Journal of
Renewable Energy Research and Applications (RERA)

Vol. 5, No. 1, 2023, 21-33 *DOI: 10.22044/rera.2023.12439.1183*

New Empirical Models for Estimation of Diffuse Radiation using Satellite Data

M. Arslan[1] and M. Çunkaş[2*]

1. Konya Technical University, Vocational School of Technical Sciences, Konya, Türkiye.
2. Selçuk University, Faculty of Technology, Department of Electrical & Electronics Engineering, Konya, Türkiye.

Received Date 21 November 2022; Revised Date 25 February 2023; Accepted Date 14 March 2023
**Corresponding author: mcunkas@selcuk.edu.tr(M. Çunkaş)*

Abstract

Diffuse radiation is used in photovoltaic systems and other energy applications. Since global radiation is measured by local meteorological stations, it is possible to reach this radiation data. However, diffuse radiation is not usually measured, so it is not possible to obtain regular data on diffuse radiation. For this reason, efforts are underway to develop various empirical models to estimate diffuse radiation. This work aims to develop new empirical models to estimate the diffuse radiation values for Konya, Türkiye. The empirical models are used to determine the relationship between the diffuse fraction and the clearness index. Data from NASA-surface meteorology and solar energy and the measured global solar is used. The three most suitable developed models are selected, and it is suggested to estimate the diffuse radiation. The developed models consist of 2nd, 3rd, and 4th-order polynomial regression models. The proposed models are tested to evaluate their performances by using eight statistical methods. These are Mean Bias Error (MBE), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Sum Squared Relative Error (SSRE), Relative Standard Error (RSE), Coefficient of determination (R^2), and Correlation Coefficient (r). For the suggested models, the statistical parameter R^2 values are calculated as 0.999705413, 1, and 1, and the RSE values are determined as 0.0084537, 0.000122, and 6.05E-06. The proposed models can contribute to the researchers working on the applications of photovoltaic systems. The approaches could be used to estimate diffuse radiation elsewhere under similar climatic conditions.

Keywords: *Diffuse solar radiation, empirical models, satellite data, clearness index, solar energy*.

1. Introduction

Energy consumption is increasing day by day due to technological developments in the world. There is a growing interest for renewable energy sources such as solar energy to meet the energy demand. Renewable energy sources can reduce environmental pollution, and provide sustainable energy. With the rapid depletion of fossil origin sources, clean and renewable energy sources form the basis of sustainable energy, and find wide application areas [1].

29% of solar radiation coming to Earth is returned in space by clouds, shiny surfaces, and the atmosphere itself, 23% is absorbed by gases, dust, and other particles in the atmosphere, and the remaining 48% is absorbed on earth. This shows that 71% of incoming solar energy is absorbed by the atmosphere and earth. The radiation reaching the earth after scattering from solar radiation due to reflection is called diffuse radiation [2].

Diffuse radiation is used in a variety of applications including solar power systems, climate modelling, and agricultural productivity. There are several methods to estimate diffuse radiation: empirical models, physical models, and statistical models. Empirical models are built using historical data on solar radiation and atmospheric conditions to estimate the amount of diffuse radiation. The isotropic sky model, which assumes that the sky brightness is the same in all directions, is the commonly preferred empirical model for estimating diffuse radiation. Physical models using radiative transfer theory need detailed atmospheric data such as water vapor, the content of aerosols, and atmospheric temperature profiles. These models can be complex and computationally intensive but they often can give more accurate results than empirical models. Some statistical models use artificial intelligence techniques to estimate diffuse radiation by identifying patterns in historical data. These models need to be trained on large datasets to increase their accuracy [3].

The common feature of solar energy systems is to benefit from the sun's radiation at the maximum level. In photovoltaic applications all over the world, total radiation values on the horizontal surface are considered. After determining the diffuse and beam components of the total radiation on the horizontal surface, calculations are made for the titled surfaces, and then the performance of the titled solar PV panels can be evaluated using the calculated values [4]. In photovoltaic projects, it is necessary to know radiation components such as daily direct irradiance and diffuse radiation for solar energy. While total radiation values can be easily obtained for many places in Türkiye and the world, it is very difficult to reach diffuse radiation values. The cost, which arises because of the maintenance and calibration of the diffuse radiation devices at certain periods, does not allow for obtaining measurement data in many places [5]. Empirical models were used to obtain this data, especially since the diffuse radiation value could not be measured. Some models were proposed for diffuse radiation by using global radiation measurements, meteorological data or satellite data [6]. In these studies, basic parameters such as extra-terrestrial radiation, sunshine hours, and average temperature were used [7]. The most used parameters are the clearness index and relative sunshine duration. The diffuse radiation can be calculated theoretically from clear sky conditions.

Consideration of more parameters will give more satisfactory results but this will increase the complexity of the radiation model [8]. Many models were derived for estimating solar radiation using solar radiation and different parameters [9]. The parameters such as instantaneous diffuse and beam radiation have been used [10, 11]. The properties between the clearness index and diffuse fraction were examined annually, monthly, daily, and hourly [12]. Another diffuse radiation model has been proposed that takes into account the total cloud cover instead of some meteorological data [13]. Besides, regression models and parametric models based on ASHRAE were used for hourly diffuse radiation estimation [14]. The effects of horizontal visibility, relative humidity, air temperature, and other meteorological or environmental variables on the diffuse radiation fraction were considered in constructing the hourly diffuse radiation model [15]. The use of artificial intelligence techniques in diffuse radiation estimation emerged as an alternative to regression methods. If global solar radiation is known, the machine learning methods can be used to estimate diffuse solar radiation. The estimation methods using machine learning were reported to be competitive in performance compared to other traditional methods [16, 17]. Hybrid models have been developed using the radiative transfer model and machine learning techniques [18]. However, empirical models are very common because of their readily available inputs, relatively simple forms of functions used, and easy handling [19].

Turkey has recently significantly increased its share of renewable energy sources in electricity generation, and continues to increase it. In addition, solar energy investments occupy a significant place among renewable energy sources. Diffuse radiation is considered an important parameter in the installation and economic operation of solar systems. There is little research on diffuse solar radiation estimation and modeling in Türkiye. Tiris, Tiris, and Türe [20] estimated diffuse radiation using the sunshine and cloudiness index for the Gebze location in Türkiye. Aras, Balli, and Hepbasli [21] proposed new hybrid models to estimate solar radiation for the central anatolian region of Türkiye. Ulgen and Hepbasli [22] obtained empirical correlations by using diffuse fraction and clearness index for the Izmir province in Türkiye. They also developed some empirical relationships to estimate the diffuse fraction for the three largest cities [23]. Tarhan and Sarı [24] determined two separate mathematical models to predict the global and diffuse radiations for the central Black Sea region. Arslanoglu [25] evaluated the feasibility of diffuse solar radiation models for Bursa. Tırmıkçı and Yavuz [26] studied the diffuse solar radiation estimation for the Antalya city. They used the clearness index and sunshine fraction. Bakirci and Kirtiloglu [27] derived empirical models to estimate diffuse radiation for Erzincan, Türkiye, using satellite data. They defined diffuse fraction as a function of the clearness index. Rusen and Konuralp [28] estimated global and diffuse radiation using satellite-based forecasting methods (HELIOSAT, Meteonom, and PVGIS) for nine locations in Turkey, and compared the performances of these methods. Bakirci [1] investigated models to estimate diffuse solar radiation in some regions of Türkiye, and determined the three most appropriate models for Erzurum and Gebze. Also Bakırcı [29] obtained new correlation models for diffuse radiation estimation in 13 locations in Eastern Anatolia region.

