Artificial Neural Network Models for Solar Radiation Estimation Based on Meteorological Data

Ersan Omer Yuzer[a] and Altug Bozkurt[b*]

[a] Department of Electricity and Energy, Çölemerik Vocational High School, Hakkari University, Hakkari, Turkey, E-mail: ersanomeryuzer@hakkari.edu.tr

[b] Department of Electrical Engineering, Faculty of Electrical and Electronics, Yildiz Technical University, Istanbul, Turkey, E-mail: abozkurt@yildiz.edu.tr
*Corresponding author

Abstract: The presence of solar energy in a particular area is closely related to meteorological parameters in that region. In this study, solar radiation estimation was carried out by using meteorological data recorded in different time series. Artificial neural networks (ANN) models were developed to determine the most effective parameters for solar radiation estimation. During the training and testing of ANN, site-specific meteorological data recorded by a meteorological station established in Hakkâri, Turkey, which has difficult climatic conditions, were used. To estimate solar radiation, basic input variables such as ambient temperature (T), wind speed (w), relative humidity (H), and atmospheric pressure (P), were modified by keeping the time series constant. To obtain the best estimation result, the number of input parameters of the input layer was applied with different possible input combinations, and the hidden layer neuron was changed to be multiples of the input layer (n, 2n, n^2). The performance of all models was analyzed using statistical tools. ANN model, which has all possible combinations of input variables and determines the number of neurons in the hidden layer by framing the number of input variables, yielded the best estimation result. The performance indicator showed the mean square error (MSE) as the lowest value of 2.56 with all data entries and modeling the number of neurons in the hidden layer as n^2. The mean absolute percentage error (MAPE) and relative root mean square error (rRMSE) values were obtained within the limits of high estimation accuracy in the network combination of T, P and H parameters as 1.99% and 1.91%, respectively. This study has revealed that increasing the variety and number of meteorological parameters affects solar radiation estimation success, but only basic meteorological parameters achieve very high estimation results.

Keywords: Artificial neural networks; estimation; meteorological data; photovoltaic; solar radiation

1 Introduction

In recent years, electricity generation based on PV technology has become one of the most sought-after fields [1]. Local knowledge of solar radiation is essential for many applications, and solar radiation estimation has become more important for energy and smart grid applications today [2]. Solar radiation can be obtained from meteorological stations and measuring devices, as well as empirical relationships can be estimated using artificial intelligence techniques and hybrid techniques [3-7]. Since it is difficult to reach solar system power data, researchers use solar radiation data in their estimation studies [8]. ANN models are powerful tools considered high-performance estimation models for estimating solar radiation [9,10]. However, there are also various ANN model structures that have proven successful in modeling systems across different fields [11-15]. These studies make significant contributions to the literature.

ANN, which are machine learning approaches, have become one of the most widely used models due to their higher accuracy and higher application areas compared to physical, persistence and statistical models. For this reason, many researchers have used ANN models to estimate solar radiation. Solar radiation can be affected by some or all of the parameters such as latitude, longitude, altitude, month, year, day, ambient temperature, wind speed, relative humidity and precipitation [16]. In recent years, parameters used in solar radiation estimation studies in many countries around the world are given in Table 1.

Tablo 1

Parameters used in solar radiation estimation studies

Parameters / **Reference No**	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]	**Present Study**
Second-[s]																	√
Minute-[m]																	√
Hour-[h]											*						√
Days-[D]			*	*				*	*					*			√
Month-[M]	*			*					*								√
Year-[Y]	*								*					*			√
Latitude (°)	*			*				*									
Longitude (°)	*			*				*									
Altitude (m)	*							*									
Ambient temp.(°C)-[T]	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	√
Wind Speed (m/sn)-[w]	*		*	*		*		*	*	*	*	*	*	*	*	*	√
Mean relative humidity (%)-[H]	*	*	*	*	*	*	*	*	*	*		*	*	*	*	*	√
Mean station level press. (mb)-[P]	*	*		*								*		*	*	*	√
Sunshine duration (hour)					*		*	*		*							
Sun hours (hour)		*	*														
Sun declination (°)								*						*			

Sunrise hourly angle (°)														*			
Extraterrestrial radiation (W/m^2)				*					*					*			
Clearness index (KI)							*	*					*				
Zenith Angel (°)											*						
Rainfall									*							*	

The studies presented in Table 1 have used basic meteorological parameters. However, the impact of the temporal variations of these parameters on solar radiation determination is not clear. Moreover, many studies using time parameters do not include minutes and seconds. This study has ensured the inclusion of these parameters, leading to more successful estimation results in solar radiation estimation. Additionally, the effectiveness of basic meteorological parameters in estimation solar radiation over time has been determined.

Compared to traditional models, ANN models are widely used in solar radiation estimation studies and can yield more accurate estimation depending on algorithm training [33, 34]. Upon reviewing the literature, numerous ANN models have been presented for solar radiation estimation using location and meteorological parameter data [35-42].

The output power of a photovoltaic (PV) system is highly intermittent and fluctuating due to the change of solar radiation depending on the meteorological parameter of nature, such as T, H, w, and P [43]. Therefore, accurate estimation of solar radiation can provide better planning for grid applications, energy management, operational cost minimization, safe operation, quality, and a balance between supply and demand. This study adopts the determination of the effects of meteorological parameters measured depending on time series as input variables in solar radiation estimation by ANN application. In addition, the study allows examining the effects of the number of hidden layers determined according to the number of input parameters of the ANN model created as an estimation method. The contributions of this study are as follows:

- Developing ANN models to estimation solar radiation while maintaining the measurement time of T, w, P, and H parameters measured at 5-second intervals by the meteorological station located at the selected location, with the measurement time year (Y), month (M), day (D), hour (h), minute (m), second (s) kept constant.
- ANN, which is commonly used in solar radiation estimation, represents the first example of using climate conditions in the study area to estimation solar radiation. This showcases the innovative use of ANN input parameters and model variation to overcome the challenge of estimation solar radiation in the region.
- Developing an ANN model for solar radiation estimation using data obtained from a single geographical location and ensuring its generalizability to other regions.

- Determining the most suitable network architectures in ANN models used to estimation solar radiation for designing or assessing solar energy facilities in regions without meteorological measurement stations.
- This research focuses on Hakkari province. However, an estimation model with higher performance can be created with a larger dataset. Estimation solar radiation varying according to different locations in different regions using the trained network model is valuable for designing PV projects in these regions.
- To develop ANN models using up-to-date data sets collected from on-site metrics.
- To train and validate ANN models for different combinations of input variables to determine the best combination of inputs.
- To verify results using statistical tools.

The reminder sections of this study are organized as follows: section 2 provides details about our research methodology based on the study region, data set, ANN model, method, and performance evaluation criteria. The estimated results are presented, discussed and analyzed in section 3. In the final section, results of this study are given.

