table 1
Regression coefficients, root mean square error (RMSE), mean bias error (MBE) and mean percentage error (MPE) for the two models based on data values from IMA and NASA.

Model number	a	b	c	d	RMSE MJ/m²/day	MBE MJ/m²/day	MPE %
Model No.1	0.390	0.354	-	-	0.476	0.016	0.220
Model No.2	0.331	0.349	-	-	0.864	0.177	-0.968
Model No.3	0.603	0.018	0.0027	-0.0008	0.642	0.399	-1.148
Model No.4	0.527	0.016	0.0025	-0.0006	0.960	0.218	-1.022

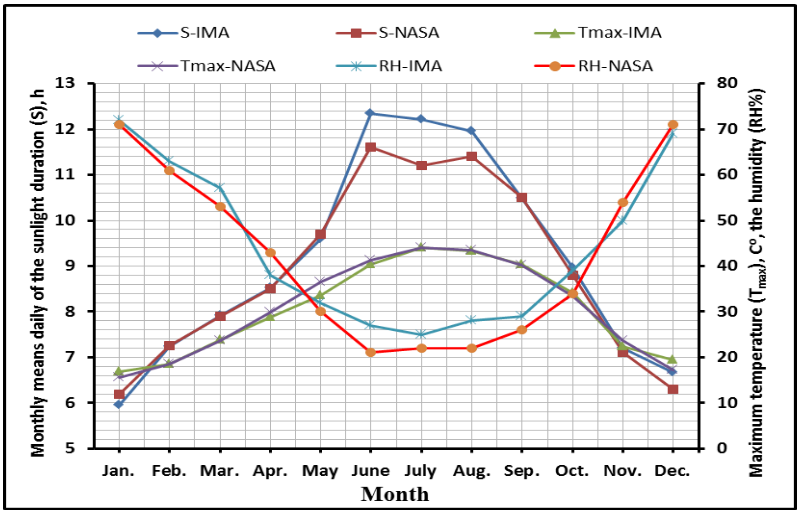

Fig. 1 Variation of monthly means daily maximum temperature, relative humidity and sunshine duration by using data of IMA for a period (1961-2016) and NASA for the period (1984-2004) for Baghdad city..

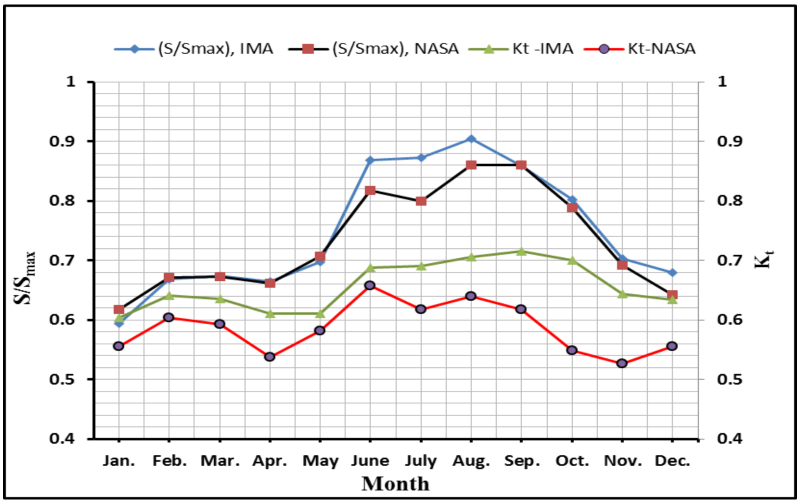

Fig. 2 Variation of clearness index and S/S_{max} for Baghdad city by using IMA and NASA data.