As mentioned above, some studies have been carried out to estimate diffuse radiation for a few locations in Turkey. However, there is still a need to develop accurate and reliable new forecasting models for different regions of Turkey. This work aims to correlate the diffuse fraction and the clearness index to estimate diffuse radiation. Diffuse radiation data cannot be measured in all regions of Türkiye due to difficulties in measurements. For this reason, diffuse radiation models have been formed for Konya, Türkiye. Two approaches are present for the estimation of diffuse radiation. First new correlations between diffuse fraction and clearness index were proposed using NASA-SSE data. This emphasizes the importance of using NASA-SSE data with empirical approaches in diffuse radiation estimation. Second new diffuse radiation models were developed by using 12 empirical models available in the literature. This approach had a significant contribution to developing new models for the accurate and reliable estimation of diffuse radiation. 14 diffuse radiation models were developed, and the validity of these models was tested using eight statistical methods. The three most suitable diffuse radiation models among developed models were selected, and it was suggested to use this model. The main contributions of this article are as follows:

- Previous studies have addressed the estimation of diffuse radiation based on the parameters such as global radiation, sunshine duration, and calculated extraterrestrial solar radiation. In very few studies, the diffuse radiation model was formed using NASA-SSE data. In this work, new models were proposed using both existing models and NASA-SSE data, and these models were compared with each other.
- Diffuse radiation has been studied for some locations in Türkiye, but no study has been conducted for Konya using both empirical models and NASA-SSE data. Thus there is no comprehensive study evaluating and comparing the different empirical models available.
- 14 new models were developed based on diffuse fraction and clearness index. Eight statistical methods were used to evaluate the performances of the models. Considering these statistical test results, the three best models have been proposed for Konya, Türkiye. The proposed models can also be used to predict the diffuse radiation for any location in the Central Anatolian region of Türkiye.
- The model development approach in this article can contribute to the researchers and scientists working on photovoltaic systems. The new models and approaches could be used to predict diffuse radiation.

2. Materials and methods

Since solar radiation cannot be measured, solar data is estimated by developing empirical models. In the literature, many models were proposed to determine the amount of solar radiation. These models are based on various parameters such as ambient temperature, sunshine duration, and cloudiness [27]. The most important parameter is the monthly average daily solar radiation values. It is necessary to know the monthly average daily value (H) of the global radiation to be used in solar applications. This value is available for many locations around the world. Another important parameter is the monthly average daily diffuse radiation (HD). However, since this value cannot be measured, it is estimated by developing some mathematical models. For reason, the diffuse fraction (KD) in equation (2), which is a function of the clearness index (KT) given in equation (1), is used for the estimation of diffuse radiation. The monthly average daily extraterrestrial radiation (H_0) is calculated by equation (3) [30].

$$\boldsymbol{K_T = \frac{H}{H_0}} \tag{1}$$

$$\boldsymbol{K_D = \frac{H_D}{H}} \tag{2}$$

$$\boldsymbol{H_0 = \frac{24}{\pi} I_{SC}\left(1 + 0.033cos\frac{360D}{365}\right) * \left(cos\varphi cos\delta sin\omega_s + \frac{2\pi\omega_s}{360} sin\varphi sin\delta\right)} \tag{3}$$

where I_{SC} is the solar constant, and its value is 1367 W/m^2. D is the number of days of the year since January, φ is the latitude angle, δ is the solar declination angle, and ω_S is the sunset hour angle. Solar declination angle and sunset hour angle can be calculated using equation (4) and equation (5), respectively. The irradiance values of the sun represent monthly average daily values [23].

$$\boldsymbol{\delta = 23.45sin\left[\frac{360D + 284}{365}\right]} \tag{4}$$

$$\boldsymbol{\omega_s = cos^{-1}[-\tan(\delta)\tan(\varphi)]} \tag{5}$$

Many statistical methods are used to compare solar radiation forecasting models. In this study, eight different statistical methods, which are widely used in the literature, were applied to test the performance of the developed diffuse radiation models. These are Mean Bias Error (MBE), Mean Absolute Error (MAE), Mean Squared Error

(MSE), Root Mean Squared Error (RMSE), Sum Squared Relative Error (SSRE), Relative Standard Error (RSE), coefficient of determination (R^2), and correlation coefficient (r).

MBE provides information on the long-term performance of the developed models. A negative MBE shows an underestimation, while a positive MBE indicates an overestimation. RMSE gives information about short-term performance, and its ideal value is zero [31].

$$\mathbf{MBE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i) \tag{6}$$

$$\mathbf{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \tag{7}$$

where n is the total number of observations, x_i is the measured data, and y_i is the calculated data. *r* is preferred to determine the relationship between measured and predicted values [27].

$$\mathbf{r} = \sqrt{(S_t - S_r)/S_t} \tag{8}$$

$$S_t = \sum_{i=1}^{n}(x_a - x_i)^2 \tag{9}$$

$$S_r = \sum_{i=1}^{n}(y_i - x_i)^2 \tag{10}$$

$$x_a = \frac{1}{n}\sum_{i=1}^{n}x_i \tag{11}$$

where S_t is the standard deviation and x_a is the average of the measured values.

MAE is an indicator of the fit used to obtain the models, and its ideal value is zero. MSE measures the mean square difference between the predicted values and the actual value. MSE is a risk function and is always positive. It can be said that estimators with MSE values close to zero perform better [18].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i| \tag{12}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 \tag{13}$$

SSRE is zero, and gives the positive value of the sum of the squares of the relative deviation. RSE shows the degree of accuracy of the estimate, and is an alternative measure of fit and always represents a proportion. R^2 is applied to determine the linear relationship between calculated and measured values, and it is desired to be close to 1 [32].

$$\mathbf{SSRE} = \sum_{i=1}^{n}\left(\frac{y_i - x_i}{x_i}\right)^2 \tag{14}$$

$$\mathbf{RSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\frac{y_i - x_i}{x_i}\right)^2}{n}} \tag{15}$$

$$\mathbf{R^2} = \frac{\sum_{i=1}^{n}(y_i - y_a).(x_i - x_a)}{\sqrt{[\sum_{i=1}^{n}(y_i - y_a)^2.\sum_{i=1}^{n}(x_i - x_a)^2]}} \tag{16}$$

2.1. Estimation of diffuse solar radiation

2.1.1. Data collection and climate characteristics

Türkiye is located between 36° and 42° latitudes in the northern hemisphere. Türkiye's annual solar energy potential is predicted to be around 380 GWh [33]. Konya is geographically located between 36.41' and 39.16' north latitudes and 31.14' and 34.26' east longitudes. Konya is one of the places with the highest solar radiation value. The solar radiation values of Konya are above the annual averages and reach the highest values.

First, diffuse and global solar radiation values were obtained from the Nasa-SSE web portal [34]. This data enables rapid evaluation of solar projects. The NASA-SSE dataset includes data from NASA satellites for 22 years (1983–2005). Secondly, the data for the years 2010 to 2020 was taken from the Turkish State Meteorological Service. Extraterrestrial solar radiation, H_0 was computed for each day, and then monthly averages were obtained using these values.

2.1.2. Estimation of diffuse radiation using NASA satellite data

Generally, the most widespread method is to obtain diffuse fraction (cloudiness index), KD by using the clearness index, KT. Then diffuse radiation values, HD is produced using the diffuse fraction value. Figure 1 shows the model development flowchart for diffuse radiation using NASA-SSE. Monthly average global, H, and diffuse, HD radiations are obtained from the NASA-SSE data set. Extraterrestrial solar radiation Ho is calculated and given in table 1. The clearness index, KT was computed using global and diffuse radiation data. Here, equations for diffuse fractions were developed utilizing a curve-fitting tool in MATLAB. The diffuse radiation was derived from these equations and compared with NASA-SSE data. Then the model that best represents the reference diffuse radiation values is selected.

