Electrical Engineering (2024) 106:3671–3689
https://doi.org/10.1007/s00202-023-02067-z

ORIGINAL PAPER

Instant solar irradiation forecasting for solar power plants using different ANN algorithms and network models

Ersan Omer Yuzer[1] · **Altuğ Bozkurt**[2]

Received: 24 August 2022 / Accepted: 30 September 2023 / Published online: 1 November 2023

Abstract
Solar irradiation is a crucial parameter in the design and operation of solar energy systems. However, its long-term measurement everywhere is hindered by the maintenance and cost of measurement devices. Therefore, numerous research studies have been conducted to determine solar irradiation, leading to the development of various prediction models. Recently, artificial neural network (ANN) models have been shown to enable researchers to make more accurate predictions. This study aims to identify the most effective algorithms and functions for accurately predicting instantaneous solar irradiation using ANN models with different network structures. Five commonly used training algorithms and two different ANN architectures are examined in this study. These models are tested with various transfer functions, and the impact of the number of neurons in the hidden layer on prediction results is also investigated. Meteorological data collected at 5-s intervals from a meteorology station in Hakkâri Province between 2019 and 2021, totaling one million data points, are used for model training. The ANN model with a network structure consisting of 100 neurons, trained with the Levenberg–Marquardt algorithm and "tansig" transfer function, achieved the best prediction performance with a correlation coefficient (R) of 0.9783 and a mean absolute percentage error of 6.79%. For an 80-10-10 data split, the mean-squared error, normalized root-mean-squared error, and mean bias error were found to be 0.024, 7.206, and 0.800, respectively. The solar irradiation prediction performance varied based on the training algorithm and particularly the transfer functions used. Similar approaches can be employed in regions where measurement devices cannot be installed, enabling successful prediction results even without direct irradiation measurements.

Keywords Algorithms · Artificial neural networks · Forecasting · Solar energy · Solar irradiation

List of symbols

ANFIS	Adaptive neural fuzzy inference system
ANN	Artificial neural network
AR	Autoregressive process
ARIMA	Autoregressive integrated moving average process
ARMA	Autoregressive moving average
B	Cloud cover
CNN	Convolutional neural network
d	Day length
D	Day
DL	Deep learning
DT	Minimum and maximum temperature difference
ELM	Extreme learning machine
GA	Genetic algorithm
GBT	Gradient boosting trees
H	Relative humidity
H_0	Extra-atmospheric solar irradiation
Kt	Openness index
L_a	Latitude
L_o	Longitude
LM	Levenberg–Marquardt
LSTM	Long-short-term memory
M	Month
MA	Moving average process
MAPE	Mean absolute percentage error
MBE	Mean biased error
MLP	Multi-layer perceptron
MSE	Mean square error
NN	Neural network

✉ Ersan Omer Yuzer
ersanomeryuzer@hakkari.edu.tr

[1] Çölemerik V.H.S., Department of Electricity and Energy, Hakkari University, Hakkari, Turkey

[2] Faculty of Electrical and Electronics, Department of Electrical Engineering, Yildiz Technical University, Istanbul, Turkey

nRMSE	Normalized root-mean-square error
NWP	Numerical weather prediction
PV	Photovoltaic
r	Rain
R	Correlation coefficient
R^2	Coefficient of determination
RF	Random forest
RMSE	Root-mean-square error
S	Sunbathing time
SVM	Support vector machine
T_{avg}	Average temperature
T_{max}	Maximum temperature
T_{min}	Minimum temperature
w	Wind speed
WT	Wavelet transform
Y	Year
Z	Altitude

1 Introduction

Today, renewable energy sources have become more important due to climate change affecting the world entirely. The fact that solar energy is a clean, renewable, and usable energy source has led all the countries of the world to produce energy from this source due to the global climate crisis of the world. Therefore, to increase the efficiency of solar energy systems and to evaluate their applicability, engineering designs and scientific studies need up-to-date data on the amount of solar irradiation which could not be easily measured in some regions [1, 2]. These data are not always available due to the high cost of solar irradiation measurement sensors and require constant maintenance and calibration [3–6].

Reliable solar irradiation information is essential for the design and development of solar energy systems [7]. Additionally, knowing the solar irradiation is crucial for determining the most suitable location for installing a photovoltaic (PV) system [8]. Therefore, accurate knowledge of solar irradiation plays a significant role in ensuring the security of the power grid and effectively storing reserves of backup energy sources [9, 10].

Accurate irradiation forecasting assists in grid planning and improves power quality [11, 12]. However, the non-stationary behavior and variability of solar irradiation make this task quite challenging [13]. To overcome this challenge, solar irradiation prediction models are required, and many models have been developed. Currently, the methods used for solar irradiation prediction can be categorized as shown in Fig. 1, with the selection of these methods primarily dependent on the prediction time horizon [14].

When examining the prediction methods in Fig. 1, it is known that hybrid methods demonstrate better performance in terms of prediction accuracy compared to basic methods [16]. However, they come with a computational complexity that requires a longer time to reach a result. Additionally, their performance is highly dependent on carefully selected historical input data [17]. Therefore, this study focuses on investigating the effectiveness of network architecture, data set distribution, algorithms, and functions in ANN models using commonly used meteorological parameters for solar irradiation prediction. The best results obtained from the algorithms and functions used in the ANN model for solar irradiation prediction are tested with different data distributions.

ANN is widely used for predicting different solar irradiation components [13, 18–21]. Recent studies have confirmed that both experimental methods and models are acceptable approaches for solar irradiation prediction. However, the use of methods like ANN modeling has been shown to enable researchers to achieve higher accuracy in their predictions [18, 22].

ANNs are commonly used in models aimed at predicting solar irradiation in different time intervals and locations around the world. Meteorological parameters such as sunshine duration (S), relative humidity (H), temperature values (T_{max}, $T_{min,}$ and T_{avg}), extraterrestrial solar irradiation (H_0), latitude (L_o), longitude (L_a), altitude (Z), wind speed (w), and wind direction are among the most frequently used data for solar irradiation prediction [23–25].

Gutierrez-Corea et al. propose an ANN approach for short-term prediction of solar irradiation in Spain, using nine input sequences consisting of ten values from ten different locations in Spain. They use the "logsig" activation function. The results obtained show errors ranging from 22.6 to 32.1% for forecast horizons of 1 to 6 h [26]. Similarly, Bosch et al. indicate that in their solar irradiation prediction study using data obtained from 12 different stations in Spain, ANN is an effective and straightforward methodology for calculating solar irradiation levels over complex mountainous terrains using data from a single radiometric station [27].

Mellit and Pavan proposed an ANN model using daily solar irradiation and air temperature data obtained in Italy. They used the LM training algorithm and reported that the model showed good prediction performance with a minimum correlation coefficient of 94% [28]. Voyant et al. conducted a study in France using data obtained from a ground station to predict solar irradiation using the ANN model. They evaluated the impact of external data and data selection on the prediction. Additionally, they examined the effects of the number of hidden layers and neurons, activation function, learning algorithm, and comparison functions used during the learning stage [29].

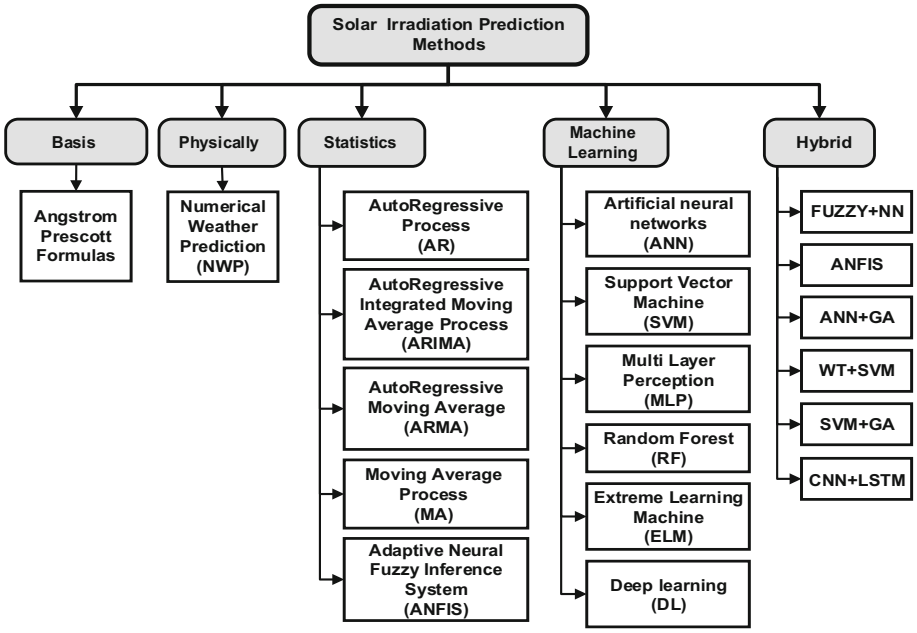


Fig. 1 Solar irradiation prediction methods [11, 15]

Mohandes et al. [30] conducted a solar irradiation prediction study in Saudi Arabia using data obtained from 41 stations and employing the "logsig" activation function. Wang et al. [19] performed solar irradiation prediction in China using an ANN model with the LM algorithm and "logsig" and "purelin" activation functions. Yadav and Chandel [31] predicted solar irradiation in India using an ANN model with the LM algorithm and the "tansig" activation function. In addition to these studies, there are numerous other works on solar irradiation prediction carried out by Quej et al. [32], in Mexico, Marzo et al. [33], in Chile, Halabi et al. [34], in Malaysia, Marzouq et al. [35], in Morocco, Jahani et al. [36], in Iran, Antonopoulos et al. [37], in Greece, and Guermoui et al. [38], in Algeria.

Turkey is a country surrounded by seas on three sides, with an average annual total irradiation of 1527.46 kWh/m^2 and an average annual sunshine duration of 2741 h. The regions located in the south and east, in particular, are suitable areas for solar energy investments. Therefore, successful results are obtained in solar irradiation prediction studies, especially for the regions located in these areas. These studies and their details are presented in Table 1.