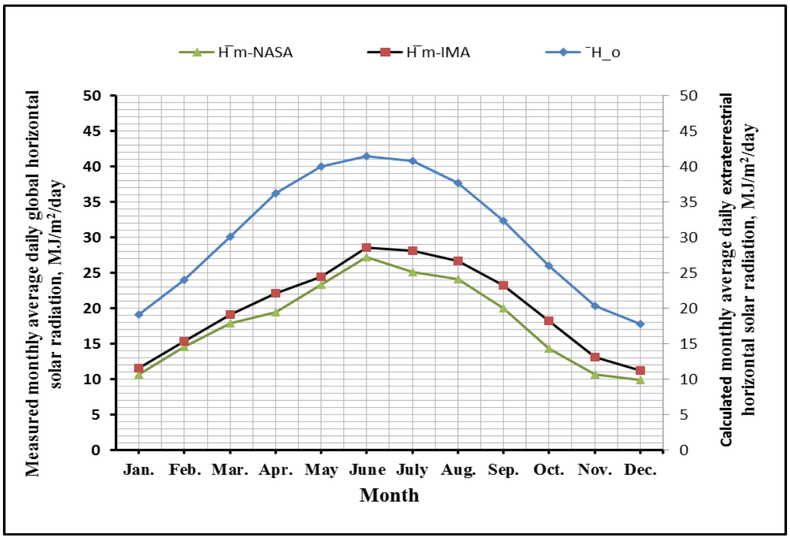

Fig. 3 Monthly avearge variation of calculated daily extraterrestrial radiation, measured daily global solar radiation by IMA and NASA.

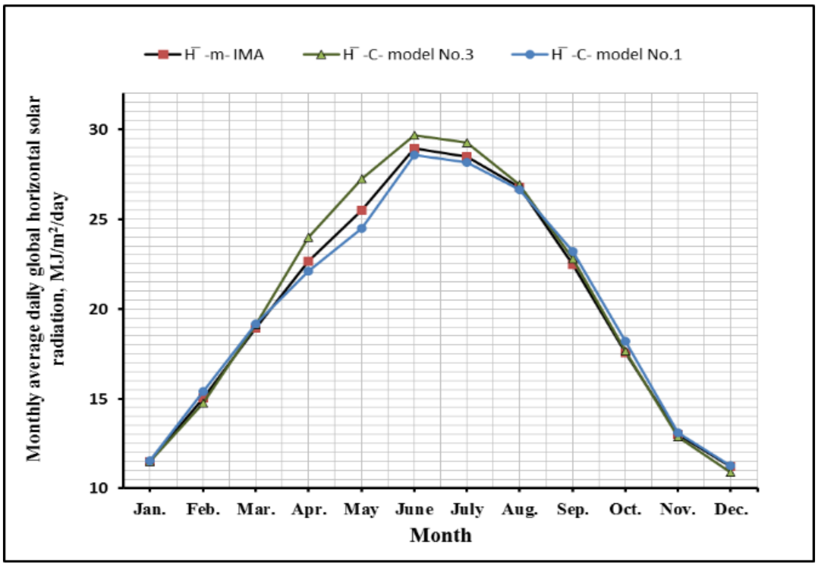

Fig. 4 Monthly avearge variation daily global solar radiation; measured by IMA and calculated by model No. 1 and model No. 3.

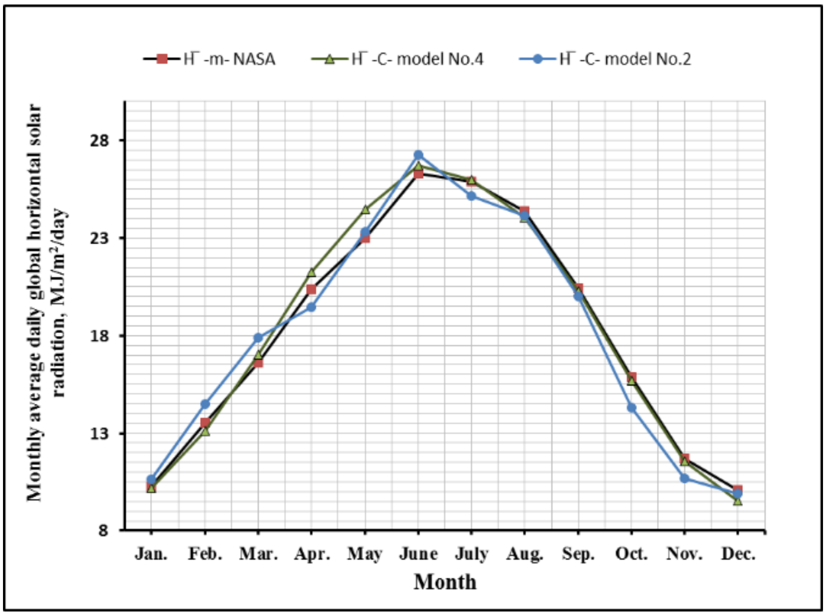

Fig. 5 Monthly avearge variation daily global solar radiation; measured by NASA and calculated by model No. 2 and model No. 4.