2 Methodology

2.1 Study Region and Data Set

Hakkâri province, located in the southeastern region of Turkey, was selected as the place of study, and the location information is given in Fig. 1. The precision pyranometer installed at the specified position (measuring range: 1200 W/m^2, spectral precision: 400-1100 nm) provides radiation data to a workstation at all times of time (day and night). In this workstation, the radiation data is recorded every 5 seconds.

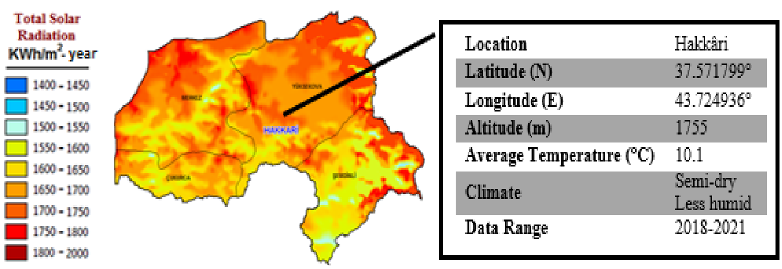

Figure 1
Location and information where the study was conducted [44, 45]

The data set to be used in the study was selected as basic meteorological parameters (T, w, P and H), which can be measured by satellite and ground-based meteorological measuring devices in almost all countries or easily obtained by advanced calculation methods (empirical formulas, etc.). A total of one million independent measurement values were obtained from approximately 7.800.000 sample data obtained between January 1, 2018 and December 31, 2021 for solar radiation estimation. From literature studies, it has been observed that this number of samples is sufficient to train a solar radiation estimation model based on neural networks [46, 47]. In Fig. 2, the multi-data set layout for training, verification and testing of solar radiation estimation model is shown.

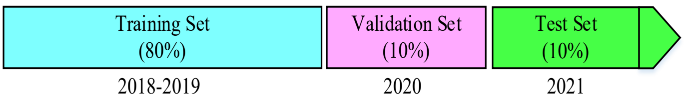

Figure 2
Data set layout for training, verification and testing of solar radiation estimation models

In this study, the dataset is divided into three groups. The dataset will correspond to the years 2018-2019, 2020, and 2021, respectively, and will be set for training, validation, and testing in proportions of 80%, 10%, and 10%, respectively, selected randomly between these years [48-51]. If the validation data poorly represents the entire dataset, it may cause premature termination of training or expose the network to unnecessary training, which is considered a significant issue during the training phase of the network. Therefore, it is essential for the network to maintain dependence on the training dataset but not to be similar to it. If the relationships present in the training dataset are also found in the validation dataset, the evaluation loses its significance. Therefore, during the data preparation stage, the entire dataset is randomly shuffled, and the dataset is divided into training, validation, and testing sections. This practice has resolved many problems such as overfitting, underfitting, and convergence in the model. Randomizing the network prevents similar data from being consecutively introduced into training, thereby preventing consecutive updates of weights in a similar manner and shaping it more robustly. The test data, on the other hand, are entirely different from the training and validation data, containing meteorological data with different temporal characteristics. These data are determined randomly for each training iteration.

2.2 Method

Test set consists of ALMEMO brand data logger and related pyranometer, anemometer, humidity and temperature sensor. Data from the meters is transferred to a workstation (Intel Xeon Silver 4114 CPU @ 2.20GHz, 32.0 GB RAM) along with date and time values and used as input and output data on the ANN model created in a MATLAB environment. The general architectural framework of the model created for this study is shown in Fig. 3.

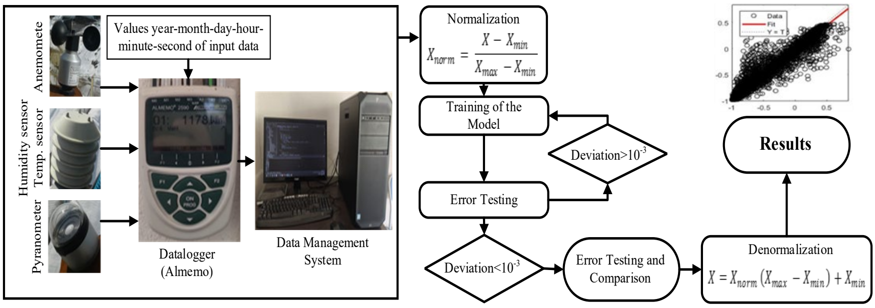

Figure 3
The flowchart of the system used in the study

The system shown in Figure 3, four input parameters, namely T, H, P and w are measured and recorded based on the time of measurement. The output of the system is the solar radiation, which is the data targeted for estimation. The first step in developing the ANN model involves defining the input and output parameters. The input parameters are determined by combining other parameters one by one while keeping the time parameters at which meteorological parameters are obtained constant, as described in the data collection process. However, due to the large size of the dataset, extreme values may disrupt the integrity of the data. Therefore, normalization is applied to make the distribution of values regular (Equation 2). In this study, the data is normalized within the range of -1 to +1. The neural network is trained with the provided data with a specified deviation rate (10^{-3}). Since the coefficient determined for training affects the learning rate in each backpropagation, the influence value of each entered data is propagated through the loops. As each input value cannot entirely influence the network based on its own knowledge, the update status of the network's weights is spread across the entire dataset. Thus, instead of carrying the state of only the most recent data entered or the data, the network carries the information of the entire dataset. Verification of the neural network is done with test data, and the validation dataset has no effect on the weights. The number of neurons in the hidden layer of the neural network (n, 2n, n^2) is adjusted in different ways depending on the number of input parameters, and the network undergoes multiple training iterations to improve the obtained results (Equation 1). The results obtained are subjected to various statistical error assessments. Finally, the normalized data is denormalized and rearranged to obtain the results (Equation 3).

If the validation data represent a low-level of the entire dataset, the training may end prematurely or be subjected to unnecessary training. This is considered a significant problem during the network's training phase. Therefore, it is necessary for the validation dataset to be related to the training dataset but not identical. If the relationships present in the training dataset are also found in the validation dataset, the evaluation loses its significance. Therefore, in the data preparation stage, the total dataset is randomly shuffled and divided into training, validation, and testing sections. This practice has addressed many problems such as memorization, inadequate learning, and resemblance in the model. Random adjustment of the network has prevented consecutive similar data from being trained, thus preventing weights from being updated similarly and shaping the network more robustly.

2.2.1 Artificial Neural Networks (ANN)

In general, ANN models use data-driven, self-adaptive methods and can be used to perform nonlinear modeling without prior knowledge of the relationship between input and output variables. In fact, it is a widely used type of nonlinear model that can solve a wide range of problems and is considered universal estimators. Initially, ANNs consisted of an input layer, at least one hidden layer and an output layer, which created a simple but competitive algorithm [51]. Over time, models with different functions have been tried especially for the hidden layer, and the success performance of the results in the output layer situation has been tried to be increased. Since approximately 80% of the methods used in meteorological estimation in recent years are based on neural networks, a multi-layered neural network approach was applied in this study to estimation solar radiation [52]. The features of the ANN model used in the study are given in Table 2.