In this study, an attempt has been made to predict the instantaneous solar irradiance, which is the most crucial parameter for electricity generation in PV systems. Real meteorological data, including ambient temperature, relative humidity, atmospheric pressure, wind speed, and solar irradiance, recorded at 5-s intervals depending on the time series from the meteorological station, were utilized. Data from the meters are transferred to a workstation (Intel Xeon Silver 4114 CPU @ 2.20 GHz, 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. Various training algorithms and activation functions based on ANN were employed. With the use of ANN, two different network structures with 5 different functions were developed to analyze the relationship between meteorological input data (such as ambient temperature, relative humidity, atmospheric pressure, wind speed) and solar irradiance based on the time series (year, month, day, hour, minute, and second). The obtained results enable the determination of the effectiveness of network structure, data set distribution, algorithms, and functions in solar irradiance prediction using basic meteorological data. In this context, the main contributions of the study are as follows:

- Determining the effectiveness of activation functions used in the ANN model for solar irradiance prediction.
- Developing an ANN model for solar irradiance prediction using data obtained from a single geographical location and ensuring its generalizability to other regions.
- Determining the impact of data distribution used during the training and testing stages of the ANN model on solar irradiance prediction.
- Investigating the prediction performance of the ANN model in solar irradiance prediction using statistical indicators to assess accuracy.
- Determining the most suitable network architectures in ANN models used to predict solar irradiance for designing or assessing solar energy facilities in regions without meteorological measurement stations.
- Developing various ANN models with the capability to analyze the relationship between meteorological parameters and solar irradiance.

Table 1 Solar irradiation estimation studies and evaluation metrics in Turkey

Reference	Prediction model	Parameters used	Assessment metrics		
			R^2	MAPE	RMSE
Ağbulut et al. [4]	ANN	H_0, d, T_{max},T_{min},B	0.932	15.92	2.157 MJ/m^2
Demir et al. [39]	SVM, LSTM, ELM	T_{max},T_{min},w,H,L_o,L_a,Z	0.895	15.17	2.297 MJ/m^2
Sözen et al. [40]	ANN	L_o,L_a,Z,T_{avr},S	0.998	6.73	NA
Ozgoren et al. [41]	ANN	H_0,L_o,L_a, Z,S,H,T	0.993	5.34	NA
Kaba et al. [42]	DL	H_0,T_{max},T_{min},S,B	0.98	NA	0.78
Yildirim et al. [43]	LSTM, MLP, ANFIS		0.945	NA	101.20 W/m^2
Yildirim et al. [44]	ANN	H_0,L_o,L_a,Z,S,H,T	0.961	NA	0.14
Bilgili et al. [45]	ANN	S,T,w,M,d	0.965	7.88	NA
Alizamir et al. [46]	MLP, ANFIS, GBT	T_{max},T_{min},w,H	0.864	NA	7.352 MJ/m^2
Ozan Şenkal [47]	ANN	L_o,L_a,Z,T	0.934		0.320 MJ/m^2
Karaman et al. [48]	ELM	w, T, S	0.991	NA	0.0297

- Determining the effectiveness of training algorithms and transfer functions on prediction accuracy in ANN models modeled with different network architectures.

The remaining sections of the article are structured as follows: Sect. 2 provides details about ANN, the study area, the data set used, and the criteria for performance evaluation based on research methodology. In Sect. 3, the results obtained from the ANN models are presented, discussed, and analyzed. The final section presents the main findings of the study.

2 Methodology

In this study, many ANN-based approaches to predict solar irradiation were designed and applied, and their results were confirmed. In the first stage of the study, models were developed to analyze the relationship between time-dependent solar irradiance and a set of meteorological parameters in a specific geographical area. In these models, the effectiveness of different training algorithms (trainscg, trainlm, trainrp, traingd, and trainoss) and different transfer functions (logsig, poslin, tansig, purelin, radbas, hardlim, tribas, and satlin), which are easy to implement, widely utilized in the literature, and have demonstrated predictive accuracy, has been investigated. The functions that give the best results in these models were determined. In the second stage, the performance in forecasting success of the created network models was analyzed, and the best-performing model was determined. Each model was dedicated to estimating the amount of solar irradiation at times, including all seasonal periods of the year. All components of all approaches were designed, implemented, and validated according to the flowchart in Fig. 2 using the multi-layer perceptron (MLP) neural network model created with the MATLAB2021a software. The process of preparing the information and its uses in modeling are described as follows:

- To prevent prediction models from being affected by different ranges, the input data are normalized as depicted in Fig. 2 (in the range − 1 to + 1) [13].
- To create neural networks by defining different functions with neural network layers and neurons.
- To train networks according to defined functions.
- To test networks according to the defined data.
- To receive output data from the neural network.
- To return normalized data to output data.
- To evaluate the performance of the neural network and compare its output with the measured data.

Using meteorological parameters measured by sensors depending on time parameters, the study, network training, and modeling are based on two different network models. In both network models, the input parameters were selected as similar. The first type of network model (10-10-1) is called ANN_I, which consists of small hidden layers that are the same as the number of input parameters. The second type of network model (10-100-1) is called ANN_II, in which a large hidden layer is formed by squaring the number of input parameters. Thus, in addition to different training algorithms and transfer functions in solar irradiation estimation, the effect of the network structure was determined by taking into account the model, the number of neurons, and different hidden layers. The operating conditions for the design of network models are given in Table 2. Training of the ANN architecture was continued until the minimum MSE value was reached.

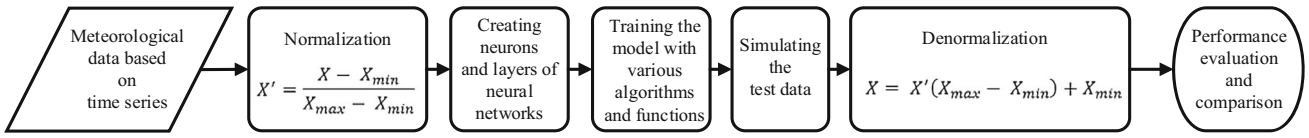


Fig. 2 The flowchart used in the study

Table 2 Operating conditions set for models

Operation network type	Multilayered Back-propagation Algorithm (MLP)
Training function	Trainscg, Trainlm, Trainrp, Traingd, Trainoss
Performance function	MSE
Number of inputs	10
Number of outputs	1
Number of hidden layers	1
Number of neurons in the hidden layer	10, 100
Transfer function	Logsig, Poslin, Tansig, Purelin, Radbas, Hardlim, Tribas, Satlin

2.1 Artificial neural networks (ANN)

ANN models demonstrate higher accuracy in solar irradiation prediction compared to empirical models and other forecasting methods. Qazi et al. [49], by evaluating the work of various researchers, have proposed the application of ANN techniques to achieve better results in future studies aimed at forecasting solar irradiation. The ANN technique is one of the most popular performance applications and is widely used in many areas such as optimization, regression, control, and classification [50–52]. The basic unit of an artificial neural network is a neuron, which uses a transfer function to formulate output. Choosing the number of neurons in the hidden layer and the number of hidden layers for any artificial neural network model is very complex. Often, one hidden layer is sufficient for complex applications. The training behavior of neurons depends on their activation function. Each input is multiplied by a weight as a relationship parameter between a neuron and several layers of neurons. A transfer function is applied to obtain the neuron result in the final stage [51]. The multi-layer network used in this study is presented schematically with input parameters in Fig. 3.

As can be seen in Fig. 3, the parameters of ambient temperature, relative humidity, wind speed, and atmospheric pressure, which are measured depending on the time parameters, are included in the first layer. According to the weights assigned to neurons and the number of neurons in the latent layer and the last layer, the total amount of solar irradiation is accepted as the network output.

2.2 Study region

The study area is located in the province of Hakkari in the southern part of the Eastern Anatolia Region with a surface area of 9551 km^2, accepting a sunshine duration of approximately 14 h especially in summer, and high solar irradiation potential. Ahlborn brand "Almemo 2590 Datalogger" device with measurement sensors is used in the meteorological station at Hakkari University Çölemerik Vocational School campus (N: 37.571799, E: 43.724936) located at an altitude of 1755 m. above sea level. Some important geographical details of the study area and its appearance on the map of Turkey are given in Fig. 4. The selected location has a cold climate zone and has a very high value in terms of solar irradiation value and sunshine duration. From this point of view, the place chosen for the study is quite important. In addition, the region chosen as the study area in the current solar irradiation estimation studies in the literature has very suitable areas for the installation of solar power plants and the efficient operation of the system. For this reason, the results obtained from this study are a guide for academic and investment studies to be made for the operation and operation of solar energy systems. The fact that no study has been carried out on the data obtained with real measurement values in Hakkari adds innovation and contributes to the study.

2.3 Data

The actual data measured from the meteorological station located at the campus of Hakkari University are used during the design and verification phases. Solar irradiation is constantly changing due to instantaneously changing weather conditions, and it can be obtained better with the data measured in shorter intervals [53]. The data set consists of one million data with a five-second period between 2019 and 2021. These measured data versus the measurement time include solar irradiation, ambient temperature, wind speed, relative humidity, and atmospheric pressure from the measurement sensors. The chosen study area, Hakkari, represents a region in which all four seasons with their respective meteorological characteristics occur throughout the year. Conducting measurements at short intervals, as frequent as

