Providing an accurate global model for monthly solar radiation forecasting using artificial intelligence based on air quality index and meteorological data of different cities worldwide

Shirin Riahi[1] · Elham Abedini[2] · Masoud Vakili[3] · Mobina Riahi[4]

Abstract

This study aims to present an exact model for predicting solar radiation worldwide through a general model. In this study, mean monthly global solar radiation would have been predicted by applying artificial intelligence methods including artificial neural network, adaptive neuro-fuzzy inference system and hybrid genetic algorithm for different cities worldwide. Investigating different models under various situations showed that the adaptive neuro-fuzzy inference system created the most accurate and precise model for predicting solar radiation. Statistics indexes, such as the determination coefficient, mean absolute percentage error, root mean square error and mean bias error, for the best model selected are 0.999, 5.50E−04, 5.90E−05 and 0.425, respectively. It can be claimed that according to the amount of the statistical indexes, which was mentioned above, the provided model has approximately more formidable accuracy and credibility in comparison with other models, which other researchers did.

Introduction

Environmental pollution has become a serious modern-day problem which mainly causes by burning fossil fuels. As a result, scientists and researchers are looking for appropriate alternatives which are more environmentally friendly. Renewable energy sources like wind, solar, thermal and biomass energies seem to be suitable solutions. It is widely believed that in comparison with other green energy resources, solar energy is the most suitable option instead of fossil fuels as it is not only sustainable but also available and abundant almost everywhere. It is predicted that solar energy will be widely utilised shortly for fulfilling the world's energy demands Khanlari et al. (2020a, b).

One of the essential parameters in optimising and designing solar radiation types is solar radiation, as it plays a crucial role in the initiation and installation of panels and solar collectors. Therefore, predicting and measuring this parameter in the desired locations should be prioritised (Bakirci 2009; Vakili et al. 2016). However, it is pretty challenging to access solar radiation in different locations because the number of meteorological stations is limited due to the maintenance cost of instruments, calibration requirements and necessary facilities (Khatib et al. 2012). For example, in Turkey, 129 out of 1798 stations can record solar radiation data until 2020, and in China, there were 756 stations until 2012 which only 122 of them could record this data (Zang et al. 2012). Therefore, different modeling methods have been introduced for anticipating solar radiation in recent decades, which are divided into three general categories,

✉ Masoud Vakili
Msd.vakili@gmail.com

1   Department of Physics, Shahid Beheshti University, Evin, Tehran, Iran

2   Department of Photonics, Laser and Plasma Research Institute, Shahid Beheshti University, Evin, Tehran, Iran

3   Graduate of Department of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran

4   Department of Computer Engineering, Sirjan University of Technology, Kerman, Iran

namely empirical models (Ustun et al. 2020), regression models (Guermoui et al. 2020) and modeling by artificial intelligence (AI) (Üstün et al. 2020).

Empirical models are based on mathematical formulas. Although these models are accepted as accessible and valuable techniques for calculating and predicting monthly average global solar radiation, they fail to anticipate short-term solar radiation data due to quick changes in weather conditions (Ağbulut et al. 2021). Moreover, only a handful of parameters can be used to predict this technique, like minimum and maximum of daily temperature, sunny hours and the amount of cloudiness. At the same time, there is an intimate relationship between solar radiation and the rate of humidity. However, the complicated and nonlinear relations of both independent and dependent variables, particularly in a humid climate, cannot be explained by these models (Kisi and Parmar 2016; Fan et al. 2018a). Furthermore, constant coefficients exist in empirical equations whose amount depends on the study area and experience (Yadav and Chandel 2014). Thus, empirical models cannot be considered a functional approach in many situations.

Recently, artificial intelligence (AI) models have attracted considerable attention in many fields. Different AI approaches are being used to anticipate solar radiation data such as wavelet transform (Zhang et al. 2020), support vector machine (SVM) (Hou et al. 2018), extreme learning machine (Bouzgou and Gueymard 2017), artificial neural network (ANN) (El Mghouchi et al. 2019), kernel nearest neighbour (k-NN) (Ağbulut et al. 2021), adaptive neuro-fuzzy inference system (ANFIS) (Zou et al. 2017), multilayer perceptron (MLP) (Guermoui et al. 2018), fuzzy logic (Baser and Demirhan 2017) and Gaussian process regression (GPR) (Zhou et al. 2020). ANN does not cause any restriction in choosing parameters for the prediction and estimation of solar radiation. Recent studies have shown that anticipating solar radiation AI algorithms is more accurate and precise than empirical techniques (Liu et al. 2020). So ANNs have done a great job modeling nonlinear, complex and time-varying input-output systems (Gani et al. 2016).

A regression model is another method for predicting solar radiation data. Comparing this technique with the artificial neural network has shown that the result derived from ANN modeling is more precise and accurate as the principal absolute percentage error amount of regression modeling results are higher than that of the ANN model. Furthermore, the values of R^2 for regression models are lower than that of obtaining from ANN. This result evidences that ANN models are more effective ways for prediction (Kumar et al. 2015).

Nowadays, the statistical approach, physical model, hybrid method and artificial intelligence are commonly used for predicting solar irradiance (Wang et al. 2020). However, this study mainly was focused on AI techniques. In hybrid models, different methods are combined in order to gain benefit from every single anticipating model. For example, for predicting hourly solar radiation, a hybrid method based on fuzzy logic (FL), genetic algorithm (GA) and artificial neural network can be used (Garud et al. 2020). As mentioned before, models based on artificial intelligence play a crucial role in anticipating solar radiation. Table 1 indicates a summary of other researchers investigating solar radiation prediction in different locations with different methods.

Recently, the artificial neural network has drawn scientists' attention more than other methods to predict solar radiation. Some results of several works of literature are mentioned as follows. Yıldırım et al. (2018) predicted monthly global solar radiation of four different Turkey stations using ANN and regression analysis. For training the algorithm, longitude, relative humidity, maximum possible sunshine duration, temperature extraterrestrial solar radiation, sunshine duration and the year were used. In this study, ANN models presented the best outcome with 0.14 and 0.961 of RMSE and R^2, respectively. Khosravi et al. (2018) employed some developed methods namely, multilayer feed-forward neural network (MLFFNN), group method of data handling (GMDH), adaptive neuro-fuzzy inference system (ANFIS), ANFIS optimised with genetic algorithm (ANFIS-GA), ANFIS optimised with particle swarm optimisation algorithm (ANFIS-PSO) and ANFIS optimised with ant colony (ANFIS-ACO) to predict daily global solar radiation of Iran. Comparing these models, GMDH has the best result as R^2, RMSE and MSE for this model were 0.9886, 0.2466 and 0.0608, respectively. Meenal and Selvakumar (2018) predicted daily global solar radiation using three empirical, ANN and SVM models. More than 0.99 correlation was seen in the SVM model. In another study, Kaba et al. (2018) used a deep learning algorithm in different Turkey stations to predict monthly average daily global radiation. Some parameters were used to train the DL algorithm like cloud cover, extraterrestrial solar radiation, minimum and maximum temperatures and sunshine duration. The best result of R^2 was 0.98 in this study. Jahani and Mohammadi (2019) applied ANN, the combination of ANN with genetic algorithm and empirical models, to forecast daily global solar radiation in Iran. However, the hybrid model outperformed the other approaches with MBE, R^2 and RMSE at 38.4, 0.92 and 185 J/cm^2/day, respectively. Marzouq et al. (2019) anticipated daily global solar radiation by ANN, k-NN and empirical models. In this study, the value of R^2 was equal to 0.96 and 0.97 for k-NN and a proposed hybrid model (k-NN and ANN), respectively. Rabehi et al. (2020) used boosted decision tree and multilayer perceptron (MLP) and a hybrid linear regression with these two models for anticipating daily global solar radiation in the south of Algeria. R^2 and root mean square error were accounted for 0.977 and 0.033, respectively, in the MLP model, which shows the highest accuracy compared to other technique. Belmahdi et al. (2020) anticipated daily global

Table 1 List of literature studies on forecasting solar radiation that used machine learning methods