Table 2

Feature table of the ANN model tracked in the study

Network Type	**Multi-Layer Back-Propagation Network**
Training Function	Trainlm (Levenberg–Marquardt)
Input Transfer Function	Logsigmoid (Logsig)
Output Transfer Function	Purelin
Performance Function	Mean Square Error (MSE)
Number of Inputs	Variable ranging from 1 to 10
Number of Hidden Layers	1
Number of Hidden Neuron	n, 2n, n^2 n: number of inputs
Learning rate	0.001

ANN architecture usually consists of an input layer, hidden layers, an output layer, link weights and biases, an activation function, and an aggregation node. In this study, three-layer feedback ANN architecture was used including input layer, hidden layer, and output layer. The Levenberg-Marquardt (trainlm) back-propagation algorithm, widely used in ANNs, was preferred as a learning algorithm due to the speed and stability it provided [53, 54]. The ANN model was trained using the Levenberg-Marquardt (LM) backpropagation algorithm. The LM algorithm is a modified version of the Newton method, which exhibits superior performance when applied to time series and transient series. This algorithm provides a balanced solution between the speed of the Newton method and the exact convergence of the steepest descent algorithm [32]. The activation function is a logsigmoid transfer function (logsig) with output between -1 and 1 and specified in Equation 1 [55, 56].

$$Logsig = \frac{1}{1 + e^{-x}} \tag{1}$$

The data measured by the weather station have different dimensions that can have a negative impact on the estimation accuracy of the ANN. Therefore, a data normalization operation must be performed on each type of argument [59]. All data collected from the weather station for both input and output datasets were normalized in the -1 and 1 range using Equation 2 to scale different variables at a similar range [58, 59].

$$X_{normalize} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2}$$

The normalized data is then subjected to denormalization with the help of Equation 3, converted to their actual values and made ready for comparison [10].

$$X = X_{normalize}(X_{max} - X_{min}) + X_{min} \tag{3}$$

In the study, ten parameters including Y, M, D, h, m, s as well as T, P, w and H were used as inputs, and instantaneous solar radiation was estimated as output. The best results were obtained by changing the number of neurons in the intermediate layer depending on the number of input data and executing various network combinations (n, 2n, n^2) over the same data sets. Fig. 4 shows the input parameters of the ANN model developed for solar radiation estimation and the hidden and output layer properties.

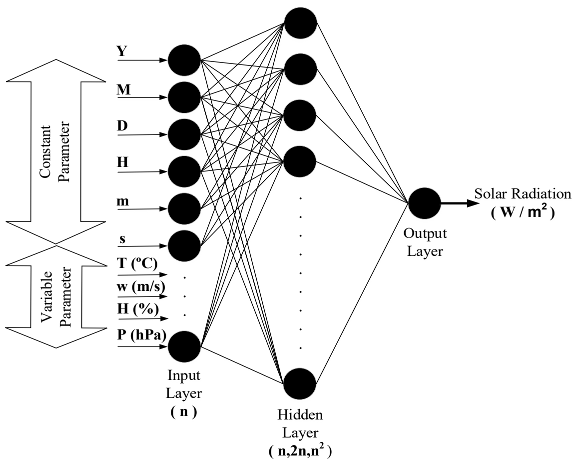

Figure 4
ANN architecture showing input, hidden, and output layer features

For a three-layer network model, the number of neurons in the hidden layer can be calculated using Equation 4. By decreasing or increasing the number of neurons found by certain values (for example, ±5), changes in estimation error can be observed, and the optimal number of neurons can be determined [60, 61]. As a result, lower estimation results can be obtained with network structures created using a very low number of neurons in the hidden layer. Therefore, the number of neurons in the hidden layer (n, 2n, n^2) was determined depending on the number of parameters in the input layer, and the results were discussed in Section 3.

$$H_n = \frac{I_n + O_n}{2} + \sqrt{S_n} \quad (4)$$

Here, H_n refers to the number of neurons in the hidden layer; S_n refers to the number of data samples; I_n and O_n refers to the number of input and output parameters.

In this study, four main conditions were considered in model development using ANN. ANN models that are created to have different hidden layers depending on the input combination are shown in Table 3. Input data such as T, w, P and H, measured depending on the time series, were used to train ANN in a way that creates different possible combinations. The target for ANN is an estimate of instantaneous solar radiation data. Eight hundred thousand data (80%) out of one million data were used to train the ANN, and two hundred thousand data (10%+10%) were used to verify and test the trained network, and the best ANN model was determined with the help of statistical tools. In addition, as a result of 10 iterations of each training, it was ensured that the network structure was created according to the most successful average result. Thus, the effect of the chance factor caused by the randomly selected data set was tried to be eliminated.

Table 3

ANN models with different input combinations

Number of input data	**Combination of Inputs**		**Network combination (n,2n,n^2)**
	Constant	**Variable**	
7-input model (n=7)	Y,M,D,h,m,s	P-H-w-T	7 – 7 – 1 7 – 14 – 1 7 – 49 – 1
8-input model (n=8)	Y,M,D,h,m,s	P,H-P,T-P,w - H,T-H,w-T,w	8 – 8 – 1 8 – 16 – 1 8 – 64 – 1
9-input model (n=9)	Y,M,D,h,m,s	P,H,T-P,H,w-P,w,T-H,w,T	9 – 9 – 1 9 – 18 – 1 9 – 81 – 1
10-input model (n=10)	Y,M,D,h,m,s	P,H,w,T	10 – 10 – 1 10 – 20 – 1 10 –100 – 1

2.3 Statistical Methods for Model Validation

The most important criterion in evaluation of performance success of estimation methods is the accuracy of the estimation method. Therefore, the results of each model are compared and verified using commonly used statistical parameters.

The equations of the following statistical methods have been used to compare the performance success of the estimation models employed in this study [62-64].

$$MSE = \frac{1}{N}\sum_{a=1}^{N}(y_a - \hat{y}_a)^2 \tag{5}$$

$$rRMSE = \frac{\sqrt{\frac{1}{N}\sum_{a=1}^{N}(y_a - \hat{y}_a)^2}}{\overline{y_a}} * 100 \tag{6}$$

$$MBE = \frac{1}{N}\sum_{a=1}^{N}(y_a - \hat{y}_a) \tag{7}$$

$$MAPE = \frac{1}{N}\sum_{a=1}^{N}\frac{(y_a - \hat{y}_a)}{y_a} * 100 \tag{8}$$

$$R^2 = 1 - \frac{\sum_{a=1}^{N}(y_a - \hat{y}_a)^2}{\sum_{a=1}^{N}(y_a - \overline{y_a})^2} \tag{9}$$

In equations, N refers to the number of data; y_a refers to actual measurement value; ($\hat{y}_a$) refers to estimation value; ($\overline{y}_a$) refers to the average of solar radiation values.