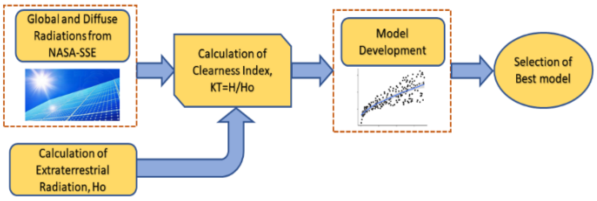


Figure 1. Model development flowchart for diffuse radiation using NASA-SSE.

Table 1. Solar radiations and clearness index.

Months	H_D(SSE) [MJ/m^2 day]	H (SSE) [MJ/m^2 day]	H_0 [MJ/m^2 day]	$K_T = H/H_0$
1	3.132	8.208	16.0612836	0.5110426
2	4.14	11.124	21.0140476	0.5293602
3	5.508	15.336	28.1370411	0.5450467
4	7.092	18.576	35.4719675	0.5236811
5	7.992	22.032	40.7576153	0.5405616
6	7.668	25.668	43.0257691	0.5965727
7	6.768	26.604	41.9896443	0.6335848
8	6.12	23.724	37.6852992	0.6295293
9	4.932	19.836	30.907469	0.6417866
10	4.284	13.572	23.2994694	0.5825025
11	3.384	9.072	17.2405714	0.5262007
12	2.88	6.84	14.5945841	0.468667

The models in the literature seem to have been developed depending on diffuse fraction and clearness index. The clearness index has more importance on diffuse radiation compared to the relative sunshine duration [25]. Thus seven empirical models were formed by considering the correlation between the clearness index and diffuse fraction. These are linear, quadratic, third and fourth-order polynomial, logarithmic, and exponential models. The model development is implemented using the MATLAB curve fitting toolbox. The models are as follows:

$$\mathbf{K_D = 0,9258 - 1,0526K_T} \qquad R^2 = 0.9988 \qquad (17)$$

$$\mathbf{K_D = 0.3529 + 0.9944K_T} \qquad R^2 = 0.9995 \qquad (18)$$

$$\mathbf{K_D = -07995 + 7.33064K_T - 13.54K_T^2 + 6.964K_T^3} \qquad R^2 = 0.9997 \qquad (19)$$

$$\mathbf{K_D = 2,0709e^{-3,272K_T}} \qquad R^2 = 0.9973 \qquad (20)$$

$$\mathbf{K_D = 0,1144K_T^{-1,821}} \qquad R^2 = 0.954 \qquad (21)$$

$$\mathbf{K_D = -0,069 - 0,587lnK_T} \qquad R^2 = 0.9976 \qquad (22)$$

$$\mathbf{K_D = 55.33 - 394.67K_T + 1081K_T^2 - 1311.5K_T^3 + 593.14K_T^4} \qquad R^2 = 0,9983 \qquad (23)$$

Table 2 shows the monthly average diffuse radiation values for seven equations and NASA-SSE. In table 2, the diffuse radiation produced from equation (19) varies between 3.189872 MJ/m^{2}day and 2.871569 MJ/m^{2}day. The statistical test results for diffuse radiation values are given in table 3. The third order polynomial model, Equation (19) shows the highest estimation accuracy with MBE = -0.00285085 MJ/m^2, MAE = 0.035414734 MJ/m^2, MSE = 0.001854687 MJ/m^2, RMSE = 0.043066072 MJ/m^2, SSRE = 0.000857583, RSE = 0.00845371, R^2 = 0.999705413, and r = 0.999687975. This is followed by Equations (18) and (17) with -0.00410764 and -0.006818814 MJ/m^2 of MBE, 0.041738416 and 0.074243664 MJ/m^2 of MAE, 0.00256197 and 0.007493633 MJ/m^2 of MSE, 0.050615905 and 0.086565772 of RMSE, 0.001018766 and 0.002979797 of SSRE, 0.009213965 and 0.01575806 of RSE, 0.999580532 and 0.99885125 of R^2, and 0.999568961 and 0.99873872 of r.

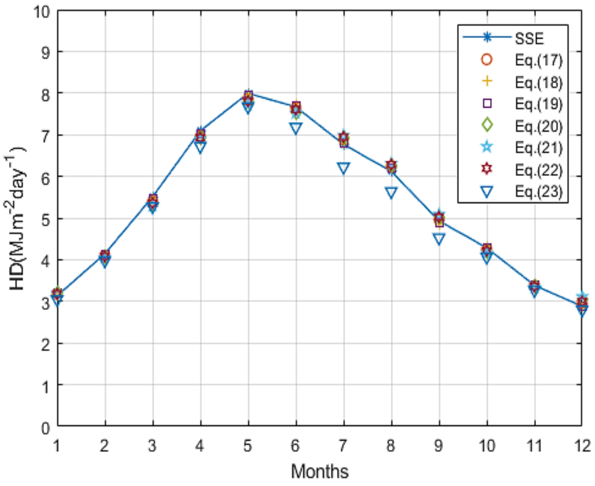


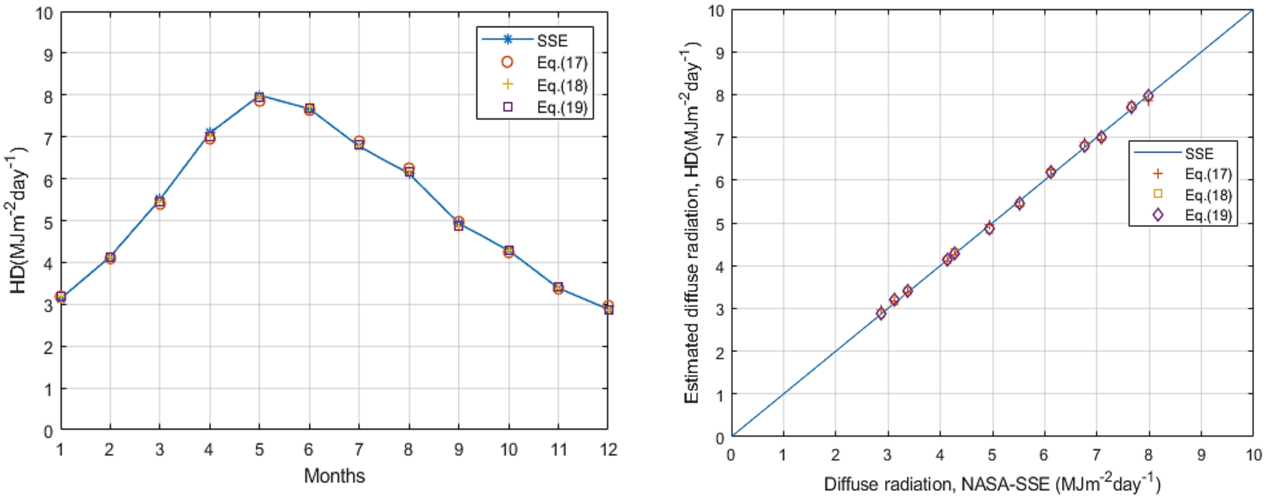
Figure 2. Diffuse radiations obtained using NASA-SSE data.

The bold values in table 3 represent the most suitable models determined according to statistical indicators. As it can be seen in table 3, the most suitable equations with NASA-SSE data are Equations (17), (18), and (19) considering the regression constants, the results of R^2, and the r correlation coefficient. Figure 2 shows diffuse radiation obtained from NASA-SSE data and developed models.

A comparison of equations (17), (18), and (19) which are the best models of diffuse radiation, is shown in figure 3. If it is necessary to define a model that provides the highest agreement with NASA-SSE data, equation (19) can be selected among these models. After that equation (18) and equation (17) are the ones with low error rates, respectively. As shown in figure 5, there are good agreements between the estimations and the

NASA-SSE data.