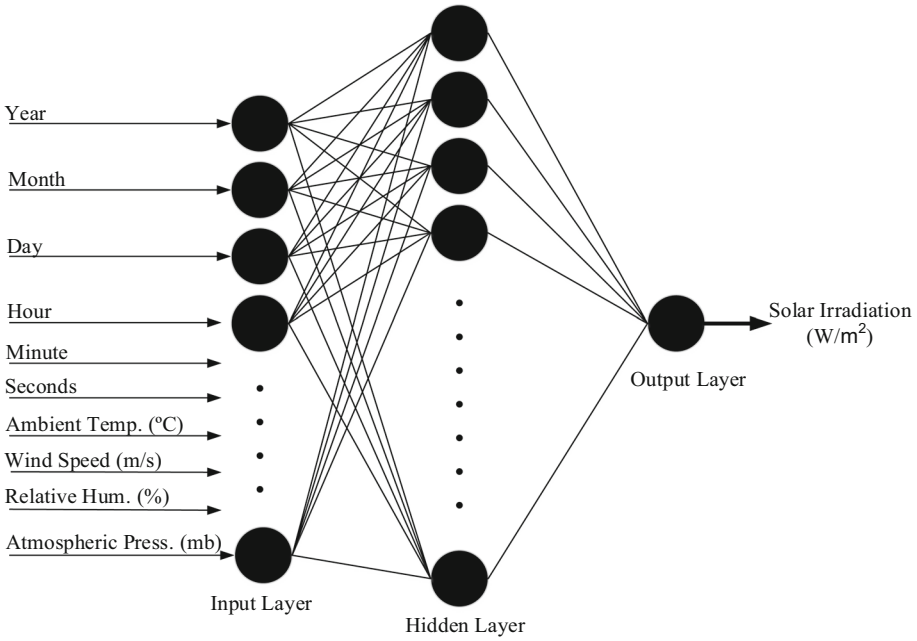
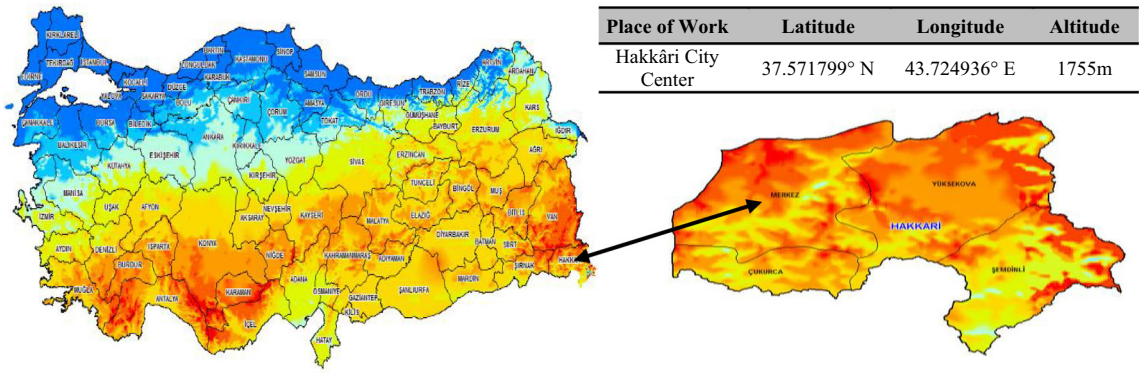


Fig. 3 Schematic diagram of the multi-layered ANN with the inputs used in the study



Place of Work	Latitude	Longitude	Altitude
Hakkâri City Center	37.571799° N	43.724936° E	1755m

Fig. 4 Location and information of the study zone on the map of Turkey

Table 3 Meteorological record samples (data set sample)

Year	Month	Day	Hour	Min	Sec	Ambient temp. (°C)	Wind speed (m/s)	Relative humidity (%)	Atmospheric pressure (hPa)	Solar irradiation (W/m^2)
2019	03	12	06	22	49	2.3	0.0	100	812.2	2
2019	12	30	08	22	38	− 4.9	0.1	89.1	836.4	19
2020	06	26	11	49	38	30.5	1.0	21.6	818.4	957
2020	11	12	06	47	23	23.0	3.6	92.4	828.4	9
2021	01	19	07	54	17	− 1.0	0.3	100	821.5	11
2021	09	04	08	05	35	21.6	1.1	34.8	812.9	467

every 5 s, and covering all seasons enables the collection of comprehensive data. Interestingly, the obtained data not only exhibit the unique characteristics of the study area but also display similarities with other geographically comparable regions within Turkey. Table 3 provides an example of data obtained from other regions with similar characteristics to the study area. As a result, this study benefits from an extensive data set comprising long-term and multi-seasonal measurements, along with diverse meteorological parameters, indicating that the results obtained are likely applicable to other regions as well.

The data set is divided into two subgroups: A design subset including 800,000 pieces of data and a simulation test subset including 200,000 pieces of data. Table 3 shows some examples of a data set sample. In this table, the first six columns represent time parameters, and the last five columns represent five meteorological records based on time parameters.

As shown in Fig. 5, the data obtained from the weather station are divided into two parts: training and simulation. The training data set is designed as 5% testing, 5% validation and 90% training data, 10% testing, 10% validation and 80% training data, 15% testing, 15% verification and 70% training data, 20% testing, 20% verification and 60% training data and 25% testing, 25% validation and 50% training data, respectively. After completing the training stage, the simulation was performed with randomly selected 200,000 data which are not included in the training data set to simulate the model and compare the results. The evaluation of all models was made according to the simulation results.

2.4 Performance evaluation criteria

Prediction performance is an evaluation metric that measures how well a model matches the actual values and the accuracy of its predictions. Commonly used performance metrics serve the purpose of evaluating the results of prediction models and comparing them with one another. As a dependency scale metric, RMSE cannot be used for model comparison across multiple data sets [38]. In this context, MSE, nRMSE, MBE, MAPE, and R^2 metrics have been used to compare the performance success of prediction models. These performance evaluation criteria, along with their equations and explanations, are provided in Table 4.

In Table 4, N refers to the number of data; y_i refers to the true value of global solar irradiation; $\widehat{y_i}$ refers to forecast value; $\overline{y_i}$ refers to the average of the measured global solar irradiation.

3 Results and discussion

In this research, ANN reaches the optimum result to estimate the solar irradiation by changing the transfer function types and the number of neurons in the hidden layer using trainlm, trainscg, trainrp, traingd, and trainoss training algorithms. Thus, in solar irradiation estimation studies, the performance of transfer functions in network structures with small and large neurons is determined in different training algorithms. Tables 5, 6, 7, 8 , and 9 present the modeling results with several conditions in the mentioned categories. A network structure was modeled with ten inputs, a single layer, having 10 and 100 neurons to estimate the amount of solar irradiation using ANN. Measurements of ambient temperature, wind speed, relative humidity, and atmospheric pressure at intervals of 5 s and the time parameters in the form of year, month, day, hour, minute, and second are introduced to the input layer. There is a neuron in the output layer since only one parameter is considered as the network output. The "purelin" function was selected as the conversion function for all models, and the most efficient function type was determined by using the transfer functions "logsig," "poslin," "tansig," "purelin," "radbas," "hardlim," "tribas," and "satlin" for hidden layers. Each model created in the ANN was tested five times, and the best-performing model was determined.

Prediction performance metrics are used to evaluate the accuracy of the model and compare the performance of different models or methods. Lower error values indicate better prediction performance, while higher error values may indicate that the model needs improvement. However, each metric has its own advantages and disadvantages. The most appropriate metric should be selected depending on the type of model and application area. In this study, the evaluation results on the MSE metric on how the percentage change of the data in the input layer of the ANN model in solar irradiation estimation affects the solar irradiation estimation performance according to the transfer functions and the training algorithms included in the ANN structure are shown in Tables 5, 6, 7, 8, and 9.

When Tables 5, 6, 7, 8, and 9 are examined, ANN_II determined the best predictive value for the "tansig" transfer function as 97.58% using 15% test data with the "trainlm" training algorithm. ANN_I model determined the best predictive value for the "logsig" transfer function as 96.25% using 15% test data with the "trainlm" training algorithm. Both network models showed similar predictive values when using the "logsig" and "tansig" transfer functions. In these algorithms, the results with the lowest performance were obtained using the "purelin" and "hardlim" functions. In this study, the optimum result is reached by changing the function types and the number of neurons in the hidden layer. The statistical performance of the modeling results performed under the best test conditions is presented in Tables 10, 11.

In line with the changes made on the data set, the effect of the change of training algorithms and transfer functions on the performance results shows little change. The reason for this is that the ANN model is trained with a large number of

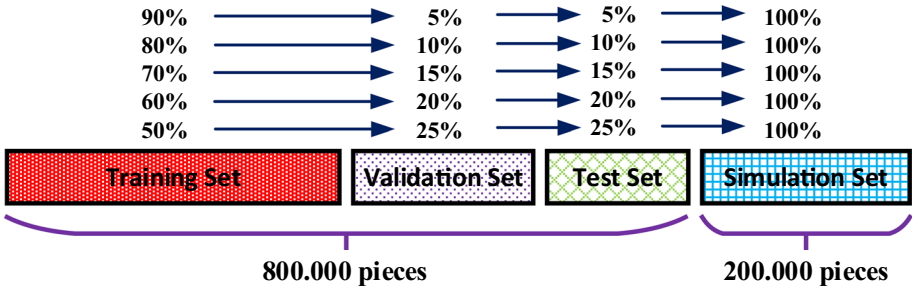


Fig. 5 Data set layout for training, verification, and testing of solar irradiation forecasting models

Table 4 Performance metrics used in the study and their explanations

Metric	Equation	Explanation
MSE	$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \widehat{y_i})^2$	It is used to find the difference between the target and output that the neural network produces in the training process. MSE is usually the parameter that is minimized by the training algorithm
nRMSE	$\frac{\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\widehat{y_i})^2}}{\overline{y_i}} * 100$	It is obtained by RMSE and the average value of the measured data. If the nRMSE is small, the forecast model has a better performance. The success of the forecast model is considered as follows Excellent: nRMSE < 10% Good: 10% < nRMSE < 20% Moderate: 20% < nRMSE < 30% Bad: nRMSE > 30%
MBE	$\mathrm{MBE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \widehat{y_i})$	It is an important metric for the long-term performance of forecast models. The smaller value of MBE indicates that the forecast model has a better performance. A positive MBE means an overestimate of the calculated value. A negative MBE indicates an underestimate in the calculated value
MAPE	$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i-\widehat{y_i})}{y_i} * 100$	The average of the absolute values of the prediction errors is the percentage of the absolute values of the actual data. The lower value of the MAPE is an indicator of the better performance of the model. The success of the forecast model is considered as follows High forecast: MAPE $\leq$ 10% Good prediction: 10% < MAPE $\leq$ 20% Reasonable prediction: 20% < MAPE $\leq$ 50% Incorrect prediction: MAPE > 50%
R^2	$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i-\widehat{y_i})^2}{\sum_{i=1}^{N}(y_i-\overline{y_i})^2}$	It provides information about how well a model can predict a set of measured data. Its value varies between 0 and 1. A value of R^2 approaching 1 is an indication of better performance

Table 5 Prediction performance of "trainscg" training algorithm based on data set and transfer functions