3 Results and Discussion

In this study, in addition to time parameters (Y, M, D, h, m, s), meteorological parameters (P, H, w, and T) have also been used as input data. A different ANN model has been developed for each combination of work created depending on the number of inputs. Additionally, in each ANN model, the estimation success of the number of neurons in the hidden layer has been investigated according to the number of input parameters. The target for all models is instantaneous solar radiation estimation. Each model was trained using the Levenberg-Marquardt [LM] algorithm. Training can be initiated with a minimum number of neurons in the hidden layer of the ANN architecture. The number of these neurons is changed to n, 2n and n^2 depending on the number of input parameters. The retraining of the ANN is continued until the best approach is reached. The main criterion for choosing among the ANN model architectures is MSE [65]. The estimation ability of the developed ANN models is tested with new data sets, and this data is not used during the model development phase.

The variation of the measurement data used in model development is shown in Fig. 5. The success of ANN is that it has the ability to find a relationship between the input parameters and the output parameter.

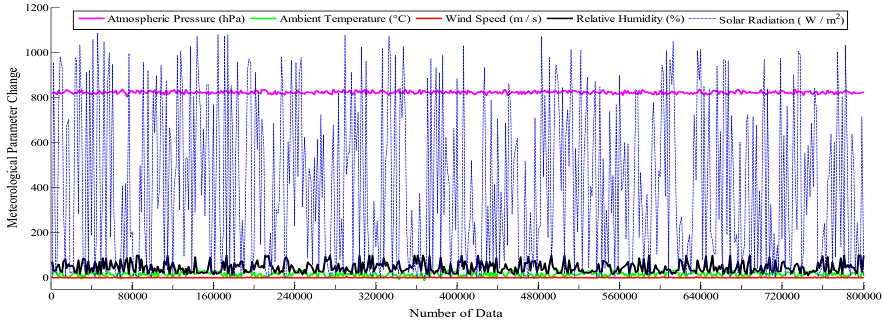

Figure 5

During the training period of the ANN; P, H, T, w and the change of solar radiation

Each input parameter is associated with instantaneous solar radiation by developing a different ANN model under the conditions specified in Table 2.

3.1 7-input ANN Model

Keeping the time parameters constant, each of the 4 entries selected was applied to the ANN separately under the conditions specified in Table 2. The number of hidden layers of the ANN was changed depending on the number of input parameters. The effect of each input on solar radiation estimation was determined. The network's best performance was achieved by training the developed ANN architecture multiple times until the MSE showed a minimum value. The same trained network was tested with new datasets to verify the network's performance. Statistical performance the ANN model created are given in Table 4.

Table 4

Statistical performance of 7-input ANN

Network combination	ANN Inputs	MSE	rRMSE (%)	R^2	MBE	MAPE (%)
7 – 7 – 1	Y,M,D,h,m,s,**P**	5.45	2.70	0.998	1.40	15.66
	Y,M,D,h,m,s,H	**4.18**	**2.37**	**0.998**	**0.80**	**9.84**
	Y,M,D,h,m,s,**T**	5.31	2.70	0.990	1.20	14.20
	Y,M,D,h,m,s,**w**	5.93	2.82	0.987	1.60	14.71
7 – 14 – 1	Y,M,D,h,m,s,**P**	5.04	2.60	0.990	0.60	**7.49**
	Y,M,D,h,m,s,H	**3.92**	**2.29**	**0.990**	1.10	12.47
	Y,M,D,h,m,s,**T**	4.61	2.49	0.983	1.30	15.31
	Y,M,D,h,m,s,**w**	5.43	2.70	0.990	1.20	14.23
7 – 49 – 1	Y,M,D,h,m,s,**P**	3.99	2.31	0.990	0.70	7.72
	Y,M,D,h,m,s,H	**3.36**	**2.12**	**0.991**	0.80	9.46
	Y,M,D,h,m,s,**T**	3.78	2.25	0.999	0.60	6.99
	Y,M,D,h,m,s,**w**	4.53	2.46%	0.987	**0.50**	**5.69**

According to the statistical results in Table 4, the neural network model created as the number of neurons in the hidden layer n^2 showed a high estimation accuracy with values of 5.69% MAPE and 98,7% R^2. At the same time, the results of all statistical indicators in this network structure are included in the successful estimation range in Section 2.3. Of all the parameters, the ANN model trained with H showed the least error value, indicating a strong relationship between H and instantaneous solar radiation at that location. In addition, T has a strong influence on the instantaneous solar radiation estimation. In both cases, R^2 shows values close to 1, and MBE shows values close to 0, indicating the accuracy of the estimation success. Figure 6 shows the comparison between the performance of the ANN model, where the number of neurons in the hidden layer is adjusted as n^2 with 7 input parameters, and the evaluation of actual and radiation values.

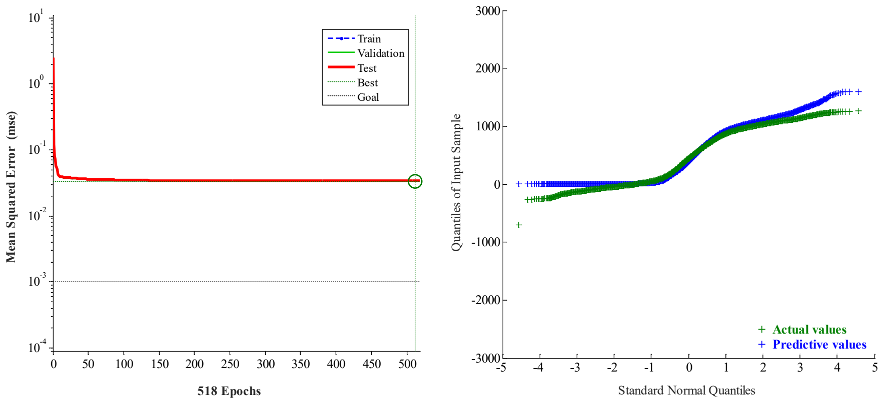

Figure 6

Model performance and comparison graph (Network combination: 7 – 49 –1 / Y,M,D,h,m,s,H)

3.2 8-input ANN Model

Two meteorological parameter inputs together with time parameters are applied to the ANN at the same time. Six such combinations were created for each ANN model, and the effect of each combination of inputs on instantaneous solar radiation was determined. Statistical performance indicators of the ANN model are given in Table 5.