References	Statistics index	Methods	Time scale	Locations
Behrang et al. (2010)	R^2, MSE, MAPE, SSE	Empirical, RBF, MLP	Daily	Dezful, Iran
Moreno et al. (2011)	R^2, MSE, RMSE, MAE	Bristow-Campbell, ANN, kNN	Daily	40 sites in Spain
Asl et al. (2011)	MAPE	MLP	Daily	Dezful, Iran
Mohandes (2012)	MAPE	ANN-PSO, ANN-BP	Monthly	41 sites in Saudi Arabia
Sumithira and Kumar (2012)	MBE, R^2, RMSE	ANN, ANFIS	Monthly	Tamil Nadu, India
Chen et al. (2013)	RMSE, NMRSE	SVM	Daily	3 sites in China
Salcedo-Sanz et al. (2014)	Bias, RMSE, MAE	Basic ELM, SVR, ELM-CRO	Daily	Murcia, Spain
Huang et al. (2014)	MeRE, MAE, r	Persistence, analog, RF, MLR	Daily	4 sites in Australia
Ramli et al. (2015)	RMSE, MRE, CC	SVM, ANN	Daily	2 sites in Saudi Arabia
Mohammadi et al. (2015a)	RMSE, NRMSE, MABE, MAPE, RPE, R^2	Empirical, SVR	Monthly	Isfahan, Iran
Olatomiwa et al. (2015)	RMSE, MAPE, R^2, r	ANN, GP, SVM-FFA	Monthly	3 sites in Nigeria
Mohammadi et al. (2015b)	R^2, MABE, NRMSE, MAPE, RMSE	ARMA, ANN, SVM-WT, GP	Daily	Bandar Abass, Iran
Gairaa et al. (2016)	MPE	ARMA, ANN	Daily	2 sites in Algeria
Belaid and Mellit (2016)	NRMSE, MBE, RMSE, MAPE, R	SVM, MLP	Daily/monthly	Ghardaia, Algeria
Alsina et al. (2016)	MAPE	ANN	Monthly	45 sites in Italy
Mehdizadeh et al. (2016)	R^2	ANFIS, GEP, ANN	Daily	Kerman, Iran
Wang et al. (2016)	RMSE, MAE, R^2	MLP, RBF, GRNN	Daily	12 sites in China
Quej et al. (2017)	R^2, RMSE, MAE	ANN, ANFIS, SVM	Daily	6 sites in Mexico
Deo and Şahin (2017)	RMSE, MAE, r	ANN, MLP, ARIMA	Monthly	Australia
Marzo et al. (2017)	r, rRMSE, rMBE	ANN	Daily	13 sites in six countries
Zou et al. (2017)	MAE, RMSE, R^2	ANFIS, E-IBCM, IYHM	Daily	3 sites in Hunan, China
Keshtegar et al. (2018)	MAE, NSE, RMSE, d, MBE	M5Tree, MARS, RSM, Kriging	Monthly	2 sites in Turkey
Rohani et al. (2018)	R^2, RMSE, E F, MAPE,	GPR	Daily/monthly	Mashhad, Iran
Fan et al. (2018b)	RMSE, MBE, MAE, R^2	SVM, XGBoost, Empirical	Daily	3 sites in China
Lotfinejad et al. (2018)	MPE, R^2	BNN, ANFIS, GRNN	Daily	4 sites in Iran
Qin et al. (2018)	RMSE, R, MAE, R^2	ANN, YHM, EPP, HSRM	Daily	837 sites in China
Zhang et al. (2019)	R^2, RMSE, MAE, rRMSE, NS	BP, PSO-BP, six statistical models	Daily	Northwest China
Rocha et al. (2019b)	MAPE, nRMSE	ANN	Daily, weekly and monthly	Northeast Brazil
Ghimire et al. (2019)	RMSE, MAE, NS, WI, MBE	ANN, SVR, GPML, GA, ARIMA, TM, TSFS	Daily	5 sites in Australia
Bamehr and Sabetghadam (2021)	MAE, RMSE, MAPE, R^2	ANN	Daily	Mashhad, Iran
Jumin et al. (2021)	R^2, MAE, RMSE, RAE, RSE	BDTR	Monthly	Malaysia

solar radiation in three cities in Morocco using different techniques, namely ANN, ARIMA and ANN-ARIMA. In this study, the statistical errors that validate the models' accuracy include MBE, NRMSE, MAPE, RMSE, TS, SD and R^2. However, the hybrid model presented a better result as the value of R^2 was 0.988. In another study, Dhakal et al. (2020) predicted DGSR in Nepal using ANN, five different machine learning methods and temperature-based empirical models. Among ANN models, ANN3 performed better than others

with $R^2 = 0.8446$ and RMSE = 1.4595, while stepwise linear regression outperformed other ML approaches. These models recommended for the locations where max and min ambient temperatures, sunshine duration, relative humidity and precipitation data are available.

In the current study, artificial neural network (ANN), hybrid genetic algorithm and ANN and adaptive neuro-fuzzy inference system (ANFIS) have been applied to estimate the daily global solar radiation of ten different cities worldwide.

The input variable includes latitude, longitude, minimum and maximum temperatures (°C), relative humidity (%), wind speed (m/s), surface pressure (kPa), amount of air pollutants (O_3, NO_2, $PM_{2.5}$, PM_{10}), dew frost point, wet bulb temperature (°C) and mean solar radiation ($MJ/m^2/day$) on a horizontal surface. This research aims to investigate the impact of various input parameters, particularly hazardous air pollutants and dew frost point wet bulb temperature, on the amount of solar radiation in different locations of the world. Therefore, an attempt has been made to introduce precise and accurate and global models that consider many useful parameters on solar radiation. According to the authors' knowledge, no unique model has estimated the mean monthly global solar radiation for any cities worldwide with this accuracy.

Case study regions and data

In this study, ten cities from all over the globe are considered to present a general model. These well-known metropolises are from different locations with different climate and rate of pollution. As shown in Fig. 1, these cities consist of Beijing, Delhi, Tehran and Tokyo selected from Asia, Hamburg, London from Europe, Johannesburg from Africa, Los Angeles from North America, Rio de Janeiro from South America and Melbourne from Oceania. The monthly climatic data of these cities, namely, latitude, longitude, min and max temperatures (°C), relative humidity (%), wind speed (m/s), surface pressure (kPa), amount of air pollutants (O_3, NO_2, $PM_{2.5}$, PM_{10}), dew frost point, wet bulb temperature(°C) and mean solar radiation ($MJ/m^2/day$) on a horizontal surface, were considered as modeling inputs between 2018 and 2019. Table 2 shows some essential geographical details of the cities along with the average data for 2 years.

Input parameters for the neural network were measured experimentally and used for training, validation and testing for 730 days in 2018 and 2019. The results are shown in Fig. 2 concerning the maximum temperature, in Fig. 3 concerning the minimum temperature, in Fig. 4 concerning the dew point of the environment, in Fig. 5 concerning the wet bulb temperatures, in Fig. 6 concerning the relative humidity, in Fig. 7 concerning the surface pressure, in Fig. 8 concerning the amount of NO_2, in Fig. 9 concerning the amount of O_3, in Fig. 10 concerning the amount of PM_{10}, in Fig. 11 concerning the amount of $PM_{2.5}$, in Fig. 12 concerning the wind speed and in Fig. 13 concerning the mean solar radiation.