Test data	Network structure	Trainscg							
		Logsig	Poslin	Tansig	Purelin	Radbas	Hardlim	Tribas	Satlin
5%	ANN_I	95.91	95.53	96.06	86.56	96.14	81.97	96.04	96.03
	ANN_II	96.15	96.40	96.20	86.56	96.86	87.35	96.68	96.76
10%	ANN_I	96.04	95.81	96.12	86.56	96.13	81.00	95.92	96.00
	ANN_II	96.45	96.52	96.47	86.56	97.13	88.13	96.86	96.95
15%	ANN_I	96.08	95.63	96.10	86.56	96.10	81.77	96.06	96.08
	ANN_II	96.89	96.47	96.97	86.56	97.20	87.18	97.11	96.91
20%	ANN_I	96.01	95.66	96.16	86.56	96.13	80.66	96.04	96.00
	ANN_II	96.91	96.53	96.89	86.56	97.10	84.26	96.43	96.18
25%	ANN_I	96.07	95.82	96.15	86.56	96.10	81.48	96.11	95.75
	ANN_II	96.85	96.36	96.95	86.56	96.88	85.66	96.49	96.07

Table 6 Prediction performance of "trainlm" training algorithm based on data set and transfer functions

Test data	Network structure	Trainlm							
		Logsig	Poslin	Tansig	Purelin	Radbas	Hardlim	Tribas	Satlin
5%	ANN_I	96.22	96.09	96.23	86.56	96.20	81.21	96.17	96.14
	ANN_II	97.49	97.25	97.47	86.56	97.47	87.41	97.27	97.33
10%	ANN_I	96.23	96.08	96.20	86.56	96.19	81.12	96.12	86.15
	ANN_II	97.47	97.25	97.47	86.56	97.16	88.07	97.29	97.31
15%	ANN_I	96.25	96.09	96.21	86.56	96.20	81.71	96.08	96.17
	ANN_II	97.41	97.24	97.58	86.56	97.24	87.93	97.19	97.22
20%	ANN_I	96.23	96.15	96.22	86.56	96.18	80.68	96.11	96.15
	ANN_II	96.89	96.91	96.90	86.56	96.54	86.53	96.98	97.12
25%	ANN_I	96.22	96.11	96.24	86.56	96.17	81.21	96.14	96.17
	ANN_II	97.18	96.95	96.91	86.56	97.23	87.09	97.11	97.25

Table 7 Prediction performance of "trainrp" training algorithm based on data set and transfer functions

Test data	Network structure	Trainrp							
		Logsig	Poslin	Tansig	Purelin	Radbas	Hardlim	Tribas	Satlin
5%	ANN_I	96.09	95.93	96.06	86.56	96.08	82.96	96.01	95.96
	ANN_II	96.89	96.23	96.83	86.56	96.98	87.52	96.82	96.58
10%	ANN_I	96.07	95.91	96.12	86.56	96.10	81.68	96.02	96.02
	ANN_II	96.80	96.70	96.70	86.56	96.88	87.61	96.90	96.71
15%	ANN_I	96.12	95.84	96.06	86.56	96.10	81.69	96.02	96.02
	ANN_II	96.90	96.63	96.63	86.56	96.93	87.37	96.93	96.83
20%	ANN_I	96.06	96.00	96.08	86.56	96.08	81.92	96.04	96.10
	ANN_II	96.68	96.91	96.58	86.56	96.59	83.69	96.74	96.99
25%	ANN_I	96.11	95.90	96.06	86.56	96.13	83.24	96.09	96.04
	ANN_II	96.92	96.25	96.71	86.56	96.78	85.69	96.56	96.68

Table 8 Prediction performance of "traingd" training algorithm based on data set and transfer functions

Test data	Network structure	Traingd							
		Logsig	Poslin	Tansig	Purelin	Radbas	Hardlim	Tribas	Satlin
5%	ANN_I	91.05	94.90	94.12	86.56	92.80	79.92	90.01	90.16
	ANN_II	91.55	94.92	94.82	86.56	93.20	80.13	92.13	90.88
10%	ANN_I	90.30	95.10	94.50	86.56	92.62	78.13	91.12	90.28
	ANN_II	92.16	95.46	95.62	86.56	93.28	79.26	93.46	92.61
15%	ANN_I	92.22	95.14	94.66	86.56	92.10	78.69	90.05	90.55
	ANN_II	92.98	95.63	94.79	86.56	92.53	80.12	91.25	90.92
20%	ANN_I	89.66	94.25	94.12	86.56	91.28	77.23	90.53	89.22
	ANN_II	90.13	94.69	94.56	86.56	91.99	79.36	91.33	90.19
25%	ANN_I	89.92	94.01	93.92	86.56	91.12	77.56	90.11	90.33
	ANN_II	90.25	94.83	94.22	86.56	92.13	78.52	90.91	91.05

Table 9 Prediction performance of "trainoss" training algorithm based on data set and transfer functions

Test data	Network structure	Trainoss							
		Logsig	Poslin	Tansig	Purelin	Radbas	Hardlim	Tribas	Satlin
5%	ANN_I	95.90	95.60	95.80	86.56	96.00	80.70	96.00	96.00
	ANN_II	96.00	96.00	96.10	86.56	96.10	81.20	96.10	96.10
10%	ANN_I	95.80	95.90	95.90	86.56	96.00	80.40	96.00	96.10
	ANN_II	95.90	96.00	96.10	86.56	96.20	87.50	96.50	96.20
15%	ANN_I	95.80	96.00	95.90	86.56	96.00	82.00	96.10	96.10
	ANN_II	96.00	96.20	96.10	86.56	96.90	87.20	96.80	96.90
20%	ANN_I	96.00	96.10	96.00	86.56	96.00	81.40	96.10	96.00
	ANN_II	96.00	96.30	96.20	86.56	96.90	87.10	96.80	96.70
25%	ANN_I	95.90	95.80	96.10	86.56	96.00	81.00	96.20	96.10
	ANN_II	96.00	96.10	96.20	86.56	96.10	87.00	96.40	96.00

data sets. The ANN model can analyze the changes in almost all meteorological data sets during the training phase and can show successful performance results during the testing phase. However, it can still be suggested to use "15%" test data, "trainlm" training algorithm, and "tansig" transfer function in solar irradiation estimation studies using a data set consisting of basic meteorological parameters.

Successful results could not be obtained in solar irradiation estimation by using the "traingd" training algorithm and "satlin" and "purelin" transfer functions used in the study. The reason for this is that the training algorithm is a network training function that updates the weight and deviation values according to the gradient descent, while the network model is running. Because "trainlm" is a network training function that updates the weight and bias values according to LM optimization, and it also contains more memory than other training algorithms; very successful results can be obtained in solar irradiation estimation studies. The results of the forecast performance shown in Tables 5, 6, 7, 8, and 9 and the evaluation metrics shown in Tables 10, 11 demonstrate this situation.

In this research, the optimum result is reached by changing the function types and the number of neurons in the hidden layer. Tables 10, 11 give the statistical performance of the modeling results for ANN_I and ANN_II performed under the best test conditions. In Tables 5, 6, 7, 8, 9, 10, and 11 all of the methodologies used throughout the study were used in the same way. Within the scope of the study, the effects of all existing functions and data set changes were explained in detail and their performance values were analyzed.

Figures 6a and 7a show slight deviations between the forecast data and the measurement data depending on the time series in the most successful functions using the ANN_I and ANN_II models. Figures 6b and 7b show the worst prediction success in the same models. In these shapes, a solid line represents the actual value, and a dashed line represents the predicted value.

In the ANN-I model, the prediction error nRMSE ranges from 8.947 to 20.749. Based on the nRMSE performance metric intervals shown in Table 4, it is evident that the worst prediction result was obtained with the "traingd" training algorithm and the "hardlim" transfer function used in the ANN model. On the other hand, the most successful model achieved an nRMSE value of 8.947, utilizing the "trainlm" training algorithm and the "tansig" transfer function.

In the ANN-II model, the prediction error nRMSE ranges from 7.354 to 20.167. In this neural network model, the most successful prediction performance was observed with an nRMSE value of 7.354, achieved using the "trainlm" training algorithm and the "tansig" and "radbas" transfer functions. However, it is worth noting that although the "tansig" transfer function resulted in a higher correlation coefficient, it is recommended for better performance. On the other hand, when the same network structure was tested with the "trainoss" training algorithm, the "hardlim" transfer function exhibited significantly poor performance.

In all the developed models except for the ones created using the "purelin" and "hardlim" functions, it is observed that the R index has a high value. However, since this index shows little variation depending on the number of neurons and layer functions, it is not appropriate to use it in the evaluation of models. During the modeling process, better results can be achieved by increasing the number of hidden layers. However, increasing the number of neurons may also lead to longer computation times and potentially reduce the accuracy of the modeling. For predicting solar irradiation using meteorological parameters, the ANN_I model with the "trainlm" training algorithm and the "logsig" transfer function, as well as the ANN_II models with the "tansig" transfer function, was selected as the optimal neural network models among

Table 10 Statistical performance of the test results of the ANN_I model according to the algorithm and data set distribution