Figure 4 shows boxplots of statistical metrics. The boxplot typically displays quartiles (or percentages) and averages. This presents a visual distribution of numerical data and variability. Box charts are a good tool when comparing distributions of multiple groups or datasets, with the advantage of taking up less space [35]. The box plots in figure 4 clearly show that the estimation error of equation (19) is very low, and is the best diffuse radiation model. Equation (23) has the highest error rate among the developed models.



(a) **(b)**

Figure 3. Comparison of diffuse radiation for the equations (17), (18), and (19).

Table 2. Diffuse radiations calculated from the seven equations.

Months	H_D(SSE)	Eqnt. (17)	Eqnt. (18)	Eqnt. (19)	Eqnt. (20)	Eqnt. (21)	Eqnt. (22)	Eqnt. (23)
1	3.132	3.1836	3.1821	3.1898	3.1930	3.1883	3.1777	3.0452
2	4.14	4.1002	4.1310	4.1379	4.0756	4.0526	4.0767	3.9792
3	5.508	5.3995	5.4659	5.4672	5.3376	5.2977	5.3574	5.2772
4	7.092	6.9580	6.9949	7.0091	6.9335	6.9017	6.9253	6.7253
5	7.992	7.8611	7.9487	7.9542	7.7815	7.7262	7.8035	7.6734
6	7.668	7.6451	7.7268	7.6920	7.5476	7.5220	7.6058	7.1822
7	6.768	6.8874	6.7921	6.8003	6.9306	6.9869	6.9432	6.2259
8	6.12	6.2431	6.1814	6.1799	6.2629	6.3038	6.2810	5.6398
9	4.932	4.9640	4.8499	4.8742	5.0306	5.0888	5.0271	4.5286
10	4.284	4.2433	4.3037	4.2874	4.1788	4.1539	4.2117	4.0668
11	3.384	3.3740	3.3953	3.4017	3.3583	3.3412	3.3566	3.2673
12	2.88	2.9581	2.8783	2.8715	3.0565	3.1105	2.9956	2.7995

Table 3. Statistical test results for monthly average daily diffuse radiation values.

	Eqnt. (17)	Eqnt. (18)	Eqnt. (19)	Eqnt. (20)	Eqnt. (21)	Eqnt. (22)	Eqnt. (23)
MBE	-0.006818814	-0.00410764	**-0.00285085**	-0.01773166	-0.01881961	-0.01148105	-0.29076208
MAE	0.074243664	0.041738416	**0.035414734**	0.124708475	0.159909865	0.110281619	0.290762082
MSE	0.007493633	0.00256197	**0.001854687**	0.01835336	0.030114738	0.015075771	0.109400763
RMSE	0.086565772	0.050615905	**0.043066072**	0.135474573	0.173535984	0.122783433	0.330757862
SSRE	0.002979797	0.001018766	**0.000857583**	0.008957568	0.014863803	0.006064889	0.036103266
RSE	0.01575806	0.009213965	**0.00845371**	0.027321494	0.035194464	0.022481268	0.054850757
R^2	**0.99885125**	**0.999580532**	**0.999705413**	0.997345827	0.995433726	0.997644582	0.998352372
r	**0.99873872**	**0.999568961**	**0.999687975**	0.996908329	0.994922111	0.997460994	0.981943602

Table 4. Solar radiations and clearness index.

Months	Mean of 12 equations (H_D)	Measured, H	Ho	$K_T = H/H_O$	$KD = H_D/H$
1	3.473722	7.5927618	16.0612836	0.472736924	0.457504
2	4.335037	12.09514185	21.01404757	0.575574116	0.358411
3	5.864856	15.8501781	28.13704113	0.563320714	0.370018
4	7.229537	20.8638711	35.47196753	0.588179133	0.34651
5	8.48711	23.01012945	40.75761532	0.564560249	0.368842
6	8.489352	26.5097709	43.02576907	0.616137061	0.320235
7	7.694236	27.85216365	41.98964435	0.663310301	0.276253
8	7.05005	24.5576754	37.68529918	0.651651332	0.287081
9	5.744565	20.25730845	30.907469	0.655417901	0.28358
10	4.709259	13.8876156	23.2994694	0.596048578	0.339098
11	3.565807	9.87299775	17.24057141	0.572660703	0.361168
12	3.146526	7.23950055	14.59458408	0.49604021	0.434633

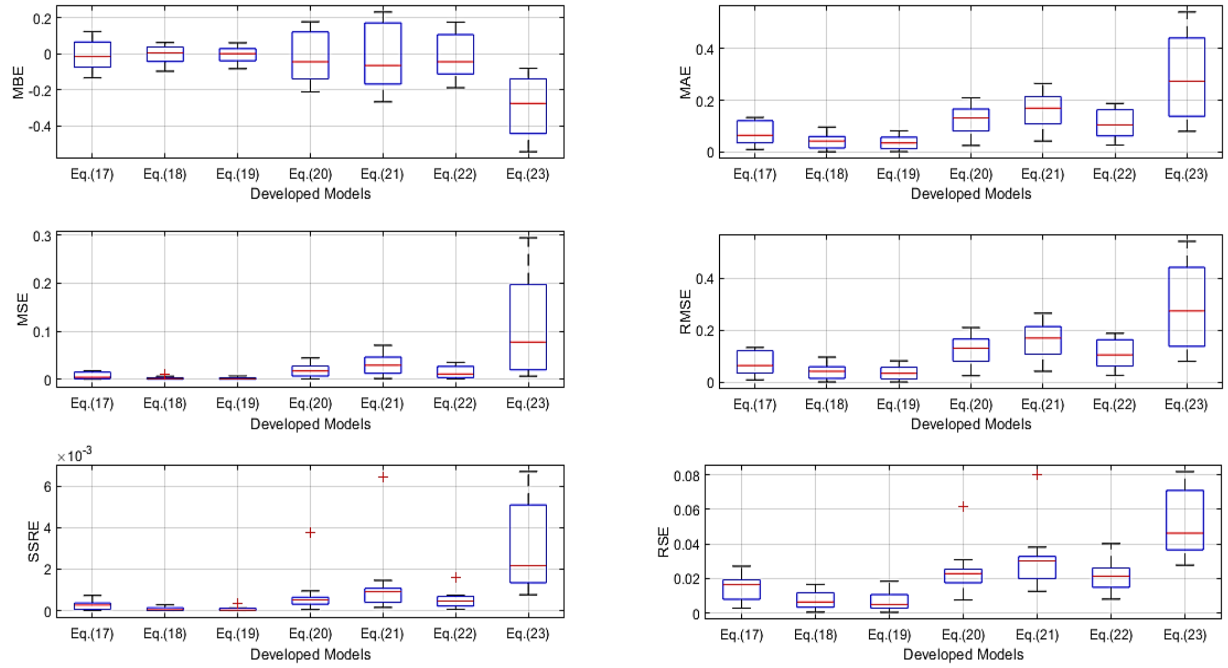


Figure 4. Boxplot plots for equations (17)-(23).

2.1.3. Estimation of diffuse radiation from existing models

New diffuse radiation models have been developed by using existing empirical models for Türkiye and other localizations in the world. Figure 5 shows the flowchart for the estimation of diffuse radiation. Equations (24) to (35) are the empirical models obtained from the literature. While selecting the existing empirical models from different regions of Türkiye and the world, attention was paid to having similar climatic conditions to be compatible with the reference data. The diffuse radiation values, H_D were generated by using twelve existing empirical models from the literature. The generated values are averaged to form reference diffuse solar radiation data. Global solar radiation, *H* was obtained from Turkish State Meteorological. Extraterrestrial solar radiation H_o is calculated. The diffuse fraction, *KD* was obtained by curve fitting methods using the clearness index, *KT*. The solar radiations, diffuse fraction, and clearness index are given in table 4.