Modeling

These days, modeling methods that include mathematical relationships and conceptual models are considered appropriate tools for anticipating natural phenomena. Because of complicated relationships between components and too many natural process variables, modeling techniques can be pretty challenging for the users (Vakili et al. 2017b). In this paper, artificial neural network, genetic algorithm and adaptive neuro-fuzzy inference system (ANFIS) have been implemented as modeling methods briefly discussed in the following.
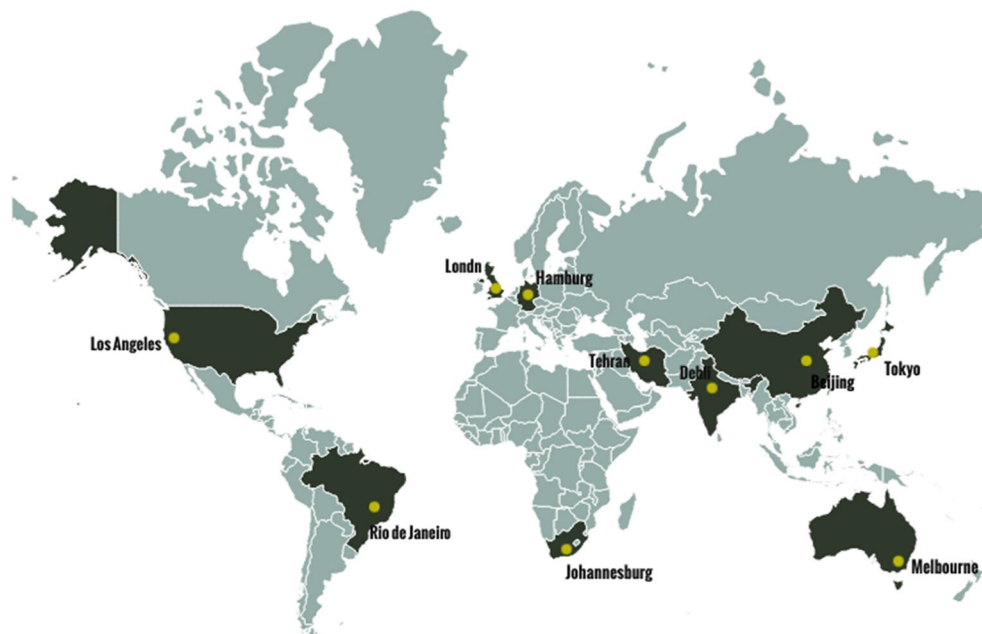
Fig. 1 Location of the cities on the map

Table 2 The average input data of the sites

Station	City/country	Latitude	Longitude	Elevation (m)	Max. Temp.	Min. Temp.	Wet bulb temp.	Dew point (C)	Relative humidity (%)	Surface pressure (kPa)	Wind speed (m/s)	PM$_{2.5}$	PM^{10}	O$_3$	NO$_2$	Mean solar radiation (MJ/m^2/day)
1	Beijing, China	39.90° N	116.41° E	44.00	22.01	9.26	3.93	3.74	49.28	100.20	1.90	105.50	52.38	45.39	17.94	13.69
2	Delhi, India	28.70° N	77.10° E	216.00	31.57	20.05	15.48	15.63	56.59	98.35	1.75	169.47	141.98	10.81	22.46	15.38
3	Hamburg, German	53.55° N	9.99° E	8.00	15.46	8.28	8.73	8.79	82.73	101.00	3.23	35.25	14.61	21.41	9.60	8.69
4	Johannesburg, South Africa	26.20° S	28.05° E	1753.00	27.03	10.40	3.49	3.53	41.59	84.87	2.54	74.83	36.54	15.77	9.00	20.73
5	London, England	51.51° N	0.13° W	11.00	16.03	7.86	8.86	8.90	83.33	100.35	3.23	46.24	20.86	21.38	27.86	9.62
6	Los Angeles, USA	34.05° N	118.25° W	87.00	24.45	12.31	5.67	6.38	52.83	94.89	1.94	47.69	27.35	30.48	16.38	18.84
7	Melbourne, Australia	37.81° S	144.96° E	31.00	18.76	7.93	6.81	6.93	68.97	99.46	3.08	28.19	16.88	18.42	9.04	16.98
8	Rio de Janeiro, Brazil	22.91° S	43.17° W	2.13	28.06	19.01	18.14	18.27	75.78	98.69	1.87	54.26	27.59	30.89	8.37	16.43
9	Tehran, Iran	35.69° N	51.39° E	1189.00	18.37	7.41	1.50	1.64	52.89	82.25	1.53	83.81	58.73	39.85	76.79	16.87
10	Tokyo, Japan	35.68° N	139.65° E	40.00	22.41	13.80	14.00	14.13	79.73	100.07	2.17	34.91	13.79	35.75	9.73	12.15

Fig. 2 The maximum temperature of each month between 2018 and 2019
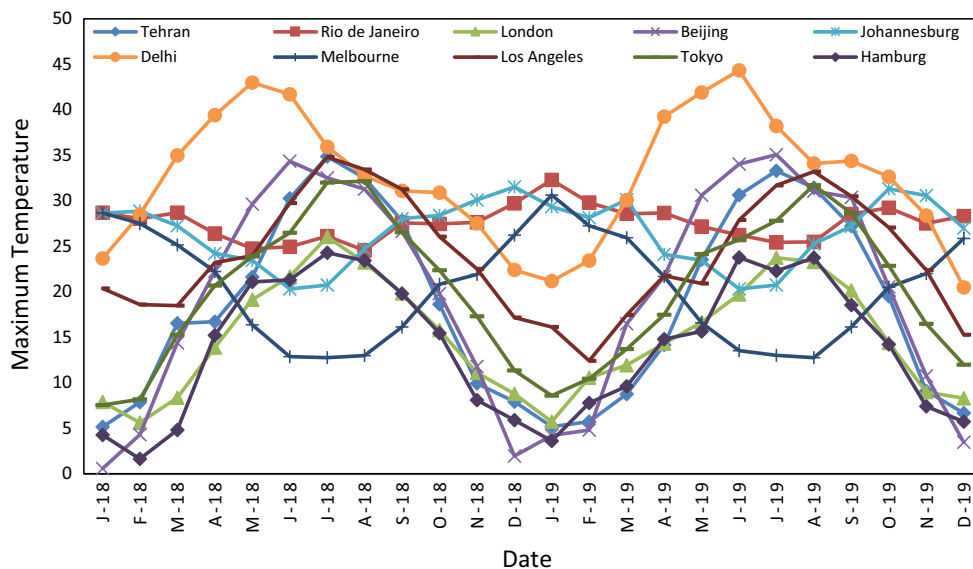
ANN method

The artificial neural network is a novel method for solving complex and time-consuming problems. It is commonly used in almost all science branches for modeling conceptual phenomena since one of the essential capabilities of this technique is to comprehend nonlinear relation between parameters which is practical. ANN is an interconnected group of nodes inspired by a simplification of neurons in the human brain's nervous system. Each connection can transmit a signal to other neurons like the synapse in a brain. In an artificial neural structure, each neuron has two states, training and acting. While training, neurons are trained to find appropriate outputs for specific input. During acting states, according to the training state, the defined input gives the proper result. Therefore, this structure has many inputs and outputs (Haykin and Network 2004). Generally, neurons are aggregated into three

layers, namely, input layer (receiving raw data), hidden layers (specifying using connection weights and inputs) and output layer (affects the performance and character of the hidden layer). Different layers transform their inputs differently, and the relation between components defines the ANN's behaviour (Tahani et al. 2016). The artificial network performs as a function, and it is a vital tool for analysing. After visually training the model, the prediction level will be started. According to the input layer's amount of belonging neurons, the input variable will receive, and the same process happens for the output layer (Vakili et al. 2017b).

There are different kinds of artificial neural networks; however, in this study, feed-forward neural network is used to solve different problems. Moreover, the backpropagation algorithm is applied to train the neural network. In this way, to each incorrect matching of the network's processing element, responsibility is allocated. The hidden layer is achieved by