Table 5

Statistical performance of 8-input ANN

Network combination	**ANN Inputs**	**MSE**	**rRMSE (%)**	**R^2**	**MBE**	**MAPE (%)**
8 – 8 – 1	**Y,M,D,h,m,s,P,H**	**4.00**	**2.31**	**0.989**	**0.40**	**4.16**
	Y,M,D,h,m,s,**P,T**	4.82	2.54	0.987	1.20	14.13
	Y,M,D,h,m,s,**H,T**	4.04	2.33	0.989	1.00	11.26
	Y,M,D,h,m,s,**P,w**	5.11	2.62	0.986	0.90	10.63

	Y,M,D,h,m,s,**w,H**	4.09	2.34	0.989	1.00	11.02
	Y,M,D,h,m,s,**T,w**	5.08	2.61	0.986	1.70	19.42
8 – 16 – 1	Y,M,D,h,m,s,**P,H**	3.75	2.24	0.990	**0.70**	**7.81**
	Y,M,D,h,m,s,**P,T**	4.39	2.43	0.988	0.80	9.70
	Y,M,D,h,m,s,H,T	**3.71**	**2.23**	**0.990**	0.90	10.41
	Y,M,D,h,m,s,**P,w**	4.81	2.54	0.987	1.10	13.11
	Y,M,D,h,m,s,**w,H**	3.81	2.26	0.989	1.30	14.83
	Y,M,D,h,m,s,**T,w**	4.38	2.42	0.988	1.30	14.60
8 – 64 – 1	**Y,M,D,h,m,s,P,H**	**3.03**	**2.02**	**0.991**	**0.40**	**4.06**
	Y,M,D,h,m,s,**P,T**	3.20	2.07	0.991	0.60	7.05
	Y,M,D,h,m,s,**H,T**	3.05	2.02	0.991	0.90	10.54
	Y,M,D,h,m,s,**P,w**	3.71	2.23	0.991	0.90	10.82
	Y,M,D,h,m,s,**w,H**	3.19	2.07	0.991	0.50	5.51
	Y,M,D,h,m,s,**T,w**	3.33	2.11	0.991	0.90	10.40

Trained on the combination of P,H from all the combinations generated, the ANN produced the fewest errors in estimation instantaneous solar radiation. The H,T combination was also successfully estimated with a MBE close to 0, an rRMSE-MAPE of <10%, and an R^2 value close to 1. Combinations in which the w parameter was included gave the highest error values. In the ANN model, which was created as the number of neurons n in the hidden layer n^2, more successful results were obtained for all statistical indicators. Figure 7 shows the comparison between the performance of the ANN model, where the number of neurons in the hidden layer is adjusted as n^2 with 8 input parameters, and the evaluation of actual and radiation values.

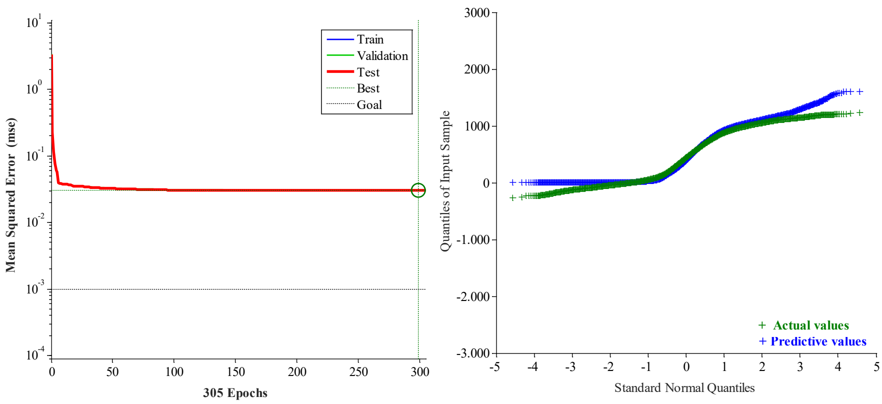

Figure 7

Model performance and comparison graph (Network combination: 8 – 64 –1 / Y,M,D,h,m,s,P,H)

3.3 9-input ANN Model

A total of four combinations with three input variables were created in addition to the time parameters to train and validate the ANN. Statistical performance of this ANN model is given in Table 6.

Table 6
Statistical performance of 9-input ANN

Network combination	ANN Inputs	MSE	rRMSE (%)	R^2	MBE	MAPE (%)
9 – 9 – 1	Y,M,D,h,m,s,**P,H,T**	3.94	2.30	0.989	**0.70**	**8.11**
	Y,M,D,h,m,s,P,H,w	**3.83**	**2.27**	**0.989**	0.80	9.78
	Y,M,D,h,m,s,**P,w,T**	4.59	2.48	0.978	1.30	14.54
	Y,M,D,h,m,s,**H,w,T**	3.97	2.31	0.988	1.20	13.54
9 – 18 – 1	Y,M,D,h,m,s,**P,H,T**	3.56	2.19	0.990	0.90	10.63
	Y,M,D,h,m,s,**P,H,w**	3.68	2.22	0.989	0.90	9.97
	Y,M,D,h,m,s,**P,w,T**	4.14	2.36	0.988	1.10	12.51
	Y,M,D,h,m,s,H,w,T	**3.52**	**2.19**	**0.990**	**0.50**	**5.62**
9 – 81 – 1	**Y,M,D,h,m,s,P,H,T**	**2.71**	**1.91**	**0.992**	**0.20**	**1.99**
	Y,M,D,h,m,s,**P,H,w**	2.89	1.97	0.990	0.40	4.70
	Y,M,D,h,m,s,**P,w,T**	2.92	1.98	0.992	1.10	12.26
	Y,M,D,h,m,s,**H,w,T**	2.86	1.96	0.992	0.40	4.54

In all three network combinations in Table 6, the ANN models including the H parameter showed good accuracy in estimation instantaneous solar radiation. All combinations of parameters P,w,T yielded high error values. It has been observed that combinations of H in all network structures show the best estimate. In particular, in the combination of 9-81-1 ANN formed as n^2 in the number of neurons in the hidden layer, results very close to the ideal values of the statistical methods specified in Section 2.3 were obtained. Figure 8 shows the comparison between the performance of the ANN model, where the number of neurons in the hidden layer is adjusted as n^2 with 9 input parameters, and the evaluation of actual and radiation values.

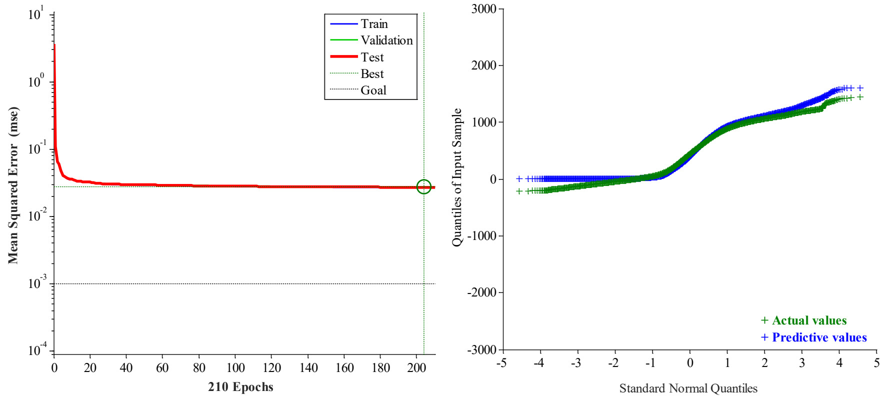

Figure 8

Model performance and comparison graph (Network combination: 8 – 81 –1 / Y,M,D,h,m,s,P,H,T)

3.4 10-input ANN Model

A combination with four input variables were created in addition to the time parameters to train and validate the ANN. Statistical performance of this ANN model is given in Table 7.