Training Alg	Transfer Func	90-5-5%				80-10-10%				70-15-15%				60-20-20%				50-25-25%			
		R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE
Trainscg	Logsig	96.58	0.040	9.302	1.333	96.82	0.039	9.185	1.300	96.82	0.039	9.185	1.300	96.84	0.038	9.067	1.267	96.84	0.039	9.185	1.300
Trainscg	Poslin	94.12	0.044	9.756	1.467	94.46	0.041	9.418	1.367	94.08	0.043	9.645	1.433	94.08	0.043	9.645	1.433	94.46	0.041	9.418	1.367
Trainscg	Tansig	96.80	0.038	9.067	1.267	96.80	0.038	9.067	1.267	96.80	0.038	9.067	1.267	96.80	0.038	9.067	1.267	96.80	0.038	9.067	1.267
Trainscg	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467
Trainscg	Radbas	95.06	0.039	9.185	1.300	95.16	0.038	9.067	1.267	95.16	0.038	9.067	1.267	95.16	0.038	9.067	1.267	95.04	0.039	9.185	1.300
Trainscg	Hardlim	47.38	0.180	19.733	6.000	47.03	0.189	20.221	6.300	47.22	0.182	19.843	6.067	47.13	0.193	20.433	6.433	47.04	0.188	20.167	6.267
Trainscg	Tribas	95.32	0.039	9.185	1.300	95.28	0.040	9.302	1.333	95.28	0.039	9.185	1.300	95.28	0.039	9.185	1.300	95.28	0.038	9.067	1.267
Trainscg	Satlin	94.82	0.039	9.185	1.300	94.82	0.039	9.185	1.300	94.82	0.039	9.185	1.300	94.78	0.040	9.302	1.333	94.92	0.038	9.067	1.267
Trainlm	Logsig	97.02	0.037	8.947	1.233	97.02	0.037	8.947	1.233	97.02	0.037	8.947	1.233	97.02	0.037	8.947	1.233	97.02	0.037	8.947	1.233
Trainlm	Poslin	95.08	0.039	9.185	1.300	95.08	0.039	9.185	1.300	95.08	0.039	9.185	1.300	95.16	0.038	9.067	1.267	95.16	0.038	9.067	1.267
Trainlm	Tansig	97.01	0.037	8.947	1.233	97.01	0.037	8.947	1.233	97.01	0.037	8.947	1.233	97.01	0.037	8.947	1.233	97.01	0.037	8.947	1.233
Trainlm	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467
Trainlm	Radbas	95.36	0.037	8.947	1.233	95.30	0.038	9.067	1.267	95.30	0.038	9.067	1.267	95.30	0.038	9.067	1.267	95.30	0.038	9.067	1.267
Trainlm	Hardlim	47.05	0.187	20.113	6.233	47.04	0.188	20.167	6.267	47.22	0.182	19.843	6.067	47.11	0.193	20.433	6.433	47.17	0.187	20.113	6.233
Trainlm	Tribas	95.36	0.038	9.067	1.267	95.36	0.038	9.067	1.267	95.32	0.039	9.185	1.300	95.34	0.038	9.067	1.267	95.36	0.038	9.067	1.267
Trainlm	Satlin	94.90	0.038	9.067	1.267	94.90	0.038	9.067	1.267	94.88	0.038	9.067	1.267	94.90	0.038	9.067	1.267	94.90	0.038	9.067	1.267
Trainrp	Logsig	96.74	0.039	9.185	1.300	96.78	0.039	9.185	1.300	96.76	0.038	9.067	1.267	96.74	0.039	9.185	1.300	96.76	0.038	9.067	1.267
Trainrp	Poslin	94.66	0.040	9.302	1.333	94.68	0.040	9.302	1.333	94.60	0.041	9.418	1.367	94.68	0.040	9.302	1.333	94.60	0.041	9.418	1.367
Trainrp	Tansig	96.48	0.039	9.185	1.300	96.62	0.038	9.067	1.267	96.46	0.039	9.185	1.300	96.48	0.039	9.185	1.300	96.48	0.039	9.185	1.300
Trainrp	Purelin	62.36	0.134	17.026	4.467	62.34	0.134	17.026	4.467	62.34	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467
Trainrp	Radbas	95.14	0.039	9.185	1.300	95.16	0.038	9.067	1.267	95.20	0.038	9.067	1.267	95.14	0.039	9.185	1.300	95.16	0.038	9.067	1.267
Trainrp	Hardlim	47.80	0.170	19.177	5.667	47.52	0.183	19.897	6.100	47.54	0.183	19.897	6.100	47.50	0.180	19.733	6.000	47.84	0.167	19.007	5.567

Table 10 (continued)

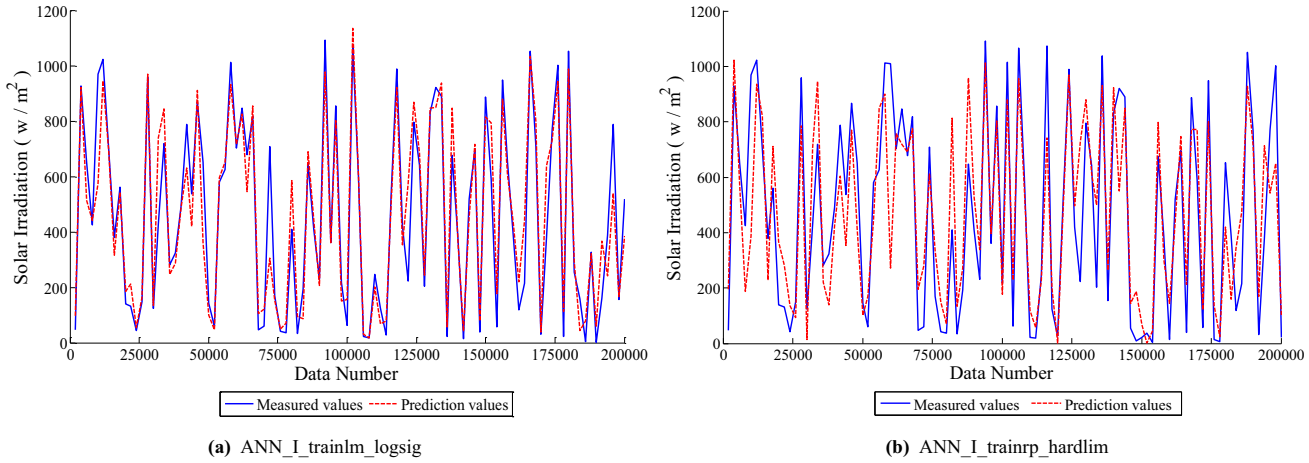
Training Alg	Transfer Func	90-5-5%				80-10-10%				70-15-15%				60-20-20%				50-25-25%			
		R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE
Trainrp	Tribas	95.34	0.039	9.185	1.300	95.36	0.039	9.185	1.300	95.36	0.039	9.185	1.300	95.34	0.039	9.185	1.300	95.34	0.039	9.185	1.300
Trainrp	Satlin	95.04	0.040	9.302	1.333	95.12	0.039	9.185	1.300	95.12	0.039	9.185	1.300	95.12	0.039	9.185	1.300	95.12	0.039	9.185	1.300
Traingd	Logsig	87.28	0.097	14.486	3.233	87.28	0.097	14.486	3.233	88.04	0.095	14.336	3.167	87.32	0.097	14.486	3.233	87.30	0.097	14.486	3.233
Traingd	Poslin	92.86	0.051	10.504	1.700	92.94	0.049	10.296	1.633	92.96	0.049	10.296	1.633	92.84	0.051	10.504	1.700	92.52	0.053	10.708	1.767
Traingd	Tansig	90.12	0.057	11.104	1.900	90.18	0.055	10.908	1.833	89.62	0.053	10.708	1.767	89.46	0.055	10.908	1.833	89.62	0.059	11.298	1.967
Traingd	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.32	0.134	17.026	4.467	62.34	0.134	17.026	4.467	62.34	0.134	17.026	4.467
Traingd	Radbas	85.24	0.113	15.635	3.767	85.94	0.099	14.635	3.300	85.96	0.097	14.486	3.233	86.08	0.103	14.927	3.433	86.08	0.103	14.927	3.433
Traingd	Hardlim	40.62	0.199	20.749	6.633	40.62	0.197	20.644	6.567	40.66	0.197	20.644	6.567	40.62	0.199	20.749	6.633	40.62	0.199	20.749	6.633
Traingd	Tribas	86.32	0.101	14.782	3.367	86.88	0.097	14.486	3.233	86.48	0.099	14.635	3.300	86.50	0.099	14.635	3.300	86.24	0.103	14.927	3.433
Traingd	Satlin	86.62	0.099	14.635	3.300	86.68	0.099	14.635	3.300	86.80	0.097	14.486	3.233	86.68	0.099	14.635	3.300	86.68	0.099	14.635	3.300
Trainoss	Logsig	95.14	0.041	9.418	1.367	94.98	0.042	9.532	1.400	95.16	0.041	9.418	1.367	95.38	0.040	9.302	1.333	95.18	0.041	9.418	1.367
Trainoss	Poslin	94.10	0.044	9.756	1.467	94.32	0.041	9.418	1.367	94.22	0.040	9.302	1.333	94.50	0.039	9.185	1.300	94.30	0.041	9.418	1.367
Trainoss	Tansig	94.40	0.041	9.418	1.367	94.41	0.041	9.418	1.367	94.41	0.041	9.418	1.367	94.40	0.040	9.302	1.333	94.34	0.039	9.185	1.300
Trainoss	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467
Trainoss	Radbas	94.86	0.040	9.302	1.333	95.20	0.040	9.302	1.333	95.14	0.040	9.302	1.333	94.86	0.039	9.185	1.300	94.86	0.039	9.185	1.300
Trainoss	Hardlim	42.50	0.193	20.433	6.433	41.96	0.196	20.592	6.533	43.28	0.179	19.678	5.967	42.12	0.185	20.005	6.167	42.02	0.190	20.274	6.333
Trainoss	Tribas	95.22	0.040	9.302	1.333	95.22	0.040	9.302	1.333	95.38	0.039	9.185	1.300	95.40	0.039	9.185	1.300	85.58	0.038	9.067	1.267
Trainoss	Satlin	95.12	0.040	9.302	1.333	95.34	0.039	9.185	1.300	95.32	0.039	9.185	1.300	95.32	0.039	9.185	1.300	95.30	0.039	9.185	1.300

Table 11 Statistical performance of the test results of the ANN_II model according to the algorithm and data set distribution