Fig. 3 The minimum temperature of each month between 2018 and 2019
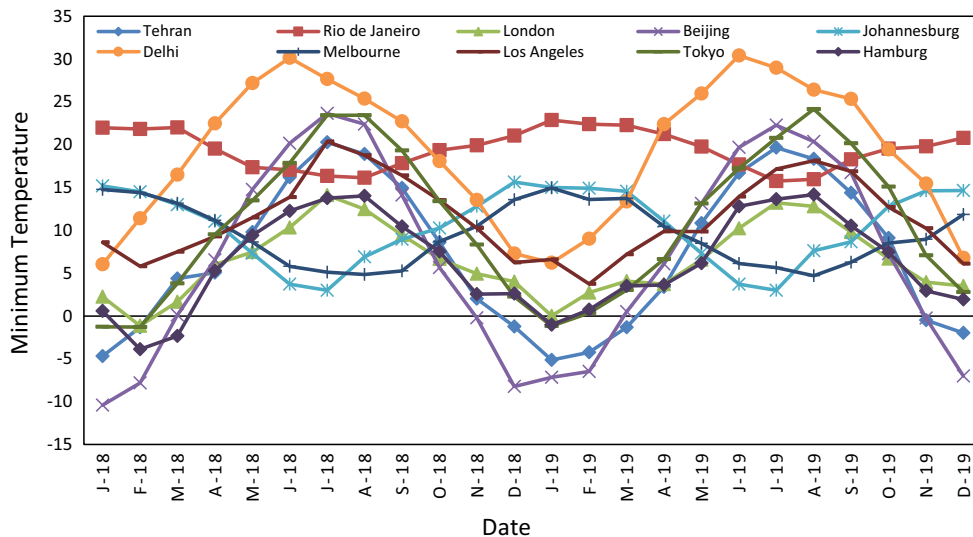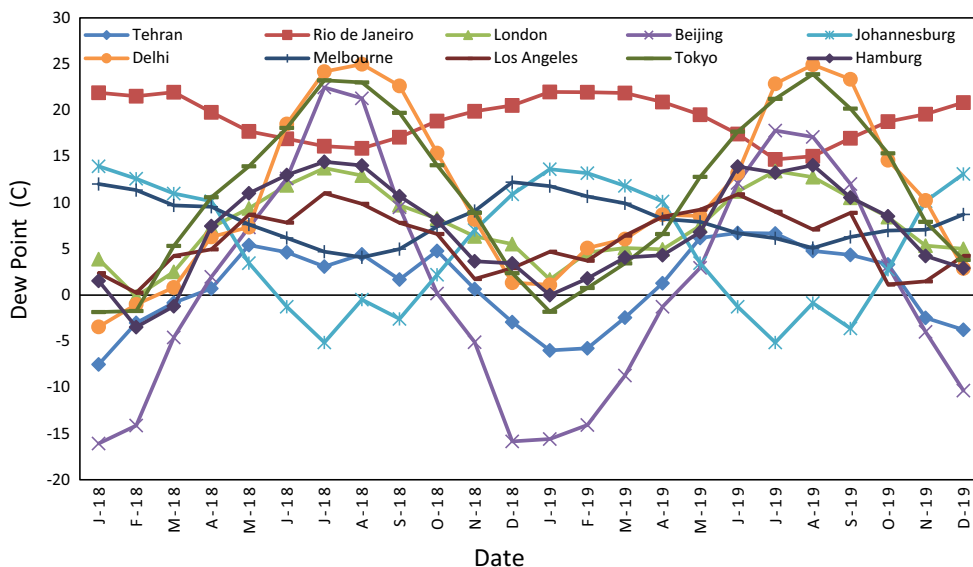
Fig. 4 The dew point of different months between 2018 and 2019

propagating the activation function gradient and the network between each hidden and input layer. Then, the inclinations and weights are adjusted in which the difference between predicted and objective mean square error was minimised.

Establishing network parameters is an empirical method. In most cases, it is an innovative technique in which networks are trained with different learning rates, various numbers of layers, different functions and momentums, and the best network is selected.

In recent decades, different activation functions have been used; however, as can be seen below, in this paper, Log-sigmoid and Tangent sigmoid for hidden layers and Purelin for output layers are used (Yahyaei et al. 2020).

$$\text{Tansig}(x) = \frac{2}{1 + \exp(-2x)} - 1 \qquad (1)$$

$$\text{Logsig}(x) = \frac{1}{1 - \exp(-x)} \qquad (2)$$

$$\text{Purelin}(x) = x \qquad (3)$$

where x is the input of the activation function.

Overall, ANN functions as a function which means the number of input variables is the same as that of input layer neurons and the number of outputs commensurate with the output layer's neurons. In this paper, the first layer's input parameters include latitude, longitude, maximum and minimum temperatures, dew point, relative humidity, wet bulb temperature, surface pressure, wind speed, $PM_{2.5}$, PM_{10}, O_3, NO_2 and mean solar radiation on a horizontal surface. According to each neuron weight and the number of neurons in the middle layers, the thermal conductivity coefficient is

Fig. 5 The wet bulb temperature of different months between 2018 and 2019
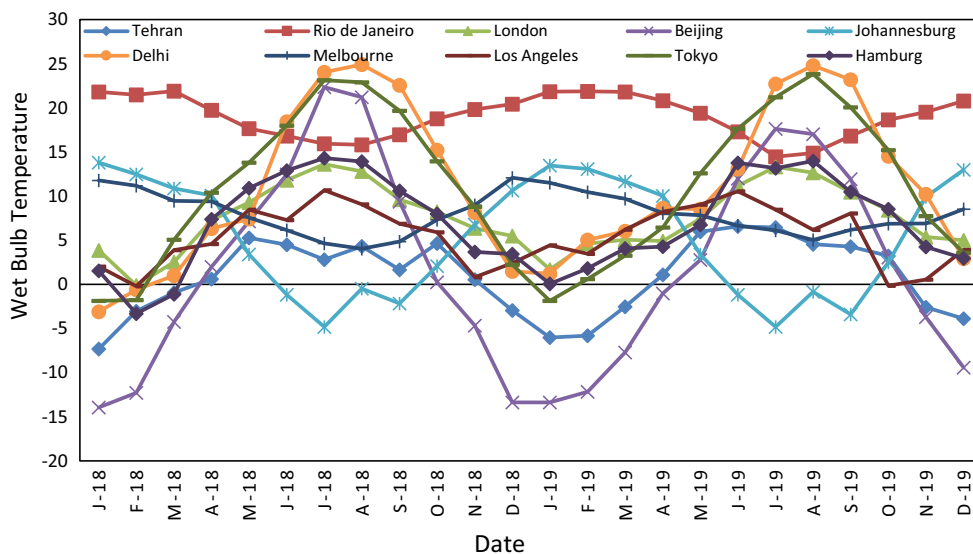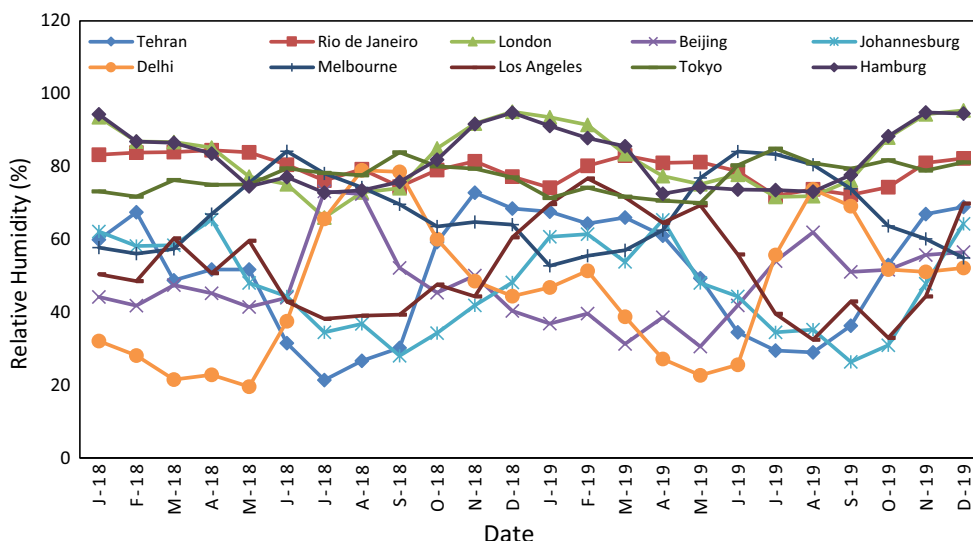
Fig. 6 The relative Humidity of different months between 2018 and 2019

presumed to be the network's output in the output layer (Fig. 14).

In this research, models are achieved by MATLAB R2019a. The input parameters, namely, latitude, longitude, min and max temperatures (°C), relative humidity (%), wind speed (m/s), surface pressure (kPa), amount of air pollutants (O_3, NO_2, $PM_{2.5}$, PM_{10}), dew frost point wet bulb temperature (°C) and mean solar radiation (MJ/m^2/day) are calculated experimentally, and 70% of them are used for training the network, 15% for validation and the rest of them for testing the model. Using Eq. 4, the data are normalised between −1 and 1 to be applied in the network (Vakili et al. 2017a).

$$y = 2 \times \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) - 1 \tag{4}$$

Hybrid genetic algorithm and ANN method

Genetic algorithm (GA) is based on the survival of the fittest or natural selection, inspired by Darwin's genetics and evolution theory. With this technique, the evolutionary process of life can be simulated. The optimisation is a typical GA method application; it is also a useful model for visual perception, feature selection, pattern recognition and machine learning (Ebrahimi-Moghadam and Moghadam 2019; Bagherzadeh et al. 2019).