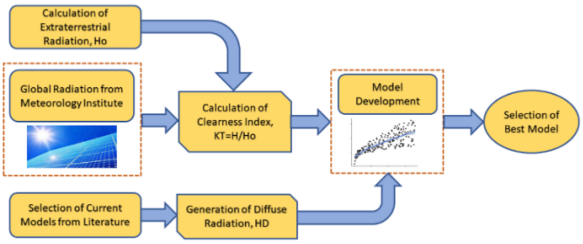


Figure 5. Model development flowchart for diffuse radiation.

The empirical models taken from the literature, which are thought to best represent Konya and the Central Anatolian region, are given below:

- Model 1 [36]:
  $\mathbf{K_D = 0.91 - 0.847K_T}$
- Model 2 [21]:
  $\mathbf{K_D = 1.1244 - 1.5582K_T + 0.3635K^2{}_T}$
- Model 3 [24]:
  $\mathbf{K_D = 0.9885 - 1.4276K_T + 0.5679K^2{}_T}$
- Model 4 [22]:
  $\mathbf{K_D = 1.7821 - 6.648K_T + 11.17K^2{}_T}$
  $\mathbf{-6.5641K^3{}_T}$
- Model 5 [37]:
  $\mathbf{K_D = 1.0 - 1.13K_T}$
- Model 6 [38]:
  $\mathbf{K_D = 1.03 - 1.17K_T}$
- Model 7 [39]:
  $\mathbf{K_D = 1.317 - 3.023K_T + 3.372K^2{}_T}$
  $\mathbf{-1.769K^3{}_T}$
- Model 8 [40]:
  $\mathbf{K_D = 1.6912 - 8.2262K_T + 25.5532K^2{}_T}$ -
  $\mathbf{37.8070K^3{}_T+19.8178K^4{}_T}$
- Model 9 [41]:
  $\mathbf{K_D = 09138 - 0.96225K_T}$
- Model 10 [24]:
  $\mathbf{K_D = 1.027 - 1.6582K_T + 1.1018K^2{}_T}$
  $\mathbf{-0.4019K^3{}_T}$
- Model 11 [42]:
  $\mathbf{K_D = 1.0896 - 1.4797K_T + 0.1471K^2{}_T}$
- Model 12 [43]:
  $\mathbf{K_D = 1.390 - 4.027K_T + 5.531K^2{}_T - 3.108K^3{}_T}$

The model development is implemented on the MATLAB curve fitting toolbox. Seven new models were proposed for the estimation of diffuse solar radiation. These are linear, quadratic, third and fourth-order polynomial, logarithmic, and exponential models. The developed models are based on the diffuse fraction, which is a function of the clearness index. The models are as follows:

$$\mathbf{K_D = 0,9052 - 0,9493K_T} \qquad \mathbf{R^2 = 0.9999} \tag{36}$$

$$\mathbf{K_D = 0,9509 - 1,1105K_T + 0,1409K_T{}^2} \qquad \mathbf{R^2 = 1} \tag{37}$$

$$\mathbf{K_D=1.0198-1.4782K_T+0.7895K_T{}^2-0.3785K_T{}^2} \qquad \boldsymbol{R^2 = 1} \tag{38}$$

$$\mathbf{K_D = 1,6437e^{-2,664K_T}} \qquad \mathbf{R^2 = 0.9993} \tag{39}$$

$$\mathbf{K_D = 0,1535K_T{}^{-1,502}} \qquad \mathbf{R^2 = 0.9980} \tag{40}$$

$$\mathbf{K_D = 0,0591 - 0,538lnK_T} \qquad \mathbf{R^2 = 0.9996} \tag{41}$$

$$\mathbf{K_D=1.1884-2.6792K_T+3.9838K_T{}^2-4.1375K_T{}^3+1.6515K_T{}^4} \qquad \boldsymbol{R^2 = 1} \tag{42}$$

Table 5 shows the diffuse radiation values for seven equations. In table 5, diffuse radiation values produced from equation (42) vary between 3,473759 MJ/m^{2}day and 3,146557 MJ/m^{2}day.

Table 5. Diffuse radiation values calculated from equations (36)-(42).

Months	Mean of 12 equations (H_D)	Eqt. (36)	Eqt. (37)	Eqt. (38)	Eqt. (39)	Eqt. (40)	Eqt. (41)	Eqt. (42)
1	3.4737218	3.465571	3.473036	3.473312	3.509779	3.563846	3.591118	3.473759
2	4.3350371	4.339828	4.334936	4.334465	4.309978	4.316698	4.256447	4.335063
3	5.8648555	5.871534	5.865266	5.864112	5.83158	5.844556	5.761127	5.864895
4	7.2295370	7.236457	7.228749	7.22863	7.19141	7.200324	7.107221	7.229575
5	8.4871104	8.496788	8.487624	8.486026	8.438647	8.456727	8.336017	8.487167
6	8.4893524	8.491108	8.487605	8.488655	8.47502	8.492136	8.422073	8.489382
7	7.6942355	7.673815	7.695195	7.692671	7.798514	7.868507	7.92043	7.694235
8	7.0500497	7.03792	7.049878	7.049279	7.110404	7.156648	7.172073	7.050056
9	5.7445646	5.733057	5.744673	5.743804	5.802457	5.844486	5.865157	5.744568
10	4.7092585	4.713055	4.708543	4.708713	4.687497	4.693321	4.637277	4.709281
11	3.5658073	3.569811	3.565802	3.565338	3.5451	3.55108	3.501028	3.565829
12	3.1465259	3.14418	3.147131	3.146411	3.159036	3.193499	3.185302	3.146557

Table 6 shows statistical test results for monthly average daily diffuse radiation. The 4th order polynomial model, equation (42) showed the highest estimation accuracy with MBE = 2.59E-05 MJ/m^2, MAE = 2.60E-05 MJ/m^2, MSE = 9.25E-10 MJ/m^2, RMSE = 3.04E-05 MJ/m^2, SSRE = 4.39E-10, RSE = 6.05E-06, R^2 = 1, and r = 1. This is followed by equations (37) and (38) with -0.00013 and -0.00072 MJ/m^2 of MBE, 0.000568 and 0.00072 MJ/m^2 of MAE, 5.36E-07 and 6.40E-07 MJ/m^2 of MSE, 0.000732 and 0.0008002 of RMSE, 1.79E-07 and 1.89E-07 of SSRE, 0.000122 and 0.0001255 of RSE, 1 and 1 of R^2, and 1 and 0.9999999 of r. The bold values in table 6 represent the most suitable models determined according to statistical indicators. As it can be seen in table 6, the most compatible equations with literature models are equations (37), (38), and (42) considering the statistical test results. If it is necessary to define a model that provides the highest agreement with literature models, equation (42) can be selected among these models. After that equation (37) and equation (38) are the ones with low error rates, respectively.

Figure 6 shows diffuse radiation obtained from developed models. Figure 7 shows the comparison of equations (37), (38), and (42) which are the best models of diffuse radiation. In figure 7, the values obtained from the models and literature data overlap with each other. The average diffuse radiation value was generated using the empirical models available in the literature, and this generated value was named as literature. The developed models were compared with this reference diffuse radiation value, and their performance was evaluated.