Training alg	Transfer Func	90-5-5%				80-10-10%				70-15-15%				60-20-20%				50-25-25%			
		R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE
Trainscg	Logsig	96.72	0.038	9.067	1.267	96.88	0.035	8.702	1.167	96.94	0.031	8.189	1.033	96.94	0.030	8.056	1.000	96.94	0.031	8.189	1.033
Trainscg	Poslin	94.86	0.035	8.702	1.167	94.88	0.034	8.576	1.133	94.88	0.035	8.702	1.167	94.88	0.034	8.576	1.133	94.86	0.036	8.825	1.200
Trainscg	Tansig	96.80	0.038	9.067	1.267	96.84	0.035	8.702	1.167	96.88	0.030	8.056	1.000	96.88	0.031	8.189	1.033	96.90	0.030	8.056	1.000
Trainscg	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467
Trainscg	Radbas	96.34	0.031	8.189	1.033	96.82	0.028	7.783	0.933	96.84	0.028	7.783	0.933	96.82	0.029	7.921	0.967	96.88	0.030	8.056	1.000
Trainscg	Hardlim	49.74	0.126	16.510	4.200	50.86	0.118	15.977	3.933	47.70	0.128	16.641	4.267	48.64	0.132	16.899	4.400	48.78	0.130	16.770	4.333
Trainscg	Tribas	95.92	0.033	8.449	1.100	95.96	0.031	8.189	1.033	96.00	0.029	7.921	0.967	96.00	0.029	7.921	0.967	95.98	0.030	8.056	1.000
Trainscg	Satlin	95.14	0.032	8.320	1.067	95.16	0.030	8.056	1.000	95.18	0.030	8.056	1.000	95.16	0.030	8.056	1.000	95.16	0.030	8.056	1.000
Trainlm	Logsig	97.22	0.025	7.354	0.833	97.20	0.025	7.354	0.833	97.26	0.024	7.206	0.800	97.26	0.024	7.206	0.800	97.24	0.025	7.354	0.833
Trainlm	Poslin	96.60	0.027	7.643	0.900	96.60	0.027	7.643	0.900	96.60	0.027	7.643	0.900	96.62	0.027	7.643	0.900	96.60	0.027	7.643	0.900
Trainlm	Tansig	97.51	0.025	7.354	0.833	97.83	0.024	7.206	0.800	97.80	0.024	7.206	0.800	97.80	0.025	7.354	0.833	97.81	0.024	7.206	0.800
Trainlm	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467
Trainlm	Radbas	96.64	0.025	7.354	0.833	96.60	0.025	7.354	0.833	96.60	0.025	7.354	0.833	96.61	0.025	7.354	0.833	96.61	0.025	7.354	0.833
Trainlm	Hardlim	54.13	0.125	16.444	4.167	54.68	0.119	16.045	3.967	54.66	0.119	16.045	3.967	54.68	0.117	15.909	3.900	54.72	0.117	15.909	3.900
Trainlm	Tribas	96.29	0.027	7.643	0.900	96.30	0.027	7.643	0.900	96.30	0.027	7.643	0.900	96.34	0.025	7.354	0.833	96.35	0.025	7.354	0.833
Trainlm	Satlin	96.88	0.026	7.500	0.867	96.90	0.026	7.500	0.867	96.90	0.027	7.643	0.900	96.86	0.027	7.643	0.900	96.86	0.027	7.643	0.900
Trainrp	Logsig	96.89	0.032	8.320	1.067	96.90	0.032	8.320	1.067	96.90	0.031	8.189	1.033	96.90	0.031	8.189	1.033	96.90	0.031	8.189	1.033
Trainrp	Poslin	95.33	0.033	8.449	1.100	95.35	0.033	8.449	1.100	95.33	0.033	8.449	1.100	95.35	0.031	8.189	1.033	95.33	0.033	8.449	1.100
Trainrp	Tansig	96.98	0.033	8.449	1.100	96.98	0.033	8.449	1.100	96.98	0.033	8.449	1.100	96.98	0.033	8.449	1.100	96.98	0.033	8.449	1.100
Trainrp	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467

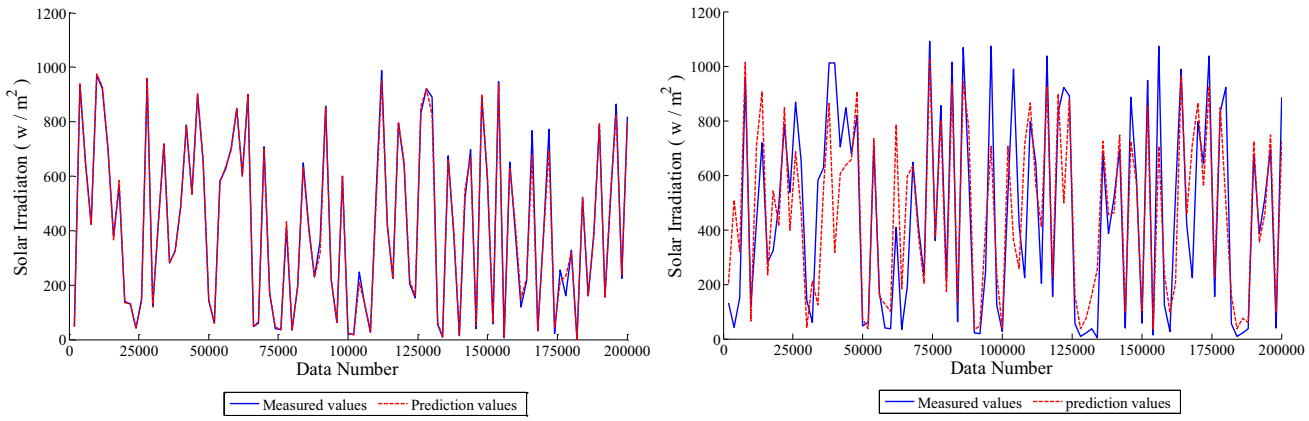
Table 11 (continued)

Training alg	Transfer Func	90-5-5%				80-10-10%				70-15-15%				60-20-20%				50-25-25%			
		R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE	R	MSE	nRMSE	MBE
Trainrp	Radbas	97.04	0.031	8.189	1.033	97.02	0.031	8.189	1.033	97.04	0.030	8.056	1.000	97.08	0.030	8.056	1.000	97.10	0.030	8.056	1.000
Trainrp	Hardlim	53.98	0.124	16.378	4.133	54.02	0.123	16.312	4.100	53.88	0.126	16.510	4.200	53.94	0.125	16.444	4.167	53.80	0.127	16.575	4.233
Trainrp	Tribas	95.82	0.032	8.320	1.067	95.84	0.031	8.189	1.033	95.84	0.030	8.056	1.000	95.86	0.030	8.056	1.000	95.84	0.031	8.189	1.033
Trainrp	Satlin	94.90	0.034	8.576	1.133	94.96	0.032	8.320	1.067	94.96	0.031	8.189	1.033	94.96	0.032	8.320	1.067	94.96	0.034	8.576	1.133
Traingd	Logsig	90.12	0.085	13.560	2.833	90.18	0.086	13.640	2.867	90.24	0.084	13.480	2.800	90.08	0.087	13.719	2.900	89.94	0.090	13.953	3.000
Traingd	Poslin	92.94	0.049	10.296	1.633	92.94	0.049	10.296	1.633	92.96	0.047	10.083	1.567	92.44	0.055	10.908	1.833	92.44	0.055	10.908	1.833
Traingd	Tansig	90.88	0.052	10.606	1.733	90.60	0.055	10.908	1.833	90.44	0.053	10.708	1.767	90.44	0.055	10.908	1.833	90.46	0.054	10.808	1.800
Traingd	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467
Traingd	Radbas	87.12	0.091	14.369	3.033	87.12	0.090	13.953	3.000	87.42	0.092	14.108	3.067	87.98	0.103	14.927	3.433	88.00	0.102	14.855	3.400
Traingd	Hardlim	46.38	0.158	18.488	5.267	45.66	0.167	19.007	5.567	45.10	0.169	19.121	5.633	43.86	0.191	20.327	6.367	42.28	0.199	20.749	6.633
Traingd	Tribas	88.23	0.082	13.319	2.733	87.56	0.087	13.719	2.900	87.52	0.089	13.876	2.967	85.72	0.101	14.782	3.367	85.70	0.102	14.855	3.400
Traingd	Satlin	89.13	0.079	13.073	2.633	89.14	0.080	13.155	2.667	88.02	0.087	13.719	2.900	87.66	0.089	13.876	2.967	87.24	0.092	14.108	3.067
Trainoss	Logsig	95.92	0.040	9.302	1.333	95.74	0.041	9.418	1.367	95.86	0.040	9.302	1.333	95.98	0.040	9.532	1.400	95.84	0.040	9.302	1.333
Trainoss	Poslin	93.82	0.040	9.532	1.400	93.90	0.040	9.302	1.333	93.96	0.038	9.067	1.267	93.82	0.037	9.418	1.367	93.82	0.039	9.302	1.333
Trainoss	Tansig	96.20	0.039	9.302	1.333	96.24	0.039	9.185	1.300	96.26	0.039	9.185	1.300	96.38	0.038	9.302	1.333	96.26	0.038	9.185	1.300
Trainoss	Purelin	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467	62.36	0.134	17.026	4.467
Trainoss	Radbas	96.32	0.039	9.185	1.300	96.36	0.038	9.067	1.267	96.88	0.031	8.189	1.033	96.88	0.031	8.189	1.033	96.30	0.039	9.185	1.300
Trainoss	Hardlim	38.62	0.188	20.167	6.267	50.22	0.125	16.444	4.167	50.10	0.128	16.641	4.267	50.08	0.129	16.705	4.300	50.20	0.130	16.770	4.333
Trainoss	Tribas	95.12	0.039	9.185	1.300	95.88	0.035	8.702	1.167	95.88	0.032	8.320	1.067	95.74	0.032	8.320	1.067	95.04	0.036	8.825	1.200
Trainoss	Satlin	94.98	0.040	9.302	1.333	95.24	0.038	9.067	1.267	96.24	0.031	8.189	1.033	96.04	0.033	8.449	1.100	95.02	0.040	9.302	1.333



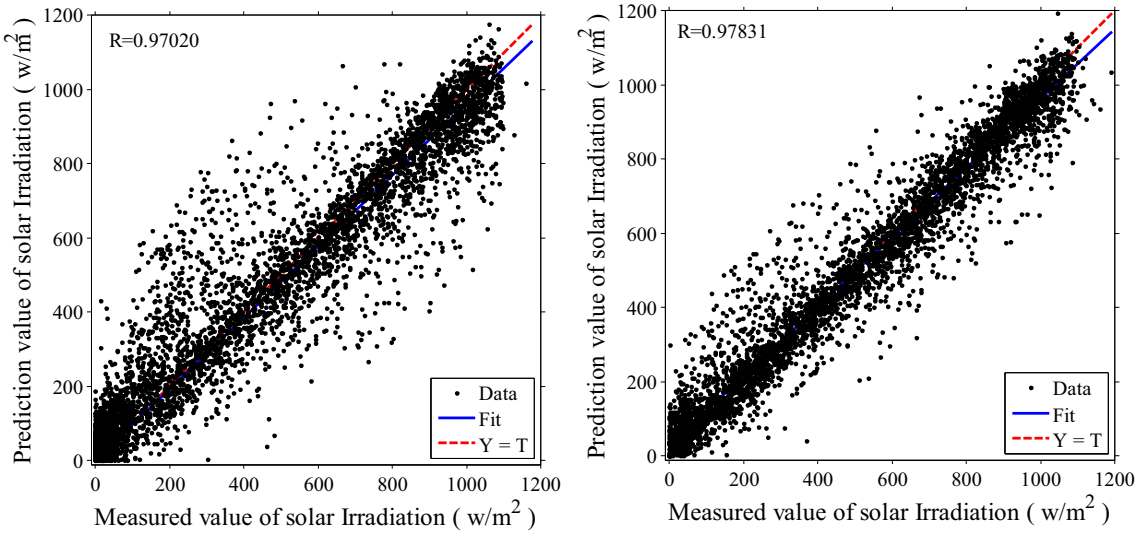
(a) ANN_I_trainlm_logsig **(b)** ANN_I_trainrp_hardlim

Fig. 6 Comparison of forecast and measurement data based on time series of the most successful and worst situation in ANN_I



(a) ANN_II_trainlm_tansig **(b)** ANN_II_traingd_purelin

Fig. 7 Comparison of forecast and measurement data based on time series of the most successful and worst situation in ANN_II



(a) **(b)** ANN_II

Fig. 8 Solar irradiation predicted by ANN_I and ANN_II using the "trainlm" training algorithm

Table 12 Simulation results for R^2 and MAPE

Model name	R^2	MAPE
ANN_I (LM_logsig)	0.9412	7.82
ANN_II (LM_tansig)	0.9570	6.79

the developed models. Figure 8a, b illustrates the variations in predicted solar irradiation by the most successful ANN models.