Table 7

Statistical performance of 10-input ANN

Network combination	ANN Inputs	MSE	rRMSE (%)	R^2	MBE	MAPE (%)
10 – 10 – 1	Y,M,D,h,m,s,P, H,T,w	3.83	2.27	0.989	1.10	13.21
10 – 20 – 1	Y,M,D,h,m,s,P, H,T,w	3.51	2.17	0.990	0.80	8.96
10 – 100 – 1	**Y,M,D,h,m,s,P, H,T,w**	**2.56**	**1.85**	**0.993**	**0.50**	**6.12**

The network model, in which all parameters were used and created as the square of the number of input parameters of the number of hidden neurons in the network depending on the number of input parameters, showed the best estimate. Although this network structure provides the best R^2 value among other network structures, it performed worse according to MAPE and MBE results than many network structures in sections 3.1, 3.2, and 3.3. However, the combination of 10-100-1 ANNs showed better rRMSE value than other network combinations. In addition, the ANN model, in which all parameters were applied as the ANN input parameter and the number of neurons in the hidden layer was selected as n^2, gave the most successful network structure result with 99.3% R^2. Figure 9 shows the comparison between the performance of the ANN model, where the number of neurons in the hidden layer is adjusted as n^2 with 10 input parameters, and the evaluation of actual and radiation values.

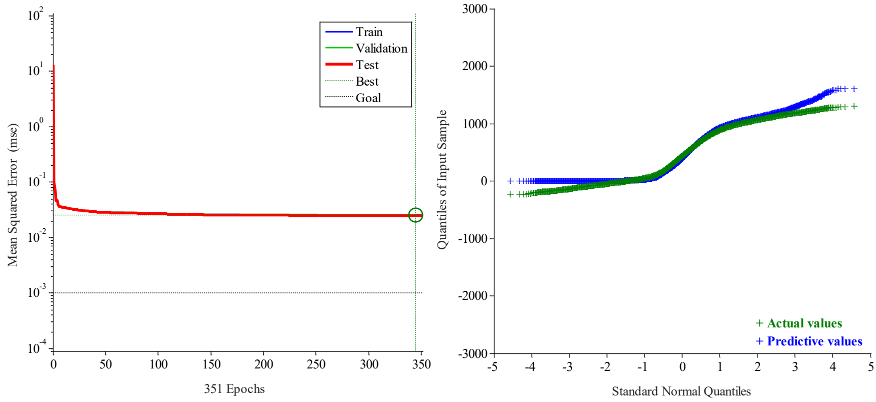

Figure 9

Model performance and comparison graph (Network combination: 9–100–1/ Y,M,D,h,m,s,P,H,T,w)

Considering all the information obtained within the scope of the study and the statistical performance indicators shown in Table 4, Table 5, Table 6, and Table 7, it can be said that the parameter H is quite effective in estimation solar radiation. The reason for this could be the presence of suspended dust particles in the air, which obstruct the solar radiation from reaching the measurement surface, depending on the parameter H. The network models developed with this parameter have generally shown successful results, and the statistical performance indicators have been observed within appropriate ranges. Although the numbers of neurons in the network structure were varied, what matters is that the parameters used as input are related to each other. Increasing the number of neurons in the hidden layer in all four ANN models created based on the number of input parameters has led to an improvement in MBE, MAPE, RMSE, and MSE results. Therefore, it is demonstrated that the ANN models with the number of neurons set as n2 in all four models perform better. Additionally, the effectiveness of the P and T parameters is shown with the network model dependent on the n2 neuron count in Table 6. In this context, the most efficient estimation approach for estimation solar radiation using short-term measured basic meteorological parameters and their obtained time parameters is the ANN model with the Y, M, D, h, m, s, P, H and T parameters. The use of the H parameter along with the P and T parameters implies an improvement in performance. The presented results support the benefits of ANN approaches commonly used in solar radiation estimation and their application to processes with variable parameters.

Conclusions

In this study, the model structures used were evaluated based on the same performance characteristics by examining the effect of fundamental meteorological parameters on solar radiation estimation performance, depending on time parameters. ANN architectures with back-propagation algorithm created depending on the number of different input combinations, and input parameters were developed. When the meteorological parameters were measured, the ANN models, considered as constant inputs and consisted of a combination of w, T, H and P parameters and increased the number of input variables, estimated the instantaneous solar radiation values with high accuracy. Combinations in which the H parameter is included are defined as combinations of highly affected inputs to estimation instantaneous solar radiation data. The best results in all combinations were obtained when the number of neurons in the hidden layer was squared (n^2) of the number of inputs. Solar radiation was estimated within high estimation limits with 2.56 MSE, 1.85% rRMSE and 6.12% MAPE values with the ANN model (10-100-1 network), in which all parameters were taken as inputs and the number of neurons in the hidden layer n^2 was selected. It also showed a very high estimation performance with a value of 99.3% R^2. At the same time, the ANN model (network 8-64-1), in which the parameters of P and H are included and the number of neurons in the hidden layer is formed to be n^2 , showed a high estimation accuracy with a MAPE value of 4.06%. In particular, it was observed that the parameters of w and T alone were not sufficient to accurately estimation solar radiation, but the accuracy increased with the inclusion of P and H parameters in these parameters. In particular, the ANN model (9-81-1 network), which was created by adding T parameter to P and H parameters, yielded a very high solar radiation estimation with 99.2% R^2 as well as 1.99% MAPE and 1.91% rRMSE results within very good estimation limits.

The ANN models developed perform well in estimation solar radiation for the selected location. The parameters of H and T recorded at a given location for any given location had a significant impact on the solar radiation at that location. The results obtained showed that the generated ANN model could be used to accurately estimate solar radiation for locations with geographical features where meteorological stations are not available. At the same time, it was shown that very high solar radiation estimation results (R^2=99.8%) can be obtained with only basic meteorological parameters. These parameters are measured by the meteorological centers of almost all countries, and it has been seen that high estimation results can be obtained without using other parameters such as precipitation amount, aperture index, and wind direction in solar radiation estimation and without larger network structures.