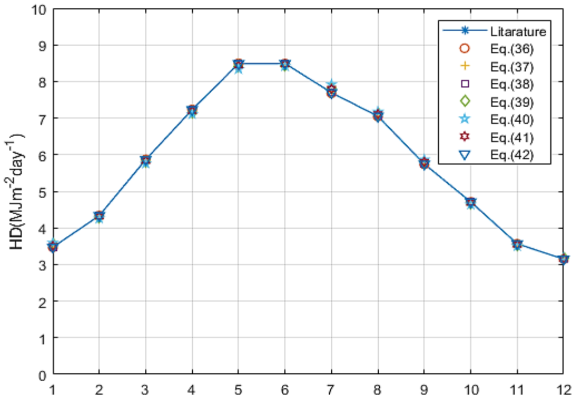
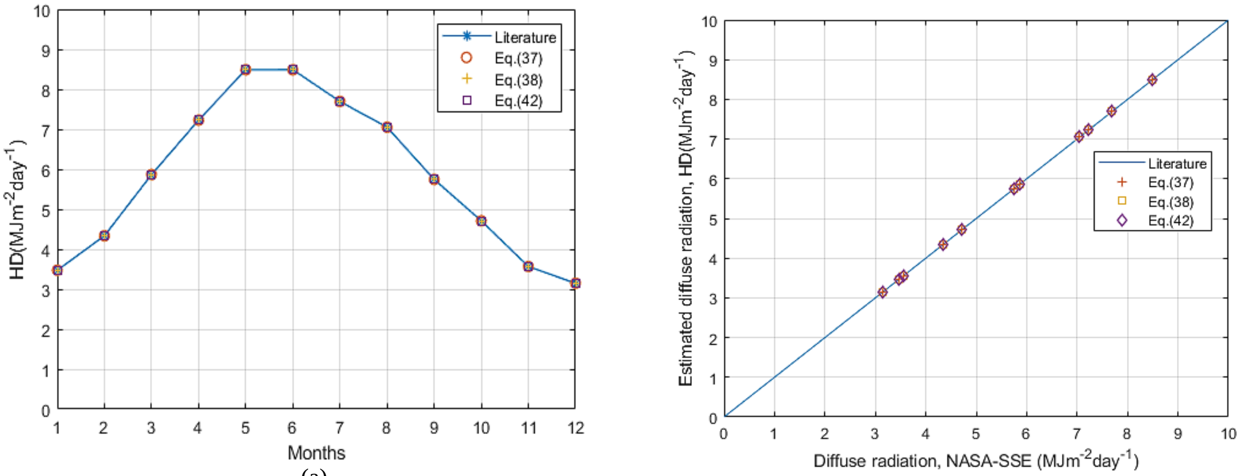


Figure 6. Diffuse radiation values for literature and developed models.

Figure 8 shows boxplots of statistical metrics. As seen in figure 8, the best model developed is equation (42). This is followed by the second-best model, equation (37), and the third-best model, equation (38). In addition to equation (40) has the highest error rate among the developed models.

Table 6. Statistical test values for equations (36)-(42).

	Eqt. (36)	Eqt. (37)	Eqt. (38)	Eqt. (39)	Eqt. (40)	Eqt. (41)	Eqt. (42)
MBE	-0.00141	-0.00013	-0.00072	-0.00272	-0.0029	0.00489	**2.59E-05**
MAE	0.007682	0.000568	0.00072	0.061191	0.107063	0.039604	**2.60E-05**
MSE	8.43E-05	5.36E-07	6.40E-07	0.004362	0.013709	0.002155	**9.25E-10**
RMSE	0.009179	0.000732	0.0008002	0.066044	0.117087	0.046421	**3.04E-05**
SSRE	2.68E-05	1.79E-07	1.89E-07	0.001558	0.00477	0.000663	**4.39E-10**
RSE	0.001494	0.000122	0.0001255	0.011395	0.019937	0.007431	**6.05E-06**
R^2	0.999988	**1**	**1**	0.999389	0.998062	0.999699	**1**
r	0.999988	**1**	**0.9999999**	0.999383	0.99806	0.999695	**1**



(a)

(b)

Figure 7. Comparison of the diffuse radiation values for equations (37, 38, 42).

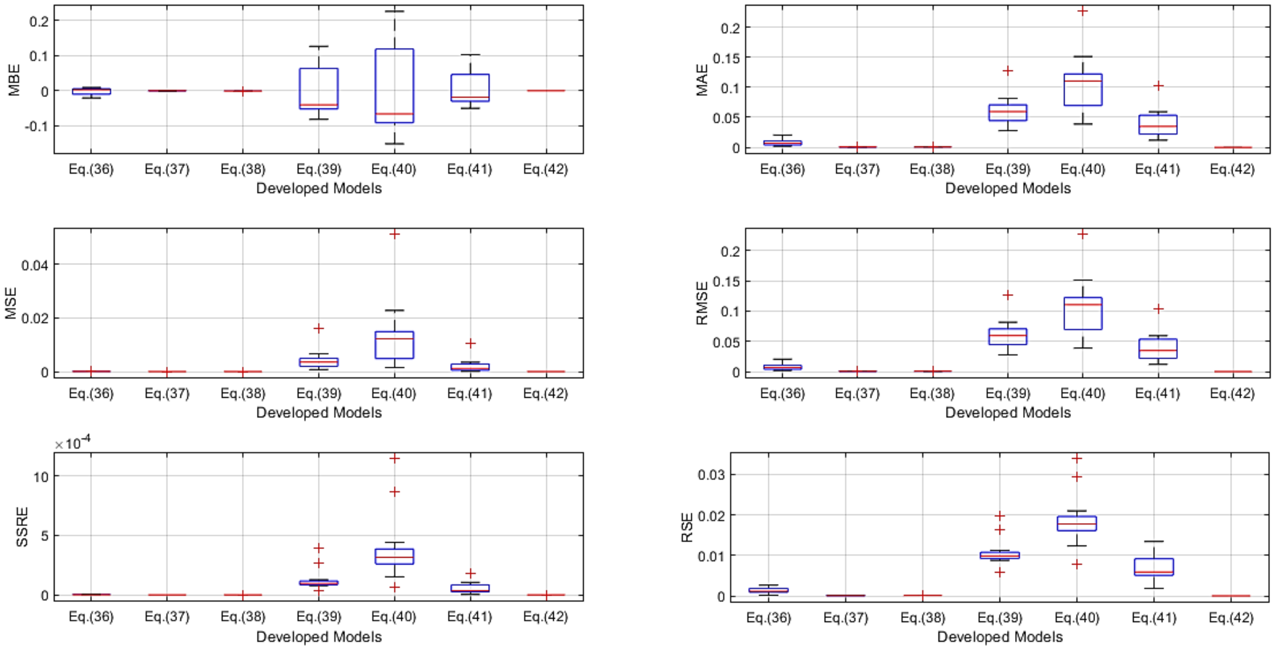


Figure 8. Boxplot plots for equations (36)-(42).

3. Findings and Discussion

In solar energy research, it is important to be able to decompose the diffuse radiation component from global radiation. Therefore, constructing an accurate and effective model is valuable in diffuse radiation. In this study, linear and non-linear correlations based on diffuse fraction and clearness index have been developed for diffuse solar radiation. In this context, 14 models were established to predict diffuse radiation using existing models and NASA-SSE data. The performances of the developed models were evaluated and compared with eight statistical metrics. To decide whether the developed models are a more accurate model, it is checked that the errors have low values, the coefficient of determination, and the correlation coefficient is close to 1. Equation (19) is the most accurate model obtained using NASA-SSE data. Equations (37) and (42) are the best diffuse radiation models proposed using existing models in the literature. The proposed models are given below.