The values calculated by the four regression models were compared with the measured ones by IMA and NASA, through a statistical test by calculating the RMSE, MBE and MPE% indicators – Eqs. 15 to 17, respectively, as shown in Table 1. Each of these equations include the product of subtracting the measured values from the calculated ones, or vice versa. All the values for these three indicators obtained were very small, which means that the calculated values using the four models are very close, compared to the values measured by the IMA and NASA.

Table 1 demonstrates that the value of the MPE% by using a single linear regression model with IMA input values (model No. 1) is positive, with a value of 0.2207. This means that the calculated GHSR values are less than the measured GHSR values in general. As shown by Eq. 17, and in the other three cases, the values of the indicators of MPE% are negative at -0.9680, -1.1487, and -1.0225 for models No. 2, No. 3, and No. 4, respectively. This indicates that the calculated values are greater than the measured values of GHSR. The values of the three statistical indicators RMSE, MBE and MPE% were 0.4769, 0.0164, 0.2207; 0.8641, 0.1773, -0.9680; 0.6420, 0.3996, -1.1487; 0.9604, 0.218, -1.0225 for models No. 1, No. 2, No. 3 and No. 4, respectively. These results mean that the single linear regression model that is related to the values provided by the IMA (model No. 1) is the best for the city of Baghdad, when calculating the *GHSR*, and it is the closest to reality.

Figure 1 includes the monthly average values of daily sunlight duration, maximum temperature, and relative humidity for each month of the year, provided by the IMA and NASA. As indicated in Figure 1, the highest values of sunlight duration are in June for both sources. The value recorded by IMA was higher than NASA's, recorded at 12.34 and 11.60, respectively. The lowest value of sunlight duration was recorded in January, and the value of NASA was higher than that of IMA at 6.20 and 5.96, respectively. The highest value of the maximum temperature was in July, reaching 44°C for both NASA and IMA, while the lowest value was in January. The value maximum temperature of IMA was greater than NASA's, recorded at 16.80 and 15.50, respectively. The highest value of relative humidity recorded by the IMA in January was 72%, while its highest value was recorded by NASA at 71%, for both December and January. The lowest value was 25% in July by the IMA, while June recorded 21% as the lowest value by NASA. The discrepancy in these values is due to the difference in the number of measurements of the year and methods used in measuring, as well as the accuracy of measurement devices for both the IMA and NASA. Similar behavior to those parameters was recorded by both sources, and was observed with little difference in their values.

In Figure 2, a comparison of clarity factor (K_t) and S/S_{max} is shown using IMA and NASA data. In most months of the year, the IMA-based values are found to be slightly larger than the NASA-based values. The highest value of S/S_{max} was 0.904 in the month of August, based on IMA, while it was 0.861 in August and September, according to NASA. The lowest value was in January for both IMA and NASA at 0.594 and 0.618, respectively. The maximum value of K_t was 0.715, according to the IMA, in September, and was 0.657 by NASA in June. The minimum value of K_t was 0.603 by the IMA in January, and was 0.526 in November, according to NASA. This means that the daily mean of global solar irradiance calculated for the horizontal surface, the measured irradiance value, and the duration of the daily sunshine provided by the IMA are greater than the values provided by NASA.

Figure 3 shows that the calculated extraterrestrial radiation ($\overline{\mathbf{H}}_{\mathbf{o}}$) is higher than the global radiation values measured ($\overline{\mathbf{H}}$) by both NASA and IMA. This is because when the radiation enters the atmosphere, a proportion of it is reflected back to the atmosphere due to dust, water vapor particles, and other factors. Maximum values of $\overline{\mathbf{H}}_{\mathbf{o}}$ and $\overline{\mathbf{H}}$**,** according to NASA and the IMA, were in June with values 41.489, 28.555, and 27.28, respectively, while minimum values were 17.762, 11.268, and 9.89, respectively, in December.

Figures 4 and 5 provide a comparison between the measured values of global radiation by the IMA and NASA, with the calculated values using Eqs. 18 to 21 (the four models). In Figure 4, it was observed that the difference between the measured and calculated values by model No. 1 and model No. 3, from April to July, was greater than the other months, where the values were 22.105, 22.641, 20.381; 24.482, 25.497, 22.992; 28.555, 28.947, 26.315; 28.182, 28.493, 25.907, respectively; the largest difference was recorded during May. As for the values of the other months, they appear almost identical. It can be said that models No. 1 and No. 3 can be relied upon to calculate the yearly daily GHSR. with great accuracy, but they are less accurate for the months from April to July. Despite that, in general, the values calculated by each of the two models correspond well with the measured values by the IMA.