References

[1] Son, N. N., & Vinh, Luu The (2022) Parameter estimation of photovoltaic model, using balancing composite motion optimization. Acta Polytechnica Hungarica, 19(11), 27-46

[2] Ben Othman, A., Belkilani, K., & Besbes, M. (2020) Prediction improvement of potential PV production pattern, imagery satellite-based. Scientific Reports, 10(1), 1-10

[3] Amrouche, B., & Le Pivert, X. (2014) Artificial neural network based daily local forecasting for global solar radiation. Applied energy, 130, 333-341

[4] Nwokolo, S. C., Amadi, S. O., Obiwulu, A. U., Ogbulezie, J. C., & Eyibio, E. E. (2022) Prediction of global solar radiation potential for sustainable and cleaner energy generation using improved Angstrom-Prescott and Gumbel probabilistic models. Cleaner Engineering and Technology, 6, 100416

[5] El Mghouchi, Y. (2022) On the prediction of daily global solar radiation using temperature as input. An application of hybrid machine learners to the six climatic Moroccan zones. Energy Conversion and Management: X, 13, 100157

[6] Amiri, B., Gómez-Orellana, A. M., Gutiérrez, P. A., Dizène, R., Hervás-Martínez, C., & Dahmani, K. (2021) A novel approach for global solar irradiation forecasting on tilted plane using Hybrid Evolutionary Neural Networks. Journal of Cleaner Production, 287, 125577

[7] Li, X., Ma, L., Chen, P., Xu, H., Xing, Q., Yan, J., ... & Cheng, Y. (2022) Probabilistic solar irradiance forecasting based on XGBoost. Energy Reports, 8, 1087-1095

[8] Singla, P., Duhan, M., & Saroha, S. (2022) A Hybrid Solar Irradiance Forecasting Using Full Wavelet Packet Decomposition and Bi-Directional Long Short-Term Memory (BiLSTM) Arabian Journal for Science and Engineering, 1-27

[9] Mishra, Sakshi, and Praveen Palanisamy. "Multi-time-horizon solar forecasting using recurrent neural network." 2018 IEEE Energy Conversion Congress and Exposition (ECCE) IEEE, 2018.8

[10] Premalatha, M., & Naveen, C. (2018) Analysis of different combinations of meteorological parameters in predicting the horizontal global solar radiation with ANN approach: A case study. Renewable and Sustainable Energy Reviews, 91, 248-258

[11] Pozna, C., Precup, R. E., Tar, J. K., Škrjanc, I., & Preitl, S. (2010) New results in modelling derived from Bayesian filtering. Knowledge-Based Systems, 23(2), 182-194

[12] Pozna, C., & Precup, R. E. (2012) Aspects concerning the observation process modelling in the framework of cognition processes. Acta Polytechnica Hungarica, 9(1), 203-223

[13] Hedrea, E. L., Precup, R. E., Roman, R. C., & Petriu, E. M. (2021) Tensor product-based model transformation approach to tower crane systems modeling. Asian Journal of Control, 23(3), 1313-1323

[14] Szedlak-Stinean, A. I., Precup, R. E., Petriu, E. M., Roman, R. C., Hedrea, E. L., & Bojan-Dragos, C. A. (2022) Extended Kalman filter and Takagi-Sugeno fuzzy observer for a strip winding system. Expert Systems with Applications, 208, 118215

[15] Duka, G., & Travin, S. (2022) Processing, neural network-based modeling of biomonitoring studies data and validation on Republic of Moldova data. Proceedings of the Romanian Academy Series A-Mathematics Physics Technical Sciences Information Science, 403-410

[16] Zhou, Y., Liu, Y., Wang, D., Liu, X., & Wang, Y. (2021) A review on global solar radiation prediction with machine learning models in a comprehensive perspective. Energy conversion and management, 235, 113960

[17] Premalatha, N., & Valan Arasu, A. (2016) Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. Journal of applied research and technology, 14(3), 206-214

[18] Wang, L., Kisi, O., Zounemat-Kermani, M., Salazar, G. A., Zhu, Z., & Gong, W. (2016) Solar radiation prediction using different techniques: model evaluation and comparison. Renewable and Sustainable Energy Reviews, 61, 384-397

[19] Al-Hajj, R., & Assi, A. (2017) Estimating solar irradiance using genetic programming technique and meteorological records. AIMS Energy, 5(5), 798-813

[20] Khosravi, A., Nunes, R. O., Assad, M. E. H., & Machado, L. (2018) Comparison of artificial intelligence methods in estimation of daily global solar radiation. Journal of cleaner production, 194, 342-358

[21] Gouda, S. G., Hussein, Z., Luo, S., & Yuan, Q. (2019) Model selection for accurate daily global solar radiation prediction in China. Journal of cleaner production, 221, 132-144

[22] Antonopoulos, V. Z., Papamichail, D. M., Aschonitis, V. G., & Antonopoulos, A. V. (2019) Solar radiation estimation methods using ANN and empirical models. Computers and Electronics in Agriculture, 160, 160-167

[23] Fan, J., Wu, L., Zhang, F., Cai, H., Ma, X., & Bai, H. (2019) 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, 105, 168-186

[24] El Mghouchi, Y., Chham, E., Zemmouri, E. M., & El Bouardi, A. (2019) Assessment of different combinations of meteorological parameters for predicting daily global solar radiation using artificial neural networks. Building and Environment, 149, 607-622

[25] Marzouq, M., Bounoua, Z., El Fadili, H., Mechaqrane, A., Zenkouar, K., & Lakhliai, Z. (2019) New daily global solar irradiation estimation model based on automatic selection of input parameters using evolutionary artificial neural networks. Journal of Cleaner Production, 209, 1105-1118

[26] Zang, H., Cheng, L., Ding, T., Cheung, K. W., Wang, M., Wei, Z., & Sun, G. (2020) Application of functional deep belief network for estimating daily global solar radiation: A case study in China. Energy, 191, 116502

[27] Narvaez, G., Giraldo, L. F., Bressan, M., & Pantoja, A. (2021) Machine learning for site-adaptation and solar radiation forecasting. Renewable Energy, 167, 333-342

[28] Rahman, S., Rahman, S., & Haque, A. B. (2021, February) Prediction of solar radiation using artificial neural network. In Journal of Physics: Conference Series (Vol. 1767, No. 1, p. 012041) IOP Publishing

[29] Yakoubi, H., Mghouchi, Y. E., Abdou, N., Hajou, A., & Khellouki, A. (2021) Correlating clearness index with cloud cover and other meteorological parameters for forecasting the global solar radiation over Morocco. Optik, 242, 167145

[30] Djeldjli, H., Benatiallah, D., Tanougast, C., & Benatiallah, A. (2024) Solar radiation forecasting based on ANN, SVM and a novel hybrid FFA-ANN model: A case study of six cities south of Algeria. AIMS Energy, 12(1)

[31] Mfetoum, I. M., Ngoh, S. K., Molu, R. J. J., Nde Kenfack, B. F., Onguene, R., Naoussi, S. R. D., ... & Berhanu, M. (2024) A multilayer perceptron neural network approach for optimizing solar irradiance forecasting in Central Africa with meteorological insights. Scientific Reports, 14(1), 3572

[32] Zaim, S., El Ibrahimi, M., Arbaoui, A., Samaouali, A., Tlemcani, M., & Barhdadi, A. (2023) Using artificial intelligence for global solar radiation modeling from meteorological variables. Renewable Energy, 215, 118904