In the charts in Fig. 8, the vertical axis is the prediction values and the horizontal axis is the actual values. According to the simulation results, the correlation coefficient (R) for the prediction values is shown in the graphs. $R = 1$ means that there is a definite linear relationship between the measured value and the actual value. As can be seen in Fig. 8a, most of the point for the most successful function in the ANN_I model, "logsig," is scattered along the diagonal line. In Fig. 8b, the points for the "tansig" transfer function in the ANN_II model are located more collectively and regularly on the diagonal line. This is an indication that the forecast results are successful. In addition, the results revealed that the prediction values had high precision. The determination coefficient (R^2) represents the percentage of data closest to the best line of compliance of the overall data set for a given station. Within the scope of this study, the R^2 and MAPE results calculated for the LM algorithm ANN_I and ANN_II that give the best results in the prediction of instantaneous solar irradiation using the parameters recorded by the weather station are shown in Table 12.

The R^2 results clearly show that the predicted instantaneous solar irradiation values are very close to the values measured by the selected ANN_I and ANN_II model. MAPE results are highly accurate as they perform below 10%. Therefore, from the statistical error analysis shown in Tables 10, 11, and 12, it can be concluded that the ANN models created with the LM training algorithm and "tansig" and "logsig" transfer functions performed well compared to the data set sample used in this study. Furthermore, the current study confirms the ability of the identified ANN models to accurately predict solar irradiation values for all regions with similar meteorological data located in and around Hakkari.

In this study, solar irradiation estimation studies were carried out using ANN, and results were obtained depending on the transfer functions affecting the prediction result and the network structure. In the study, the most commonly used training algorithms and transfer functions in the literature have been employed. During the training phase, solar irradiation prediction results were obtained using an ANN model with randomly selected test data from the total data set, which was different from the data set used for training. In addition, many tests were performed on several ANN-based network structures created. The simulation results show that highly accurate predictions can be obtained in the ANN's solar irradiation prediction. It has also been observed that the percentage of the estimate changes in the ANN when the test data and training data change. Tests were carried out on two different training algorithms with looping on the eight different transfer functions, and it was tried to determine the algorithm and function type that would give the best response to the system under similar conditions. Although these cycles for the study yielded prediction values, the algorithm that gave the best results was "trainlm" and the best transfer function was "logsig" in networks with small neurons and "tansig" in networks with large numbers of neurons. Deviation values depending on the number of repetitions as the results of the comparison are shown in Fig. 9.

4 Conclusion

This study creates a neural network model of different network structures to predict solar irradiation from a multivariate time series data set consisting of meteorological data for the Hakkari Province of Turkey. The comparison of these neural network models was then evaluated by a set of performance metrics.

This assessments provides a better view of the training and transfer functions in the network structure in predicting solar irradiation, as well as the contribution of neuron counts in the hidden layer to the performance. Identifying conditions that affect forecast accuracy helps in many applications where solar irradiation is the main factor. In the study, the ANN_II network structure of 100 neurons created using the "trainlm" algorithm and the "tansig" transfer function performed best. In addition, according to the results obtained in the network structures created by using regional meteorological data, "logsig" and "tansig" transfer functions, especially with "trainlm" training algorithm, produced a prediction success rate of approximately 97%. The same transfer functions showed successful results in the "trainscg" and "trainoss" training algorithms. The performance criteria used to evaluate the models were found satisfactory considering the size of the data set. In the analyses made with the "traingd" and "trainrp" training algorithms, the prediction success in all transfer functions showed worse results. Therefore, to use both training functions in solar irradiation forecasting using meteorological parameters is not recommended. A large, more diverse data set with more parameters in different climatic conditions can be used to create models that will help to provide even better results in the future. It can also be a guide for monitoring and optimizing the results of solar power plants that we will benefit more from in the future.

Finally, it is not yet possible to empirically measure the total solar irradiation at all locations on the Earth, and it also

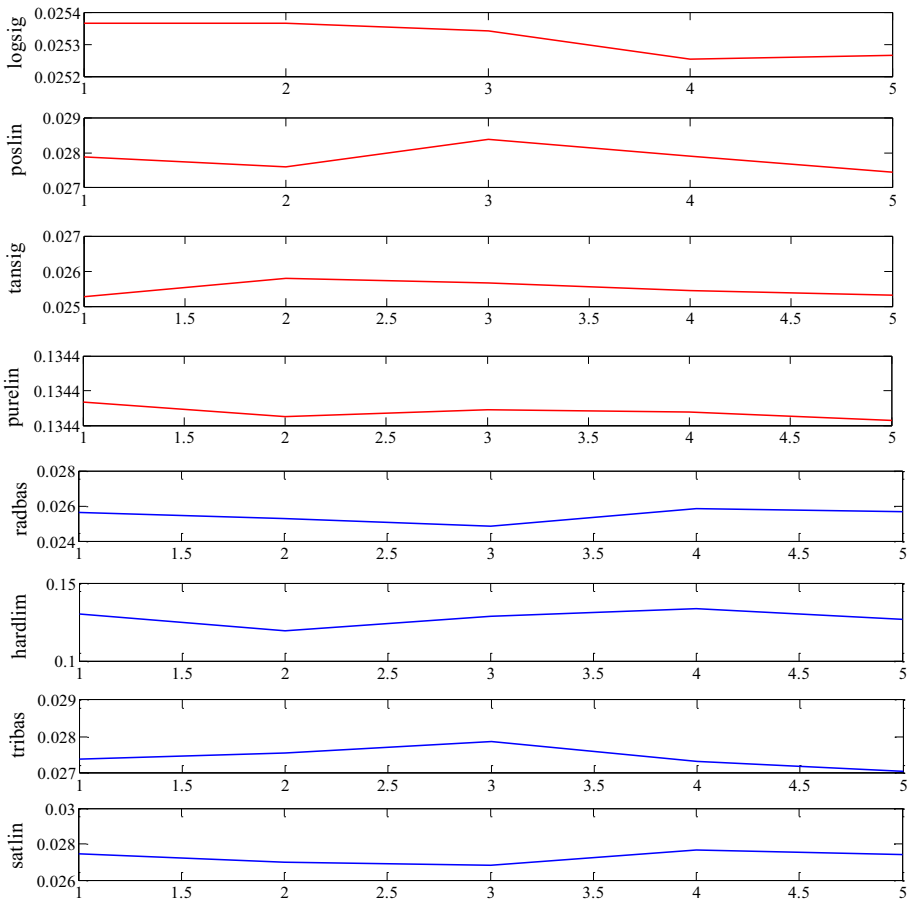


Fig. 9 Deviation values depending on the number of repetitions for functions used in the ANN_II model

requires specialized and expensive equipment with advanced systems. Furthermore, developing the neural network models will help the prediction of the solar irradiation accurately since that process. For this reason, the ANN-based solar irradiation estimation models are recommended to replace with the experimental measurements and methods based on empirical relationships.

Author contributions Credit Author Statement: AB Formal Analysis, Supervision, Resources, Writing—Review and Editing. EOY Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Writing—Original Draft, Visualization.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Vakili M, Sabbagh-Yazdi SR, 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. J Clean Prod 141:1275–1285. https://doi.org/10.1016/j.jclepro.2016.09.145
2. Kurniawan A, Shintaku E (2021) Two-step artificial neural network to estimate the solar radiation at Java Island. Int J Electric Comput Eng (2088-8708). https://doi.org/10.11591/ijece.v11i4.pp3559-3566
3. Ben Othman A, Belkilani K, Besbes M (2020) Prediction improvement of potential PV production pattern, imagery satellite-based. Sci Rep 10(1):1–10. https://doi.org/10.1038/s41598-020-76957-8
4. Ağbulut Ü, Gürel AE, Biçen Y (2021) Prediction of daily global solar radiation using different machine learning algorithms: evaluation and comparison. Renew Sustain Energy Rev 135:110114. https://doi.org/10.1016/j.rser.2020.110114
5. 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. Renew Sustain Energy Rev 91:248–258. https://doi.org/10.1016/j.rser.2018.03.096
6. Kashani MH, Inyurt S, Golabi MR, AmirRahmani M, Band SS (2022) Estimation of solar radiation by joint application of phase space reconstruction and a hybrid neural network model. Theor Appl Climatol. https://doi.org/10.1007/s00704-021-03913-5
7. Woldegiyorgis TA, Admasu A, Benti NE, Asfaw AA (2022) A comparative evaluation of artificial neural network and sunshine based models in prediction of daily global solar radiation of lalibela. Ethiop Cogent Eng 9(1):1996871. https://doi.org/10.1080/23311916.2021.1996871
8. Calcabrini A, Ziar H, Isabella O, Zeman M (2019) A simplified skyline-based method for estimating the annual solar energy potential in urban environments. Nat Energy 4(3):206–215. https://doi.org/10.1038/s41560-018-0318-6