Surviving a population in a genetic algorithm depends on individual desirability. The ones with high ability are more likely to get married and have children. As a result, their following generations will have the better capability. In the genetic algorithm, individuals represent as chromosomes with a set of properties, and during several generations, these chromosomes will be evolved. Each chromosome has undergone the assessment process, and their values determine whether

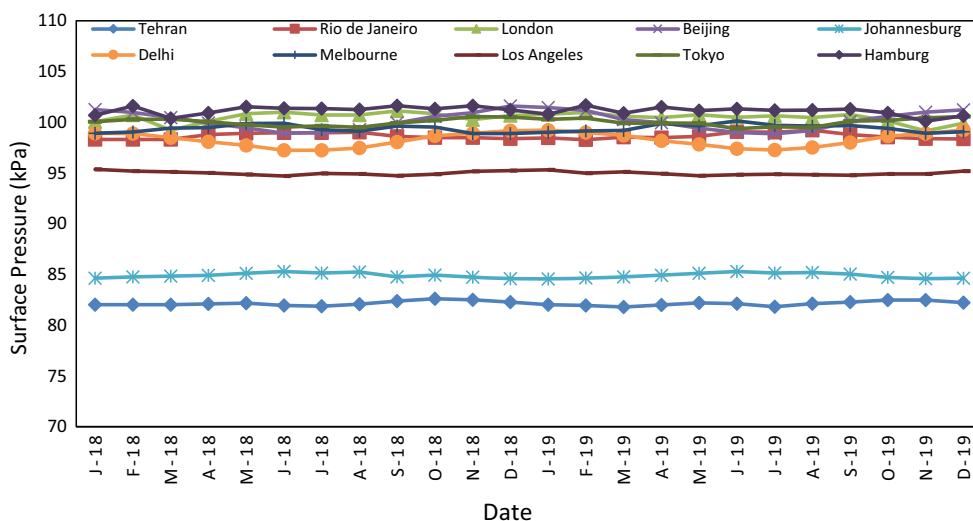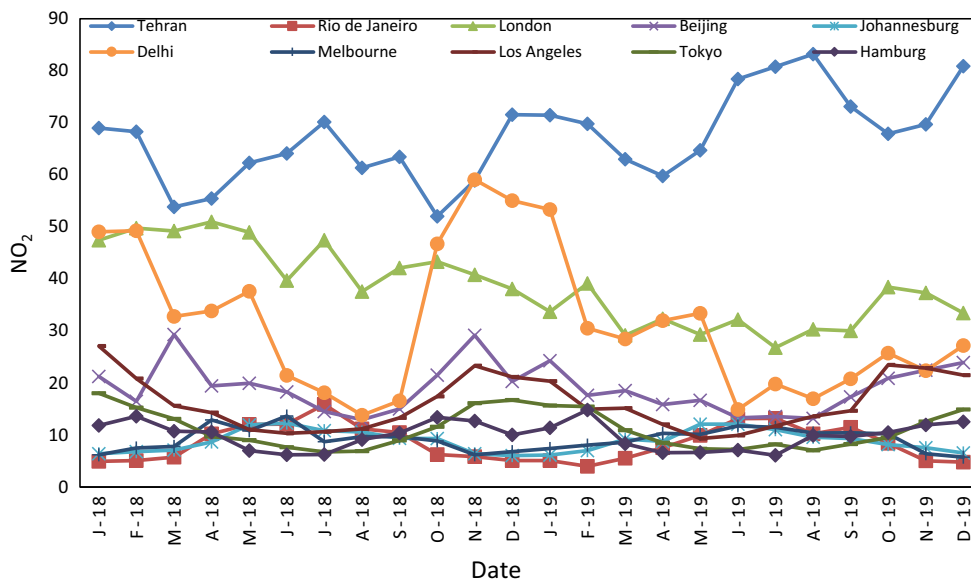**Fig. 7** The surface pressure of different months between 2018 and 2019

Fig. 8 The amount of NO₂ in the different months between 2018 and 2019

they will survive and breed or be left out. Two operators, namely, mutation and cross over, are used in the genetic algorithm's generation process. Preferable parents are selected according to a fit function (Man et al. 1999).

In each stage of the genetic algorithm, a group of space points was processed by chance. During the process, character sequence is assigned to each point, and genetic operators are applied to these sequences. Then, in order to obtain new points in the search space, the sequences will be decoded. In the end, each point's participation chance in the next stage is determined by their objective function amount. Figure 15 indicates the combination of genetic algorithm and artificial neural network.

In this study, the genetic algorithm has been used to enhance the training process of ANN. In this case, after defining

the amount of neuron correlation weights and bias of each layer by the genetic algorithm, the network trains itself, and by considering the errors, it presents a model.

In modeling by genetic algorithm, these steps are followed:

1. At first, a random population is selected. Each individual's first generation features are chosen randomly from the weight in an artificial neural network.
2. To investigate each person, the ANN is performed according to inputs, output and defined weight for each layer and neuron. In the end, the modeling output is compared with the experimental amount.
3. Individuals are ranked based on minimising the network error (elitism)

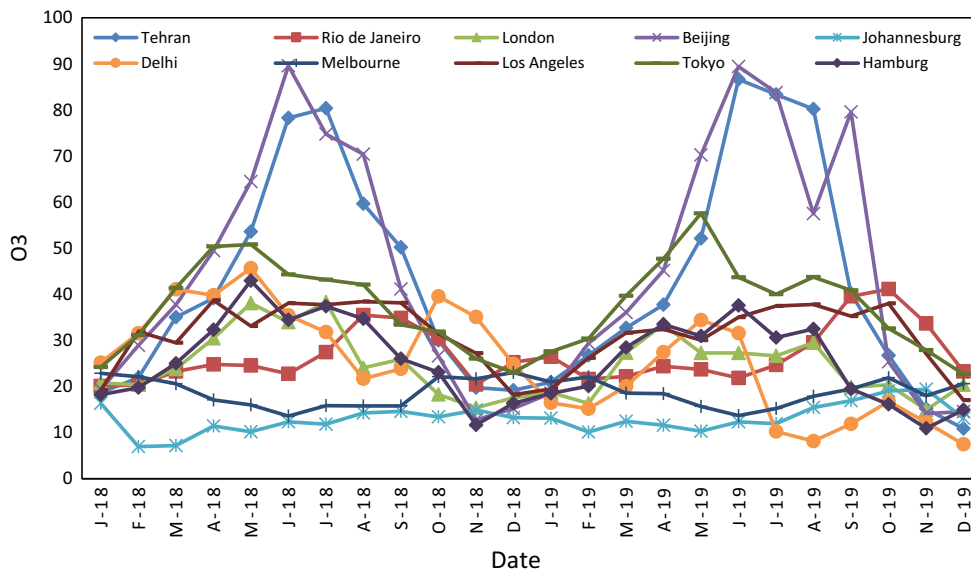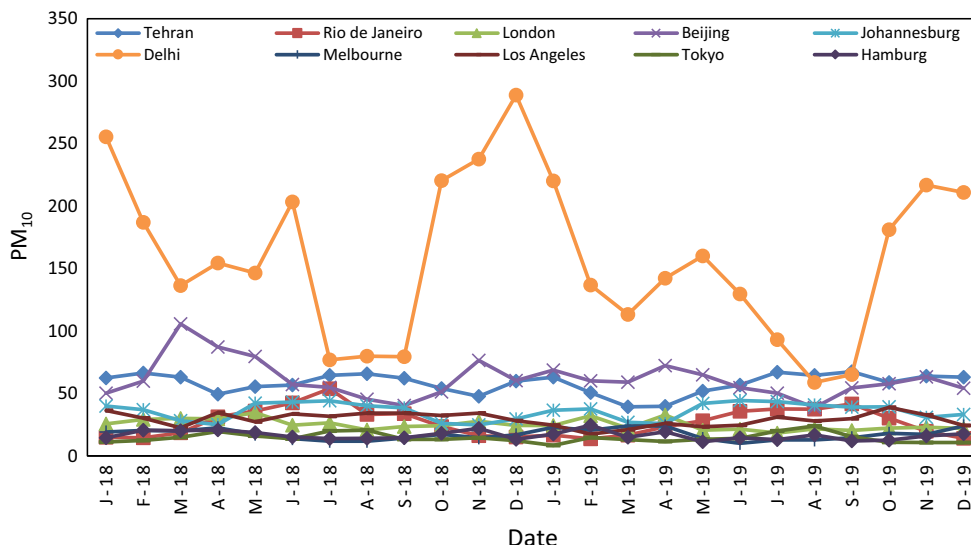**Fig. 9** The amount of O₃ in the different months between 2018 and 2019

Fig. 10 The amount of PM_{10} in the different months between 2018 and 2019

4. The next generation consists of the best person of the previous generation who has a lower error rate.

5. Best parents are selected for reproduction using genetic algorithm operators.

6. For the second generation, this technique is repeated, and with regard to the defined cycling number, the algorithm is run. The final set of weights for each neuron (the best chromosomes of the previous generation) is chosen to train the neural network.