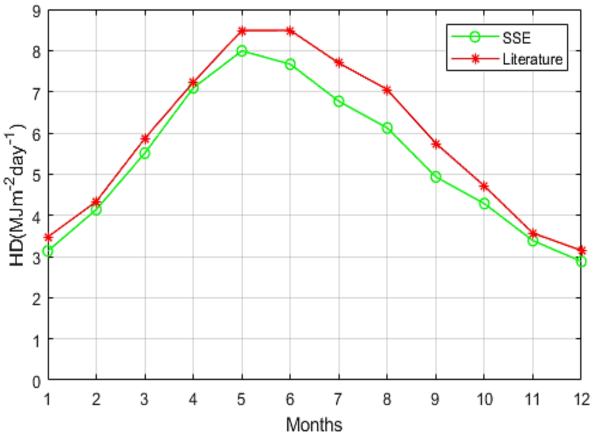
$$\mathbf{K_D = -07995 + 7{,}33064K_T - 13{,}54K_T^2 + 6{,}964K_T^3}$$

$$\mathbf{K_D = 0{,}9509 - 1{,}1105K_T + 0{,}1409K_T^2}$$

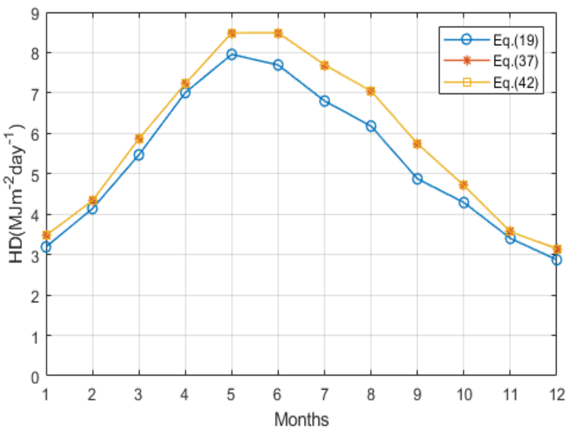
$$\mathbf{K_D = 1{,}1884 - 2{,}6792K_T + 3{,}9838K_T^2 - 4{,}1375K_T^3 + 1{,}6515K_T^4}$$

Figure 9a shows diffuse solar radiations for NASA-SSE and literature. The average diffuse radiation value was calculated using the empirical models available in the literature, and this value was named literature. Then the newly developed models were compared with this reference diffuse radiation value. Figure 9b shows the general comparison of developed models. In figure 9b, the developed models represent the best compromised diffuse radiations by producing the results closest to the reference diffuse radiation values in figure 9a.

The second, third, and fourth-order polynomial equations give better results, whereas logarithmic and exponential equations are relatively unsuccessful in estimating diffuse radiation. The models found to predict diffuse solar radiations are optimal solutions. A new approach to estimating diffuse solar radiation using both literature existing models and NASA-SSE data is presented. It is hoped that the models developed will help researchers, and investors, and will provide significant insight into this field.



(a)



(b)

Figure 9. (a) Diffuse radiations for NASA-SSE and literature b) Comparison of the best models.

4. Conclusion

It is necessary to know the diffuse radiation value in evaluating the suitability of solar energy systems to be installed in any part of the world, which is one of the components of solar radiation. In this paper, we proposed empirical models to improve the diffuse radiation estimation results. The global radiation and NASA-SSE data were used in the construction of empirical models. Two approaches were presented to construct a diffuse solar radiation model. First, seven new models were developed using NASA-SSE data. Secondly, empirical models were chosen from the literature, and the diffuse radiation values obtained from these empirical models were averaged. Then new models were constructed using the average values of diffuse radiation and the measured solar radiation values. The developed models were validated using eight statistical test methods: These are Mean Bias Error (MBE), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Sum Squared Relative Error (SSRE), Relative Standard Error (RSE), Coefficient of determination (R^2) and Correlation Coefficient (r). The best three models were proposed. Equation (19) is the most accurate model obtained using NASA-SSE data. Equations (37) and (42) are the best diffuse radiation models proposed using existing models in the literature.

The third order polynomial model, equation (19) showed the highest estimation accuracy with MBE = -0.00285085 MJ/m^2, MAE = 0.035414734 MJ/m^2, MSE = 0.001854687 MJ/m^2, RMSE = 0.043066072 MJ/m2, SSRE = 0.000857583, RSE = 0.00845371, R2 = 0.999705413, and r = 0.999687975.

Another second-order polynomial model, equation (37) showed the highest estimation accuracy with -0.00013 MJ/m^2 of MBE, 0.000568 MJ/m^2 of MAE, 5.36E-07 of MSE, 0.000732 of RMSE, 1.79E-07 of SSRE, 0.000122 of RSE, 1 of R^2, and 1 of r.

This is followed by the 4th order polynomial model, equation (42) with MBE = 2.59E-05 MJ/m^2, MAE = 2.60E-05 MJ/m^2, MSE=9.25E-10 MJ/m^2, RMSE = 3.04E-05 MJ/m^2, SSRE = 4.39E-10, RSE = 6.05E-06, R^2 = 1, and r = 1.

These three models with the lowest statistical errors yielded accurate results compared to other models. According to the results of both box plot analysis and statistical test methods, the proposed models are quite compatible with the reference diffuse radiation data. New models could be used to predict diffuse solar radiation possibly elsewhere in similar conditions.

As the computational resources become more powerful, the accuracy and reliability of the predictions will increase, and thus the models representing diffuse radiation will continue to be improved. It is also important to note that the choice of diffuse radiation estimation method may vary depending on the specific application and available data. Future research may focus on hybridizing multiple methods and combining satellite data with ground-based observations to improve the accuracy of diffuse radiation estimates. Depending on the availability of local atmospheric data, prediction methods can be developed that integrate empirical and radiation transfer models. Popular machine learning algorithms such as deep learning and support vector machines can be used for diffuse radiation estimation.

5. Nomenclature

H	Monthly average daily global radiation	MJ/m^{2}day
H_D	Monthly average daily diffuse radiation	MJ/m^{2}day
H_0	Monthly average daily extraterrestrial radiation	MJ/m^{2}day
I_{SC}	Solar constant	W/m^2
D	Number of days of the year	
φ	Latitude of the place	(^{O})
δ	Solar declination	(^{O})
ω_S	Sunset hour angle	(^{O})
K_T	Monthly average daily clearness index	
K_D	Monthly average daily diffuse fraction index	
MBE	Mean bias error	
MABE	Mean absolute bias error	

MSE	Mean squared error
RMSE	Root mean squared error
SSRE	Sum squared relative error
RSE	Relative standard error
r	Correlation coefficient
R^2	Coefficient of determination

6. References

[1] K. Bakirci, “Prediction of diffuse radiation in solar energy applications: Turkey case study and compare with satellite data,” Energy, Vol. 237, pp. 121527, 2021.

[2] NASA. "Earth’s Energy Budget
" June 20, 2021; https://earthobservatory.nasa.gov/features/EnergyBalance/page4.php.

[3] H. C. Bayrakçı, C. Demircan, and A. Keçebaş, “The development of empirical models for estimating global solar radiation on horizontal surface: A case study,” Renewable and Sustainable Energy Reviews, Vol. 81, pp. 2771-2782, 2018.

[4] S. F. Khahro, K. Tabbassum, S. Talpur, M. B. Alvi, X. Liao, and L. Dong, “Evaluation of solar energy resources by establishing empirical models for diffuse solar radiation on tilted surface and analysis for optimum tilt angle for a prospective location in southern region of Sindh, Pakistan,” International Journal of Electrical Power & Energy Systems, Vol. 64, pp. 1073-1080, 2015.

[5] J. Fan, L. Wu, F. Zhang, H. Cai, X. Ma, and H. Bai, “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, Vol. 105, pp. 168-186, 2019.

[6] J. Boland, B. Ridley, and B. Brown, “Models of diffuse solar radiation,” Renewable Energy, Vol. 33, No. 4, pp. 575-584, 2008.

[7] R. Tang and T. Wu, “Optimal tilt-angles for solar collectors used in China,” Applied energy, Vol. 79, No. 3, pp. 239-248, 2004.

[8] F. Cao, H. Li, T. Yang, Y. Li, T. Zhu, and L. Zhao, “Evaluation of diffuse solar radiation models in Northern China: New model establishment and radiation sources comparison,” Renewable Energy, Vol. 103, pp. 708-720, 2017.