9. Huang X, Li Q, Tai Y, Chen Z, Zhang J, Shi J, Liu W et al (2021) Hybrid deep neural model for hourly solar irradiance forecasting. Renew Energy 171:1041–1060. https://doi.org/10.1016/j.renene.2021.02.161
10. Gupta A, Gupta K, Saroha S (2022) Solar energy radiation forecasting method. Smart technologies for energy and environmental sustainability. Springer, Cham, pp 105–129. https://doi.org/10.1007/978-3-030-80702-3_7
11. Das UK et al (2018) Forecasting of photovoltaic power generation and model optimization: a review. Renew Sustain Energy Rev 81:912–928. https://doi.org/10.1016/j.rser.2017.08.017
12. Molina A, Falvey M, Rondanelli R (2017) A solar radiation database for Chile. Sci Rep 7(1):1–11. https://doi.org/10.1038/s41598-017-13761-x
13. Guermoui M, Benkaciali S, Gairaa K, Bouchouicha K, Boulmaiz T, Boland JW (2022) A novel ensemble learning approach for hourly global solar radiation forecasting. Neural Comput Appl. https://doi.org/10.1007/s00521-021-06421-9
14. Voyant C, Notton G, Kalogirou S, Nivet ML, Paoli C, Motte F, Fouilloy A (2017) Machine learning methods for solar radiation forecasting: a review. Renew Energy 105:569–582. https://doi.org/10.1016/j.renene.2016.12.095
15. Bhatt A, Ongsakul W, Singh JG (2022) Sliding window approach with first-order differencing for very short-term solar irradiance forecasting using deep learning models. Sustain Energy Technol Assess 50:101864. https://doi.org/10.1016/j.seta.2021.101864
16. Guermoui M, Bouchouicha K, Benkaciali S, Gairaa K, Bailek N (2022) New soft computing model for multi-hours forecasting of global solar radiation. Eur Phys J Plus 137(1):162. https://doi.org/10.1140/epjp/s13360-021-02263-5
17. Ali-Ou-Salah H, Oukarfi B, Bahani K, Moujabbir M (2021) A new hybrid model for hourly solar radiation forecasting using daily classification technique and machine learning algorithms. Math Probl Eng. https://doi.org/10.1155/2021/6692626
18. Kurniawan A, Shintaku E (2020) Estimation of the monthly global, direct, and diffuse solar radiation in Japan using artificial neural network. Int J Mach Learn Comput 10(1):253–258. https://doi.org/10.18178/ijmlc.2020.10.2.928
19. Wang F, Mi Z, Su S, Zhao H (2012) Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. Energies 5(5):1355–1370. https://doi.org/10.3390/en5051355
20. Linares-Rodríguez A, Ruiz-Arias JA, Pozo-Vázquez D, Tovar-Pescador J (2011) Generation of synthetic daily global solar radiation data based on ERA-Interim reanalysis and artificial neural networks. Energy 36(8):5356–5365. https://doi.org/10.1016/j.energy.2011.06.044
21. Mubiru J, Banda EJKB (2008) Estimation of monthly average daily global solar irradiation using artificial neural networks. Sol Energy 82(2):181–187. https://doi.org/10.1016/j.solener.2007.06.003
22. Bounoua Z, Mechaqrane A (2022) Estimation of daily direct normal solar irradiation using machine-learning methods. WITS 2020. Springer, Singapore, pp 627–638. https://doi.org/10.1007/978-981-33-6893-4_58
23. Zeng J, Qiao W (2011). Short-term solar power prediction using an RBF neural network. In: 2011 IEEE power and energy society general meeting. IEEE, pp 1–8. https://doi.org/10.1109/PES.2011.6039204
24. Mousavi SM, Mostafavi ES, Jaafari A, Jaafari A, Hosseinpour F (2015) Using measured daily meteorological parameters to predict daily solar radiation. Measurement 76:148–155. https://doi.org/10.1016/j.measurement.2015.08.004
25. El Alani O, Ghennioui H, Ghennioui A (2019) Short term solar irradiance forecasting using artificial neural network for a semi-arid climate in Morocco. In: 2019 international conference on wireless networks and mobile communications (WINCOM). IEEE, pp. 1–7. https://doi.org/10.1109/WINCOM47513.2019.8942412
26. Gutierrez-Corea FV, Manso-Callejo MA, Moreno-Regidor MP, Manrique-Sancho MT (2016) Forecasting short-term solar irradiance based on artificial neural networks and data from neighboring meteorological stations. Sol Energy 134:119–131. https://doi.org/10.1016/j.solener.2016.04.020
27. Bosch JL, Lopez G, Batlles FJ (2008) Daily solar irradiation estimation over a mountainous area using artificial neural networks. Renew Energy 33(7):1622–1628. https://doi.org/10.1016/j.renene.2007.09.012
28. Mellit A, Pavan AM (2010) A 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at Trieste, Italy. Sol Energy 84(5):807–821. https://doi.org/10.1016/j.solener.2010.02.006
29. Voyant C, Muselli M, Paoli C, Nivet ML (2011) Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation. Energy 36(1):348–359. https://doi.org/10.1016/j.energy.2010.10.032
30. Mohandes M, Rehman S, Halawani TO (1998) Estimation of global solar radiation using artificial neural networks. Renew Energy 14(1–4):179–184. https://doi.org/10.1016/S0960-1481(98)00065-2
31. Yadav AK, Chandel SS (2012) Artificial neural network based prediction of solar radiation for Indian stations. Int J Comput Appl 50(9)
32. Quej VH, Almorox J, Arnaldo JA, Saito L (2017) ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. J Atmos Solar Terr Phys 155:62–70. https://doi.org/10.1016/j.jastp.2017.02.002
33. Marzo A, Trigo-Gonzalez M, Alonso-Montesinos J, Martínez-Durbán M, López G, Ferrada P, Batlles FJ (2017) Daily global solar radiation estimation in desert areas using daily extreme temperatures and extraterrestrial radiation. Renew Energy 113:303–311. https://doi.org/10.1016/j.renene.2017.01.061
34. Halabi LM, Mekhilef S, Hossain M (2018) Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation. Appl Energy 213:247–261. https://doi.org/10.1016/j.apenergy.2018.01.035
35. 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. J Clean Prod 209:1105–1118. https://doi.org/10.1016/j.jclepro.2018.10.254
36. Jahani B, Mohammadi B (2019) A comparison between the application of empirical and ANN methods for estimation of daily global solar radiation in Iran. Theoret Appl Climatol 137(1):1257–1269. https://doi.org/10.1007/s00704-018-2666-3
37. Antonopoulos VZ, Papamichail DM, Aschonitis VG, Antonopoulos AV (2019) Solar radiation estimation methods using ANN and empirical models. Comput Electron Agric 160:160–167. https://doi.org/10.1016/j.compag.2019.03.022
38. Guermoui M, Gairaa K, Ferkous K, Santos DSDO Jr, Arrif T, Belaid A (2023) Potential assessment of the TVF-EMD algorithm in forecasting hourly global solar radiation: review and case studies. J Clean Prod 385:135680. https://doi.org/10.1016/j.jclepro.2022.135680
39. Demir V, Citakoglu H (2023) Forecasting of solar radiation using different machine learning approaches. Neural Comput Appl 35(1):887–906. https://doi.org/10.1007/s00521-022-07841-x
40. Sözen A, Arcaklioğlu E, Özalp M (2004) Estimation of solar potential in Turkey by artificial neural networks using meteorological and geographical data. Energy Convers Manag 45(18–19):3033–3052. https://doi.org/10.1016/j.enconman.2003.12.020

41. Ozgoren M, Bilgili M, Sahin B (2012) Estimation of global solar radiation using ANN over Turkey. Expert Syst Appl 39(5):5043–5051. https://doi.org/10.1016/j.eswa.2011.11.036
42. Kaba K, Sarıgül M, Avcı M, Kandırmaz HM (2018) Estimation of daily global solar radiation using deep learning model. Energy 162:126–135. https://doi.org/10.1016/j.energy.2018.07.202
43. Yildirim A, Bilgili M, Ozbek A (2023) One-hour-ahead solar radiation forecasting by MLP, LSTM, and ANFIS approaches. Meteorol Atmos Phys 135(1):10. https://doi.org/10.1007/s00703-022-00946-x
44. Yıldırım HB, Çelik Ö, Teke A, Barutçu B (2018) Estimating daily global solar radiation with graphical user interface in eastern Mediterranean region of Turkey. Renew Sustain Energy Rev 82:1528–1537. https://doi.org/10.1016/j.rser.2017.06.030
45. Bilgili M, Ozgoren M (2011) Daily total global solar radiation modeling from several meteorological data. Meteorol Atmos Phys 112(3–4):125–138. https://doi.org/10.1007/s00703-011-0137-9
46. Alizamir M, Kim S, Kisi O, Zounemat-Kermani M (2020) A comparative study of several machine learning based non-linear regression methods in estimating solar radiation: case studies of the USA and Turkey regions. Energy 197:117239. https://doi.org/10.1016/j.energy.2020.117239
47. Şenkal O (2010) Modeling of solar radiation using remote sensing and artificial neural network in Turkey. Energy 35(12):4795–4801. https://doi.org/10.1016/j.energy.2010.09.009
48. Karaman ÖA, Ağır TT, Arsel İ (2021) Estimation of solar radiation using modern methods. Alex Eng J 60(2):2447–2455. https://doi.org/10.1016/j.aej.2020.12.048
49. Qazi A, Fayaz H, Wadi A, Raj RG, Rahim NA, Khan WA (2015) The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review. J Clean Prod 104:1–12. https://doi.org/10.1016/j.jclepro.2015.04.041
50. El Mghouchi Y, Chham E, Zemmouri EM, El Bouardi A (2019) Assessment of different combinations of meteorological parameters for predicting daily global solar radiation using artificial neural networks. Build Environ 149:607–622. https://doi.org/10.1016/j.buildenv.2018.12.055
51. Premalatha N, Valan Arasu A (2016) Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. J Appl Res Technol 14(3):206–214. https://doi.org/10.1016/j.jart.2016.05.001
52. Faisal AF, Rahman A, Habib MTM, Siddique AH, Hasan M, Khan MM (2022) Neural networks based multivariate time series forecasting of solar radiation using meteorological data of different cities of Bangladesh. Results Eng 13:100365. https://doi.org/10.1016/j.rineng.2022.100365
53. Li DH, Chen W, Li S, Lou S (2019) Estimation of hourly global solar radiation using multivariate adaptive regression spline (MARS)–a case study of Hong Kong. Energy 186:115857. https://doi.org/10.1016/j.energy.2019.115857