7. The neural network will start to perform after the genetic algorithm is trained.

In this study, modeling is done by considering the initial population of 150, 50 generations, the intersection link rate of 0.5 and the mutation rate of 0.5. The algorithm stop condition is considered for lack of objective function progression for 50

consecutive generations or when the generations are finished. To substitute the newborns for the previous generations, the befitting chromosomes of the parents and children are selected as replacing the population.

ANFIS method

The adaptive network–based fuzzy inference system is a hybrid method that combines artificial neural networks and fuzzy system architecture to capture the benefits of both models in a framework. ANFIS is a powerful tool for determining nonlinear relationships of inputs and outputs (Jang 1993a). This model is based on Takagi-Sugeno fuzzy inference system in constructing five forward network layers (Takagi and Sugeno 1985). Fuzzy and neural networks establish the relation between output and input variables to improve

Fig. 11 The amount of $PM_{2.5}$ in the different months between 2018 and 2019
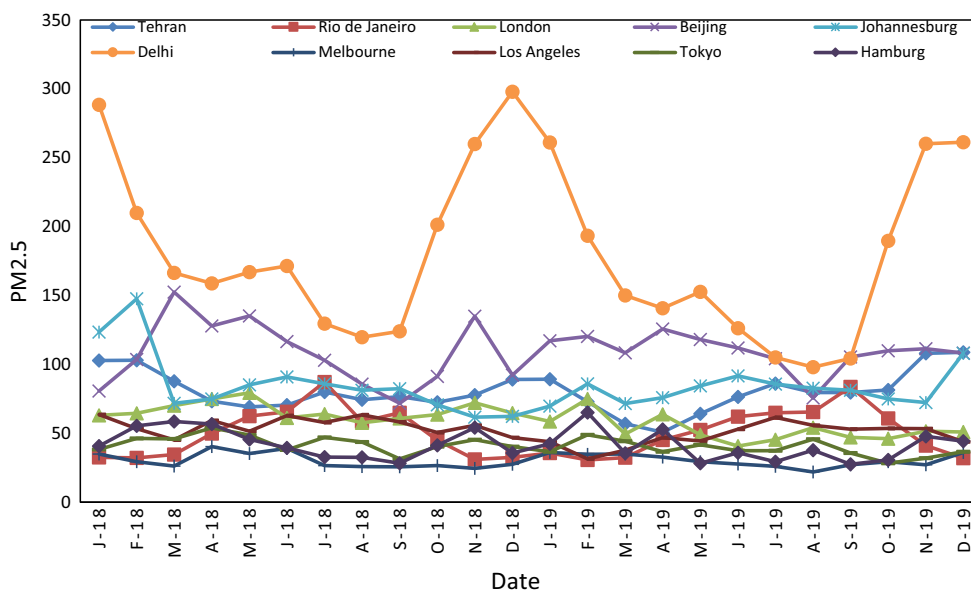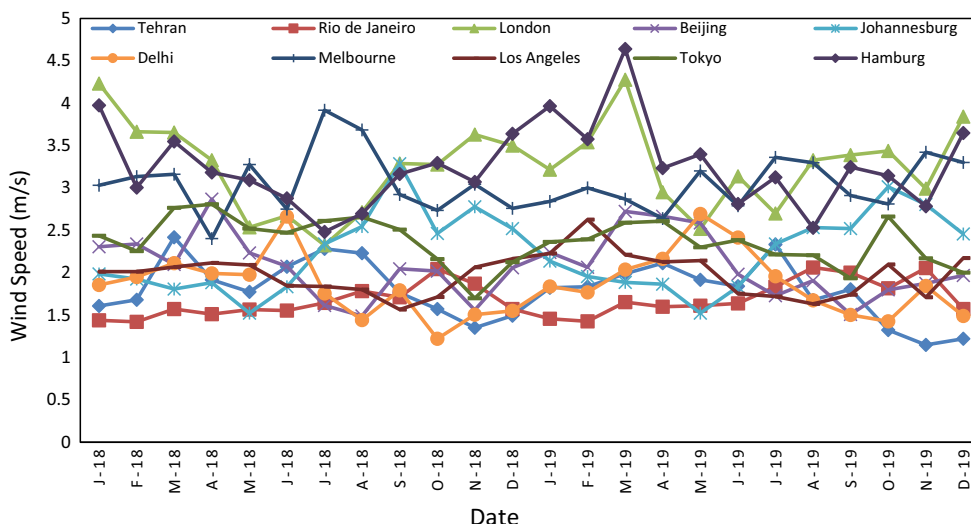
Fig. 12 The wind speed in the different months between 2018 and 2019

fuzzy membership functions and help them become more efficient. Multilayered connected networks consist of forwarding comparative nodes, which mean the learning algorithm adjusts the parameters that determine the outputs to minimise the modeling error. Figure 16 illustrates a summary of this model.

The gradient reduction rule is the fundamental principle used in adaptive networks. However, because of being slow and converging to local minima, it is not an ideal approach. Instead of hybrid methods that combine gradient-based technique with other approaches like least square estimation (LSE), it results in better performance.

Parameters of an adaptive network include all the network groups' parameters, which must be set up according to reduced-gradient training and learning method.

Five layers compose the ANFIS with N inputs number. Each input has dependency functions K_1, K_2, …, K_n, respectively. As a result, the most number of rules can be accomplished as below (Jang 1993b):

$$R = k_1 \times k_1 \times \ldots \times k_n \quad (5)$$

First layer: Input values are taken and determined by the membership functions belonging to them at this step. The fuzzy logic is applied, so this layer is commonly called the fuzzification layer. $O_{1,k}$ is the output of each layer.

$$O_{1,k} = \mu_{Ak}(x_i), \quad k = 1, 2, 3\ldots, n \quad (6)$$

$$\mu_{Ak}(x_i) = \exp\left[-\left(\frac{x_i - \tau_k}{2\sigma_k}\right)^2\right], \quad k = 1, 2, 3\ldots, n \quad i = 1, 2, .., R$$