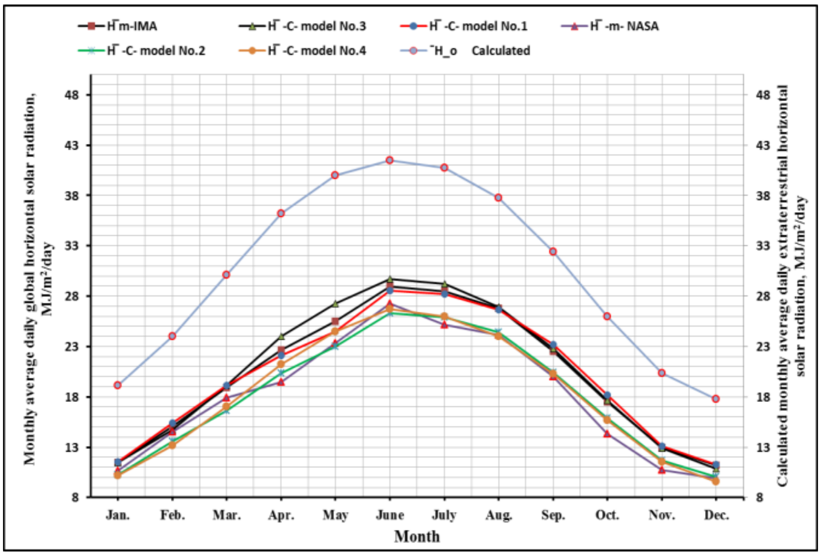

Fig. 6 Monthly average variation of calculated daily extraterrestrial radiation, measured daily global solar radiation by the IMA and NASA, calculated daily global solar radiation by models No. 1–No. 4.

Table 2
Comparison between the calculated daily diffuse solar radiation (MJ/m²/day) by the IMA-based, NASA-based, model No. 1, model No. 2, model No. 3 and model No. 4, for Baghdad city.

Month	Ima- diffuse radiation	NASA- diffuse radiation	Model No.1 - diffuse radiation	Model No.2 - diffuse radiation	Model No.3 - diffuse radiation	Model No.4 - diffuse radiation
January	3.403	3.141	3.387	3.394	3.037	3.005
February	4.106	3.871	4.015	3.936	3.615	3.505
March	5.183	4.837	5.123	5.167	4.495	4.604
April	6.411	5.644	6.567	6.957	5.911	6.161
May	7.082	6.746	7.376	7.880	6.651	7.078
June	6.558	6.265	6.648	6.821	6.044	6.136
July	6.402	5.716	6.473	6.643	5.886	5.905
August	5.717	5.181	5.747	5.770	5.232	5.156
September	4.793	4.132	4.648	4.715	4.224	4.189
October	4.000	3.141	3.852	3.876	3.491	3.444
November	3.464	2.830	3.434	3.403	3.098	3.057
December	3.065	2.691	3.046	2.960	2.744	2.601

Table 3
Comparison between the calculated daily beam solar radiation (MJ/m²/day) by the IMA-based, NASA-based, model No. 1, model No. 2, model No. 3 and model No. 4 for Baghdad city.

Month	Ima- beam radiation	NASA- beam radiation	Model No.1 - beam radiation	Model No.2 - beam radiation	Model No.3 - beam radiation	Model No.4 - beam radiation
January	8.125	7.498	8.087	8.101	7.251	7.175
February	11.287	10.638	11.034	10.817	9.935	9.634
March	13.974	13.042	13.812	13.930	12.117	12.412
April	15.693	13.815	16.073	17.029	14.469	15.082
May	17.399	16.573	18.120	19.356	16.340	17.386
June	21.996	21.014	22.298	22.877	20.271	20.579
July	21.779	19.443	22.019	22.598	20.020	20.086
August	20.930	18.968	21.043	21.128	19.158	18.876
September	18.396	15.857	17.839	18.093	16.210	16.078
October	14.211	11.158	13.684	13.770	12.404	12.235
November	9.633	7.869	9.550	9.463	8.614	8.502
December	8.202	7.198	8.151	7.921	7.343	6.958

Figure 5 shows the semi-match between the measured daily GHSR values by NASA and the values calculated by model No. 4, throughout the year, except for April, May, and June, where the values were 19.460, 20.381, 21.244; 23.320, 22.992, 24.46; and 27.280, 26.315, 26.716, respectively. As for model No. 2, it shows a larger deviation from model No. 4 for all months of the year, compared to the daily GHSR values measured by NASA. Despite that, the difference is very small between daily GHSR calculated values for both models (models No. 2 and No. 4) compared to the values measured by NASA.