[33] Mfetoum, I. M., Ngoh, S. K., Molu, R. J. J., Nde Kenfack, B. F., Onguene, R., Naoussi, S. R. D., ... & Berhanu, M. (2024) A multilayer perceptron neural network approach for optimizing solar irradiance forecasting in Central Africa with meteorological insights. Scientific Reports, 14(1), 3572

[34] Lu, Y., Wang, L., Zhu, C., Zou, L., Zhang, M., Feng, L., & Cao, Q. (2023) Predicting surface solar radiation using a hybrid radiative Transfer–Machine learning model. Renewable and Sustainable Energy Reviews, 173, 113105

[35] Reddy, K. S., & Ranjan, M. (2003) Solar resource estimation using artificial neural networks and comparison with other correlation models. Energy conversion and management, 44(15), 2519-2530

[36] Mubiru, J., & Banda, E. J. K. B. (2008) Estimation of monthly average daily global solar irradiation using artificial neural networks. Solar energy, 82(2), 181-187

[37] Yadav, A. K., & Chandel, S. S. (2014) Solar radiation prediction using Artificial Neural Network techniques: A review. Renewable and sustainable energy reviews, 33, 772-781

[38] Kumar, S., & Kaur, T. (2016) Development of ANN based model for solar potential assessment using various meteorological parameters. Energy Procedia, 90, 587-592

[39] Yıldırım, H. B., Çelik, Ö., Teke, A., & Barutçu, B. (2018) Estimating daily Global solar radiation with graphical user interface in Eastern Mediterranean region of Turkey. Renewable and Sustainable Energy Reviews, 82, 1528-1537

[40] Karaman, Ö. A., Ağır, T. T., & Arsel, İ. (2021) Estimation of solar radiation using modern methods. Alexandria Engineering Journal, 60(2), 2447-2455

[41] Guermoui, M., Abdelaziz, R., Gairaa, K., Djemoui, L., & Benkaciali, S. (2022) New temperature-based predicting model for global solar radiation using support vector regression. International Journal of Ambient Energy, 43(1), 1397-1407

[42] Gairaa, K., Voyant, C., Notton, G., Benkaciali, S., & Guermoui, M. (2022) Contribution of ordinal variables to short-term global solar irradiation forecasting for sites with low variabilities. Renewable Energy, 183, 890-902

[43] Arora, I., Gambhir, J., & Kaur, T. (2021) Data normalisation-based solar irradiance forecasting using artificial neural networks. Arabian Journal for Science and Engineering, 46(2), 1333-1343

[44] Solar Energy Potential Atlas (GEPA) July 20. 2023 [Online] https://gepa.enerji.gov.tr/MyCalculator/

[45] Turkish State Meteorological Service (TSMS) July 20. 2023 [Online] https://mgm. gov.tr/eng/forecast-cities.aspx

[46] Qing, X., & Niu, Y. (2018) Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, 148, 461-468

[47] Jiang, H., Lu, N., Qin, J., Tang, W., & Yao, L. (2019) A deep learning algorithm to estimate hourly global solar radiation from geostationary satellite data. Renewable and Sustainable Energy Reviews,114,109327

[48] Halabi, L. M., Mekhilef, S., & Hossain, M. (2018) Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation. Applied energy, 213, 247-261

[49] Pereira, S., Canhoto, P., Salgado, R., & Costa, M. J. (2019) Development of an ANN based corrective algorithm of the operational ECMWF global horizontal irradiation forecasts. Solar Energy, 185, 387-405

[50] Kumari, P., & Toshniwal, D. (2021) Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance. Journal of Cleaner Production, 279, 123285

[51] Şahin, M. (2019) Determining optimum tilt angles of photovoltaic panels by using artificial neural networks in turkey. Tehnički vjesnik-Technical Gazette, 26(3), 596-602

[52] Narvaez, G., Giraldo, L. F., Bressan, M., & Pantoja, A. (2021) Machine learning for site-adaptation and solar radiation forecasting. Renewable Energy, 167, 333-342. https://doi.org/10.1016/j.renene.2020.11.089

[53] Heng, S. Y., Ridwan, W. M., Kumar, P., Ahmed, A. N., Fai, C. M., Birima, A. H., & El-Shafie, A. (2022) Artificial neural network model with different backpropagation algorithms and meteorological data for solar radiation prediction. Scientific reports, 12(1), 1-18

[54] Pang, Z., Niu, F., & O'Neill, Z. (2020) Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons. Renewable Energy, 156, 279-289

[55] Vakili, M., Sabbagh-Yazdi, S. R., Khosrojerdi, S., & Kalhor, K. (2017) Evaluating the effect of particulate matter pollution on estimation of daily global solar radiation using artificial neural network modeling based on meteorological data. Journal of cleaner production, 141, 1275-1285

[56] Zang, H., Cheng, L., Ding, T., Cheung, K. W., Wang, M., Wei, Z., & Sun, G. (2020) Application of functional deep belief network for estimating daily global solar radiation: A case study in China. Energy, 191, 116502. https://doi.org/10.1016/j.energy.2019.116502

[57] Wang, H., Cai, R., Zhou, B., Aziz, S., Qin, B., Voropai, N., ... & Barakhtenko, E. (2020) Solar irradiance forecasting based on direct explainable neural network. Energy Conversion and Management,226,113487

[58] Ghimire, S., Deo, R. C., Downs, N. J., & Raj, N. (2019) Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia. Journal of cleaner production,216,288-310

[59] Guher, A. B., & Tasdemir, S. (2021) Determining of solar power by using machine learning methods in a specified region. Tehnički vjesnik-Technical Gazette, 28(5), 1471-1479

[60] Yadav, A. K., Malik, H., & Chandel, S. S. (2014) Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models. Renewable and Sustainable Energy Reviews, 31, 509-519

[61] Chow, S. K., Lee, E. W., & Li, D. H. (2012) Short-term prediction of photovoltaic energy generation by intelligent approach. Energy and Buildings, 55, 660-667

[62] Ghimire, S., Deo, R. C., Casillas-Pérez, D., & Salcedo-Sanz, S. (2022) Boosting solar radiation predictions with global climate models, observational predictors and hybrid deep-machine learning algorithms. Applied Energy, 316, 119063

[63] Tuncer, A. D., Sözen, A., Afshari, F., Khanlari, A., Şirin, C., & Gungor, A. (2020) Testing of a novel convex-type solar absorber drying chamber in dehumidification process of municipal sewage sludge. Journal of Cleaner Production, 272, 122862

[64] Tirmikci, C. A., & Yavuz, C. (2018) Establishing new regression equations for obtaining the diffuse solar radiation in Sakarya (Turkey). Tehnički vjesnik-Technical Gazette, 25(Supplement 2), 503-508

[65] Yang, D. (2022) Correlogram, predictability error growth, and bounds of mean square error of solar irradiance forecasts. Renewable and Sustainable Energy Reviews, 167, 112736