[9] L. Wang, Y. Lu, L. Zou, L. Feng, J. Wei, W. Qin, and Z. Niu, “Prediction of diffuse solar radiation based on multiple variables in China,” Renewable and Sustainable Energy Reviews, Vol. 103, pp. 151-216, 2019.

[10] M. U. Yousuf, M. Umair, and M. Uzair, “Estimating the average diffuse solar radiation based on multiple parameters: a case study of arid climate zone of Pakistan,” International Journal of Ambient Energy, pp. 1-11, 2020.

[11] S. C. Nwokolo, A. U. Obiwulu, J. C. Ogbulezie, and S. O. Amadi, “Hybridization of statistical machine learning and numerical models for improving beam, diffuse and global solar radiation prediction,” Cleaner Engineering and Technology, Vol. 9, pp. 100529, 2022.

[12] S. E. Berrizbeitia, E. Jadraque Gago, and T. Muneer, “Empirical models for the estimation of solar sky-diffuse radiation. A review and experimental analysis,” Energies, Vol. 13, No. 3, pp. 701, 2020.

[13] T. Zhu, J. Li, L. He, D. Wu, X. Tong, Q. Mu, and Q. Yu, “The improvement and comparison of diffuse radiation models in different climatic zones of China,” Atmospheric Research,Vol. 254, pp. 105505, 2021.

[14] L. Al-Ghussain, O. Al-Oran, and F. Lezsovits, “Statistical estimation of hourly diffuse radiation intensity of Budapest City,” Environmental Progress & Sustainable Energy, Vol. 40, No. 1, pp. e13464, 2021.

[15] F. Li, Y. Lin, J. Guo, Y. Wang, L. Mao, Y. Cui, and Y. Bai, “Novel models to estimate hourly diffuse radiation fraction for global radiation based on weather type classification,” Renewable Energy, Vol. 157, pp. 1222-1232, 2020.

[16] B. Pavlinić, and D. Tolić, "Predicting Diffuse Horizontal Radiation with Neural Networks." pp. 127-132.

[17] N. Chawphongphang, P. Chaiwiwatworakul, and S. Chirarattananon, “Global and Diffuse Solar Radiation Characteristics of Bangkok and its Forecast Using Artificial Neural Network,” International Energy Journal, Vol. 22, No. 4, 2022.

[18] Y. Lu, R. Zhang, L. Wang, X. Su, M. Zhang, H. Li, S. Li, and J. Zhou, “Prediction of diffuse solar radiation by integrating radiative transfer model and machine-learning techniques,” Science of The Total Environment, Vol. 859, pp. 160269, 2023.

[19] L. Yang, Q. Cao, Y. Yu, and Y. Liu, “Comparison of daily diffuse radiation models in regions of China without solar radiation measurement,” Energy, Vol. 191, pp. 116571, 2020.

[20] M. Tiris, C. Tiris, and I. Türe, “Correlations of monthly-average daily global, diffuse and beam radiations with hours of bright sunshine in Gebze, Turkey,” Energy Conversion and Management, Vol. 37, No. 9, pp. 1417-1421, 1996.

[21] H. Aras, O. Balli, and A. Hepbasli, “Estimating the horizontal diffuse solar radiation over the Central Anatolia Region of Turkey,” Energy conversion and management, Vol. 47, No. 15-16, pp. 2240-2249, 2006.

[22] K. Ulgen and A. Hepbasli, “Comparison of the diffuse fraction of daily and monthly global radiation for Izmir, Turkey,” Energy Sources, Vol. 25, No. 7, pp. 637-649, 2003.

[23] K. Ulgen and A. Hepbasli, “Diffuse solar radiation estimation models for Turkey’s big cities,” Energy Conversion and management, Vol. 50, No. 1, pp. 149-156, 2009.

[24] S. Tarhan, and A. Sarı, “Model selection for global and diffuse radiation over the Central Black Sea (CBS) region of Turkey,” Energy Conversion and Management, Vol. 46, No. 4, pp. 605-613, 2005.

[25] N. Arslanoglu, “Evaluation and establishment of diffuse solar radiation models for Bursa, Turkey,” Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, Vol. 38, No. 18, pp. 2788-2797, 2016.

[26] C. A. Tırmıkçı and C. Yavuz, “Establishing New Diffuse Solar Radiation Equations for Antalya, Turkey,” Academic Perspective Procedia, Vol. 1, No. 1, pp. 514-520, 2018.

[27] K. Bakirci and Y. Kirtiloglu, “Prediction of diffuse solar radiation using satellite data,” International Journal of Green Energy, Vol. 15, No. 2, pp. 76-79, 2018.

[28] S. E. Rusen and A. Konuralp, “Quality control of diffuse solar radiation component with satellite-based estimation methods,” Renewable Energy, Vol. 145, pp. 1772-1779, 2020.

[29] K. Bakirci, “Evaluation of models for prediction of diffuse solar radiation and comparison with satellite values,” Journal of Cleaner Production, Vol. 374, pp. 133892, 2022.

[30] J. A. Duffie and W. A. Beckman, “Solar Engineering of Thermal Processes John Wiley & Sons,” Inc. New York, 1991.

[31] H. Bulut, and O. Büyükalaca, “Simple model for the generation of daily global solar-radiation data in Turkey,” Applied Energy, Vol. 84, No. 5, pp. 477-491, 2007.

[32] A. G. Kaplan and Y. A. Kaplan, “Developing of the new models in solar radiation estimation with curve fitting based on moving least-squares approximation,” Renewable Energy, Vol. 146, pp. 2462-2471, 2020.

[33] Republic of Turkey. "Ministry of Energy and Natural Resources, Republic of Turkey.," June 20, 2021; https://enerji.gov.tr/bilgi-merkezi-enerji-gunes/.

[34] NASA-SSE. "Surface meteorology and Solar Energy," June 20, 2021; https://power.larc.nasa.gov/data-access-viewer/.

[35] S. ÇETİNTÜRK, N. TARAKÇIOĞLU, M. ACARER, M. Çunkaş, and E. SALUR, “Using Quadratic Multiple Linear Regression Models to Investigate the Effect of Inoculant Type and T6 Heat Treatment on Microstructural, Mechanical and Corrosion Properties of Al-Cu Alloy Produced by Casting,” Materials Today Communications, pp. 105549, 2023.

[36] M. Iqbal, “Study of canadian diffuse and total solar radiation data. i. monthly average daily horizontal radiation,” Sol. Energy;(United States), Vol. 22, No. 1, 1979.

[37] J. K. Page, "The estimation of monthly mean values of daily total short wave radiation on vertical and inclined surfaces from sunshine records for latitudes 40ıN–40ıS." pp. 378–390.

[38] C. Jacovides, L. Hadjioannou, S. Pashiardis, and L. Stefanou, “On the diffuse fraction of daily and monthly global radiation for the island of Cyprus,” Solar Energy, Vol. 56, No. 6, pp. 565-572, 1996.

[39] D. Erbs, S. Klein, and J. Duffie, “Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation,” Solar energy, Vol. 28, No. 4, pp. 293-302, 1982.

[40] E. Taşdemiroğlu, and R. Sever, “Estimation of monthly average, daily, horizontal diffuse radiation in Turkey,” Energy, Vol. 16, No. 4, pp. 787-790, 1991.

[41] K. Gopinathan and A. Soler, “Diffuse radiation models and monthly-average, daily, diffuse data for a wide latitude range,” Energy, Vol. 20, No. 7, pp. 657-667, 1995.

[42] S. Barbaro, G. Cannata, S. Coppolino, C. Leone, and E. Sinagra, “Diffuse solar radiation statistics for Italy,” Solar Energy, Vol. 26, No. 5, pp. 429-435, 1981.

[43] B. Y. Liu and R. C. Jordan, “The interrelationship and characteristic distribution of direct, diffuse and total solar radiation,” Solar energy, Vol. 4, No. 3, pp. 1-19, 1960.