All the values of calculated daily GHSR using the four mathematical models (equations 18 to 21), and the values of measured daily GHSR by NASA and the IMA, as well as the values of the calculated extraterrestrial radiation throughout the year, are included in Figure 6. It can be noted that daily GHSR calculated by models No. 1 and No. 3, which are based on some of the IMA's inputs, are close to the daily GHSR measured by the IMA. Likewise, for models No. 2 and No. 4, their calculated daily GHSR are very close to the measured NASA values. The difference between extraterrestrial radiation on the one hand and global radiation on the other – calculated by all four models or measured the by IMA and NASA – occurred as a result of absorption and reflection of part of the radiation by suspended particles in the air, water vapor particles or other factors. It can be noted that all the models behave similarly and have convergent values; Figure 6 was done for the sake of comparison and to ensure the reliability of the models proposed in this study.

Table 2 reviews the values of daily diffuse solar radiation, calculated depending on the IMA, NASA, and the values calculated by each of the four models. The lowest values of diffused radiation were in December, when their values were 3.065, 2.691, 3.046, 2.960, 2.74, and 2.601, respectively, while the maximum values in May were 7.082, 6.746, 7.376, 7.880, 6.651, and 7.078, respectively.

Table 3 presents the values of daily beam (direct) radiation that reaches the earth's surface after a reflection of part of the global radiation away from the earth. Maximum daily beam solar radiation values were 21.996, 21.014, 22.298, 22.877, 20.271, and 20.5799 in June, for values calculated depending on the IMA, NASA, and the values calculated by four models, respectively. While minimum values were 8.125, 8.087, and 7.251 in January, calculated depending on the IMA and values calculated by models No. 1 and No. 3, respectively. The minimum values for the beam, calculated depending on NASA and the ones calculated by models No. 2 and No. 4, were 7.198, 7.921, and 6.958785, respectively, in December.

5. Conclusion

In this study, the possibility of using linear regression models to calculate solar radiation was demonstrated. Each linear regression equation resulted in two models, one of which depended on the use of some parameters of NASA, and the other was based on the IMA's parameters. As a result, four models were formulated. These models were produced as a result of calculating the constants of equations, which changed as a result of using the two input sources for the computer programme, which was formulated in FORTRAN. The obtained results were analysed to calculate the GHSR using the four models, two models were single and two polynomials. It was noted that the performance of all the models was good and acceptable, compared to the values measured by NASA and the IMA. Through statistical tests for all models, statistical indicators values (RMSE; MJ/m2/day, MBE; MJ/m2/day and MPE%) were obtained for model No. 1 (0.4769, 0.0164, 0.2207), model No. 2 (0.8641, 0.1773, -0.9680), model No. 3 (0.6420, 0.3996, -1.1487), and model No. 4 (0.9604, 0.218, -1.0225). It can be concluded that the single linear regression model that depends on IMA's data (model No. 1) is closer to reality, as compared to the measured values, followed by the single linear regression model based on NASA's data (model No. 2), the polynomial model based on NASA data (model No. 4), and finally the polynomial model based on the IMA's data (model No. 3). The four models can be used to calculate the GHSR for the city of Baghdad. For more accurate values, it is recommended to use a single linear regression model based on the IMA's data (model No. 1).