$$(7)$$

Fig. 13 The mean solar radiation of the different months between 2018 and 2019
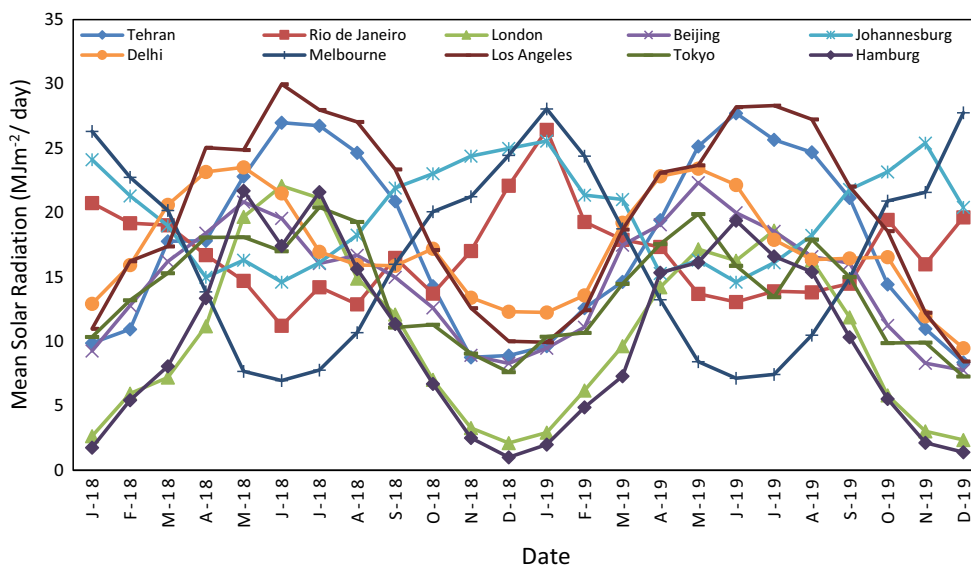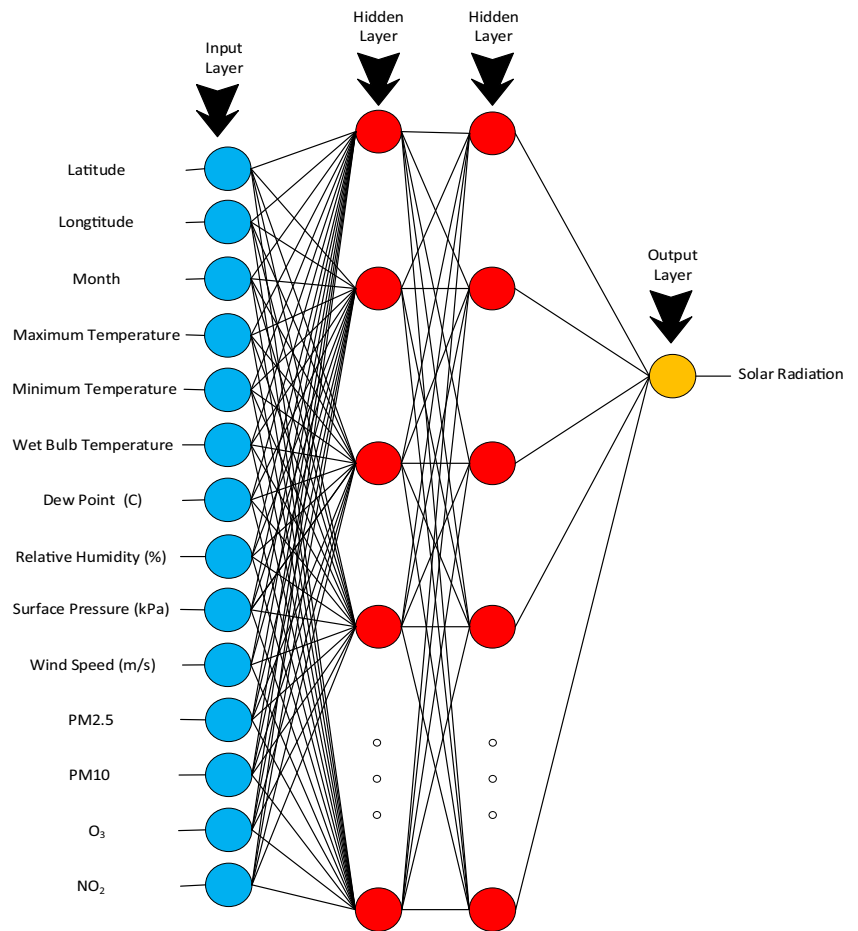
Fig. 14 Schematic diagram of the ANN structure

Fig. 15 Hybrid model of genetic algorithm and neural network structure

$$
\begin{aligned}
&i_1 \in \{1, 2, \ldots, k_1\} \\
&i_2 \in \{1, 2, \ldots, k_2\} \\
&\qquad \ldots, \\
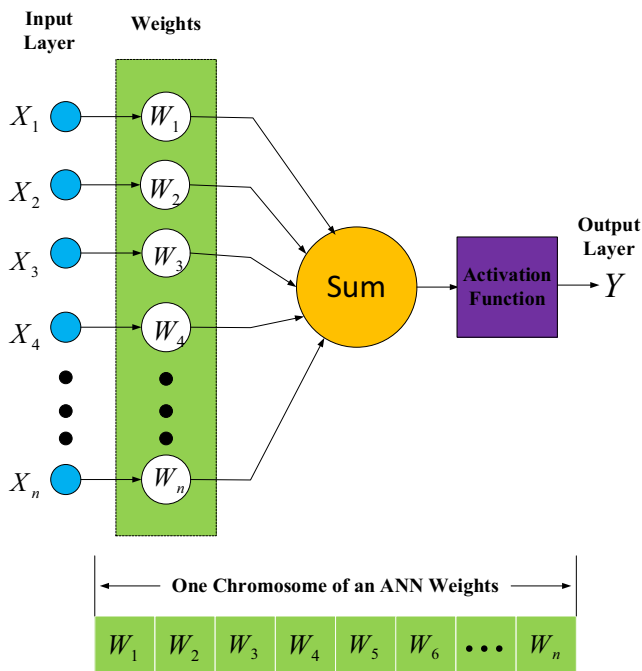&i_n \in \{1, 2, \ldots, k_n\}
\end{aligned}
\tag{8}
$$

where σ_k and τ_k are the fuzzy set A_k membership function parameters, and each node parameter defines the form of the fuzzy set membership function.

Second layer: Firing strengths are generated, and nodes are responsible for calculating their rule. These nodes called rule nodes.

$$
O_{2,k} = w_k = \mu_{Ak}(x_1) . \mu_{Ak}(x_2), \quad k = 1, 2, 3 \ldots, n
\tag{9}
$$

Third layer: The computed firing strengths are normalised by dividing each value for all firing strength, so these nodes are called intermediate nodes.

$$
O_{3,k} = \overline{w}_k = \frac{w_k}{\sum_{k=1}^{n} w_k}, \quad k = 1, 2, 3 \ldots, n
\tag{10}
$$

where $\overline{w}_k$ is the kth rule normalised firing strength.

Fourth layer: The fourth layer's nodes' role, known as the result layer, divides each fuzzy rule normalised weight in the last part output.
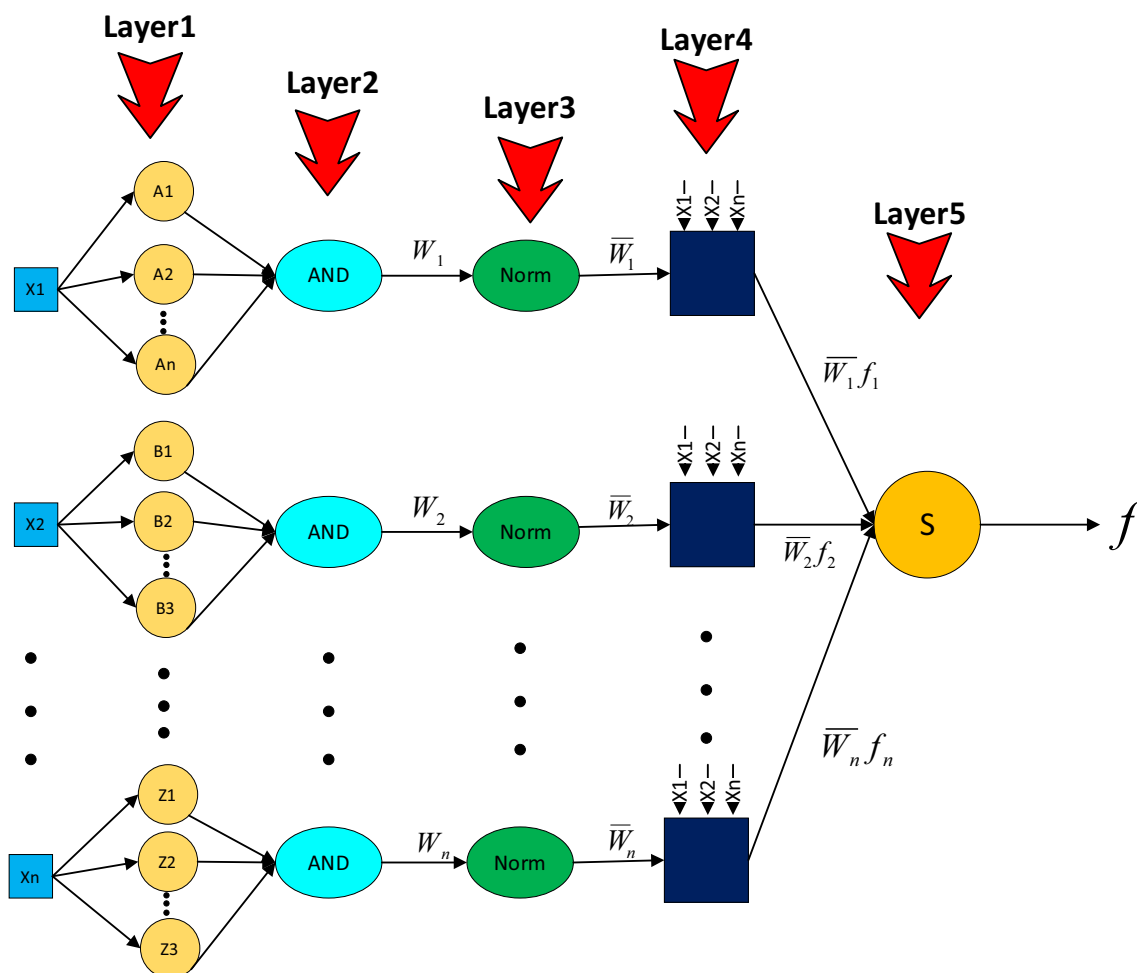
Fig. 16 Schematic of the ANFIS model