References

Abdelhafidi, N., Bachari, N.E.I. & Abdelhafidi, Z. (2021). Estimation of solar radiation using stepwise multiple linear regression with principal component analysis in Algeria. *Meteorol Atmos Phys*, 133, 205–216; doi.org/10.1007/s00703-020-00739-0

Ali Etem Gürel, Ümit Agbulut & Yunus Biçen (2020). Assessment of machine learning, time series, response surface methodology and empirical models in prediction of global solar radiation. *Journal of Cleaner Production*, 277, 122353; doi.org/10.1016/j.jclepro.2020.122353

Angstrom, A. (1924). Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation. ***Q. J. Roy.*** *Meteorol. Soc.* 50 (210),121–126; doi.org/10.1002/qj.49705021008

Bailek, N., Bouchouicha, K., Abdel-Hadi, Y.A., El-Shimy, M., Slimani, A., Jamil, B. & Djaafari, A. (2020). Developing a new model for predicting global solar radiation on a horizontal surface located in Southwest Region of Algeria. *NRIAG Journal of Astronomy and Geophysics*, 9(1), 341-349; doi.org/10.1080/20909977.2020.1746892

Bakirci, K. (2009). Correlations for estimation of daily global solar radiation with hours of bright sunshine in Turkey. *Energy*, 34(4), 485–501; doi.org/10.1016/j.energy.2009.02.005

Bamehr, S. & Sabetghadam, S. (2021). Estimation of global solar radiation data based on satellite-derived atmospheric parameters over the urban area of Mashhad, Iran. *Environmental Science and Pollution Research*, 28(6), 7167-7179; doi.org/10.1007/s11356-020-11003-8

Benghanem, M. & Joraid, A. (2007). A multiple correlation between different solar parameters in Medina, Saudi Arabia. *Renew Energy,* 32(14), 2424–2435; doi.org/10.1016/j.renene.2006.12.017

Besharat F, Dehghan AA, Faghih AR. (2013). Empirical models for estimating global solar radiation: a review and case study. *Renew Sustainable Energy Rev*, 21,798–821; doi.org/10.1016/j.rser.2012.12.043

Blal, M., Khelifi, S., Dabou, R., Sahouane, N., Slimani, A., Rouabhia, A., Ziane, A., Necaibia, A., Bouraiou, A. and Tidjar, B. (2020). A prediction models for estimating global solar radiation and evaluation meteorological effect on solar radiation potential under several weather conditions at the surface of Adrar environment. *Measurement*, 152, 107348; doi.org/10.1016/j.measurement.2019.107348

Chaichan, M.T., Kazem, H.A. and Abed, T.A. (2018). Traffic and outdoor air pollution levels near highways in Baghdad, Iraq. *Environment Development and Sustainability*, 20(2), 589-603; doi.org/10.1007/s10668-016-9900-x

Chaichan, M.T., Kazem, H. A. (2018) *Generating Electricity Using Photovoltaic Solar Plants in Iraq*, Springer, ISBN: 978-3-319-75030-9; doi.org/10.1007/978-3-319-75031-6

Duffie, J.A., Beckman, W.A. (1991) *Solar engineering of thermal process*. 2nd edn. John Wiley & Sons, New York, NY, USA.

El-Sebaii, A. A. and Trabea, A. A. (2005). Estimation of Global Solar Radiation on Horizontal Surfaces Over Egypt*, Egypt. J. Solids*, 28(1), 163-175.

Ertekin, C., Yaldiz, O. (2000). Comparison of some existing models for estimating global solar radiation for Antalya (Turkey). *Energy Conversion and Management*, 41(4), 311–330; doi.org/10.1016/S0196-8904(99)00127-2

Garba Musa Argungu, Yusuf Abubakar Sanusi, Kabir Ahmed Dabai, Muazu Abubakar, Abubakar Roko.(2020). Comparative Analysis of the Performance of Six Models for the Estimation of Global Solar Radiation for Katsina, Nigeria. *American Journal of Energy Engineering*. 8(2), 18-25; doi.org/10.11648/j.ajee.20200802.11

Gopinathan, K.K.(1988). A new model for estimating total solar radiation in Doha. *Energy Conversion and Managemen*, 28, 63–72.

Iqbal, M. (1983) *An Introduction to Solar Radiation*, Academic Press, Toronto, Canada.

Kacem Gairaa , Yahia Bakelli. (2013). A Comparative Study of Some Regression Models to Estimate the Global Solar Radiation on a Horizontal Surface from Sunshine Duration and Meteorological Parameters for Gharda / a Site, Algeria. *Hindawi Publishing Corporation ISRN Renewable Energy*; doi.org/10.1155/2013/754956

Kazem, H.A., Yousif, J.H., Chaichan, M.T.(2016). Modeling of Daily Solar Energy System Prediction using Support Vector Machine for Oman. *International Journal of Applied Engineering Research*, 11(20), 10166-10172.

Koussa, M., Malek, A., Haddadi, M. (2009). Statistical comparison of monthly mean hourly and daily diffuse and global solar irradiation models and a Simulink program development for various Algerian climates. Energy Conversion and Management,50(5),1227–1235; doi.org/10.1016/j.enconman.2009.01.035

Liu, D.L. (1996). Incorporating Diurnal Light Variation and Canopy Light Attenuation into Analytical Equations for Calculating Daily Gross Photosynthesis. Ecological Modeling, 93(1-3): 175–189; doi.org/10.1016/0304-3800(95)00223-5

Liu, P., Tong, X., Zhang, J., Meng, P., Li, J. and Zhang, J. (2020). Estimation of half-hourly diffuse solar radiation over a mixed plantation in north China. Renewable Energy, 149(2), 1360-1369; doi.org/10.1016/j.renene.2019.10.136

Moafaq, K.S.Al-Ghezi. (2019). Study the Maximum Solar Radiation by Determining the Best Direction of the Solar Collectors. *International Research Journal of Advanced Engineering and Science*, 4(3), 42-44.

Moafaq, K.S.Al-Ghezi. (2017). Calculate the Reflected Hourly Solar Radiation by Mirror Surfaces of Solar Concentrators Parabolic Trough. *International Journal of Computation and Applied Sciences IJOCAAS*, 3(3), 256-260.

Moafaq, K.S.Al-Ghezi. (2017). The Global and Scattered Radiation Evaluation for a Horizontal Surface in Baghdad City. *International Journal Of Computation and Applied Sciences IJOCAAS*, 3(1), 153-158.

Moafaq, K.S.Al-Ghezi., Khaleel, I.Abass., Ahmed, Q. Salam.,Raid, S. Jawad & Hussein, A. Kazem. (2021). The possibilities of using nano-CuO as coolants for PVT system: An experimental study. *Journal of Physics: Conference Series*, 1973(1), 012123; doi.org/10.1088/1742-6596/1973/1/012123

Mohamed Salah Mecibah, Taqiy Eddine Boukelia, Reda Tahtah, Kacem Gairaa. 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*, 36, 194–202; doi.org/10.1016/j.rser.2014.04.054

Mujabar, S. & Venkateswara, R.C. (2021). Empirical models for estimating the global solar radiation of Jubail Industrial City, the Kingdom of Saudi Arabia. *SN Applied Sciences*, 3(1), 1-11; doi.org/10.1007/s42452-020-04043-9

NASA. https://eosweb.larc.nasa.gov/cgi-bin/sse/daily.cgi

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 latitudes 40ºN-40ºS. *UN Conference on New Sources of Energy*, 4, 378-390.

Rabaia, M.K.H., Abdelkareem, M.A., Sayed, E.T., Elsaid, K., Chae, K.J., Wilberforce, T. & Olabi, A.G. (2021). Environmental impacts of solar energy systems: A review. *Science of The Total Environment*, 754, p.141989; doi.org/10.1016/j.scitotenv.2020.141989

Sabbagh, J., Sayigh, A.A.M, El-Salam, E.M.A. (1977). Estimation of the total solar radiation from meteorological data. *Sol Energy*, 19(3), 307–311; doi.org/10.1016/0038-092X(77)90075-5

Sen, Z. (2007). Simple nonlinear solar irradiation estimation model. *Renew Energy*, 32(2), 342–350; doi.org/10.1016/j.renene.2006.02.015

Sridharan, M. (2021). Generalized Regression Neural Network Model Based Estimation of Global Solar Energy Using Meteorological Parameters. *Annals of Data Science*, 1-19; doi.org/10.1007/s40745-020-00319-4

Trenberth, K.E., Fasullo, J.T. & Kiehl, J. (2009). Earth's global energy budget. *Bulletin of the American Meteorological Society*, 90(3), 311–324; doi.org/10.1175/2008BAMS2634.1

Yousif, J.H., Al-Balushi, H.A., Kazem, H.A. & Chaichan, M,T. (2019). Analysis and forecasting of weather conditions in Oman for renewable energy applications. *Case Studies in Thermal Engineering*, 13, 100355; doi.org/10.1016/j.csite.2018.11.006

Yu, L., Zhang, M., Wang, L., Lu, Y. & Li, J. (2021). Effects of aerosols and water vapor on spatial-temporal variations of the clear-sky surface solar radiation in China. *Atmospheric Research*, 248, 105162; doi.org/10.1016/j.atmosres.2020.105162