$$O_{4,k} = \overline{w}_k f_k = \overline{w}_k \left(p_0^k + p_1^k x_1 + p_2^k x_2 \right), \quad k = 1, 2, 3 \ldots, n \qquad (11)$$

Fifth layer: The fifth layer defuzzicates the outputs of the fourth layer in order to pass them to the final output of the whole network, and it is called output layer. (The number of nodes and outputs are equal.)

$$O_{5,k} = \sum_{k=1}^{n} \overline{w}_k f_k \qquad (12)$$

The final output can be obtained as below by putting Eq. 11 in Eq. 12.

$$O_{5,k} = \widehat{Y} = \sum_{k=1}^{n} \left(\overline{w}_k p_0^k \right) + \sum_{k=1}^{n} \sum_{i=1}^{n} \left(\overline{w}_k p_i^k \right) x_i \qquad (13)$$

A hybrid algorithm combined forward and backward propagations in order to train the data. The least-squares method (backwards) is applied to optimise the secondary parameters. The primary parameters are gained by applying the reduced-gradient algorithm (backward) when the secondary

parameters are estimated. Afterwards, for anticipating parameters of the model, the learning algorithm was employed as below:

Dividing data into two groups, testing and training.

1- The data are trained.
2- By putting the primary and secondary parameters in Eq. 12, the output is gained.
3- The acquired parameters are gone through trial and error using testing data.
4- The training ends when it meets the maximum number of iterations or the error is less than the predefined value. Otherwise, step 2 is rerun, and the weights are optimised.
5- The testing data error is computed in order to validate the method.

In designing fuzzy systems, selecting the correct number of rules is crucial. Too many regulations cause a highly complex system, whereas insufficient rules will result in an ineffective system that cannot meet its goals. In this paper, the rules numbers of fuzzy system are vital parameter which is defined

according to subsequent modeling error and input-output pairs. A primary reason why data are gone through clustering is to classify input-output pairs into distinct categories where the fuzzy rules are employed. The number of clusters and rules is equal. Conceptually, clustering means separating data into different clusters (subset) so that the data in a cluster are similar as possible and distinguishable from other cluster data.

In this study, a fuzzy C-means algorithm is applied for clustering. This technique allows data points to belong to more than one cluster, used in pattern recognition. This method is performed according to minimise the objective function in Eq. 14:

$$J_m = \sum_{i=1}^{n}\sum_{k=1}^{n} u_{ik}^m \|x_i - C_k\|^{*2} \qquad 1 \le m \le \infty \qquad (14)$$

where each cluster centre is C_j, u_{ik} is the membership functional degree of x_i in the jth cluster, * is the norm measured the distance between the cluster centre and data point, x_i is the Ith measured data dimension and m can be responsible for any natural number (Tilson et al. 1988).

According to the preparing data procedure for modeling, at first, the input parameters were experimentally measured, namely latitude, longitude, maximum and minimum temperatures, dew point, relative humidity, wet bulb temperature, surface pressure, wind speed, PM$_{2.5}$, PM$_{10}$, O$_3$ and NO$_2$ in order to predict daily global solar radiation. While creating the network, the inputs were accidentally divided into two groups which 70% of them were accounted for training and the rest were used to test the network.

For having accurate and precise models, the number of fuzzy rules, the effects of clusters, the impact of other membership function on each other and many other influential parameters on FCM modeling are reported in Table 3.

Table 3 Terms of the developed ANFIS model

ANFIS parameters	Description/value
Structure of fuzzy	Sugeno-type
The approach of the FIS generation	FCM
Function type of input membership	Gaussian
Output membership function	Linear
Clusters number	1-20
Input number	14
Outputs number	1
Optimisation method	Hybrid
Maximum Iterations number	100
Size of initial step	0.01
Rate of step size decrease	0.9
Rate of step size increase	1.1

Statistical analysis

In this study, for estimating the accuracy and performance of the model, various statistical measures were applied, such as determination coefficient (R^2), mean absolute percentage error (MAPE), root mean square error (RMSE) and mean bias error (MBE). For examining the model's accuracy, the last three indexes are suitable and the closer they are to zero, the more accurate the model is. To define the future correlation tendency between the two datasets, the coefficient of determination is used. Again, as much as it is close to zero, the model possesses a higher accuracy and shows the model's better performance. The statistical measures equations are as below (Vakili et al. 2021):

$$RSME = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(G_p - G_a\right)^2} \qquad (15)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\left(G_p - G_a\right)}{G_a}\right| \times 100 \qquad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(G_a - G_p\right)^2}{\sum_{i=1}^{n}\left(G_a - G_a\right)^2} \qquad (17)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}\left(G_p - G_a\right) \qquad (18)$$

where G_p is the predicted amount of solar radiation, G_a is the real amount of monthly solar radiation, G_a is the average of real solar radiation during the measurement period and n is the number of measured samples.

Results and discussion

ANN method

As mentioned before, 70% of data were responsible for training, 15% for testing and the rest were accounted for validating the model. In order to choose the best amount for the number of neurons and hidden layers, trial and error are employed. The statistical indicators are assessed according to the different layers and neurons to find the best result. The results are indicated in Table 4 (the best result of modeling mentioned in bold).

According to the result shown in Table 4, which is based on modeling by the artificial neural network, the most optimal model is selected with more accuracy than other models. Among different activation functions, including tansig, logsig and purline, the most optimal model is created using the tansig function in the first hidden layers, the logsig function in the second hidden layer and the purelin function in the output layer. Furthermore, the optimal number of neurons in the first layer is eight and in the second layer is ten. Since the only

Table 4 The results of calculated statistical indexes of modeling accuracy in ANN modeling concerning various conditions

| Model		Function	Neurons	MAPE	RMSE	R^2	MBE
ANN	1	Tansig-Tansig-Purelin	8-12-1	1.38	0.065	0.991	0.003
	2	Tansig-Logsig-Purelin	4-11-1	3.34	0.073	0.988	−0.003
	3	**Tansig-Logsig-Purelin**	**8-10-1**	**0.43**	**0.062**	**0.995**	**−0.012**
	4	Logsig-Logsig-Purelin	3-8-1	4.13	0.08	0.987	0.002
	5	Logsig-Logsig-Purelin	6-12-1	1.68	0.077	0.986	−0.002
	6	Logsig-Tansig-Purelin	6-10-1	2.02	0.07	0.991	0.004
	7	Tansig-Purelin	12-1	2.38	0.083	0.989	−0.0008
	8	Tansig-Purelin	8-1	1.42	0.07	0.989	0.01
	9	Logsig-Purelin	5-1	6.23	0.1	0.977	−0.008
	10	Logsig-Purelin	9-1	3.60	0.07	0.989	0.005
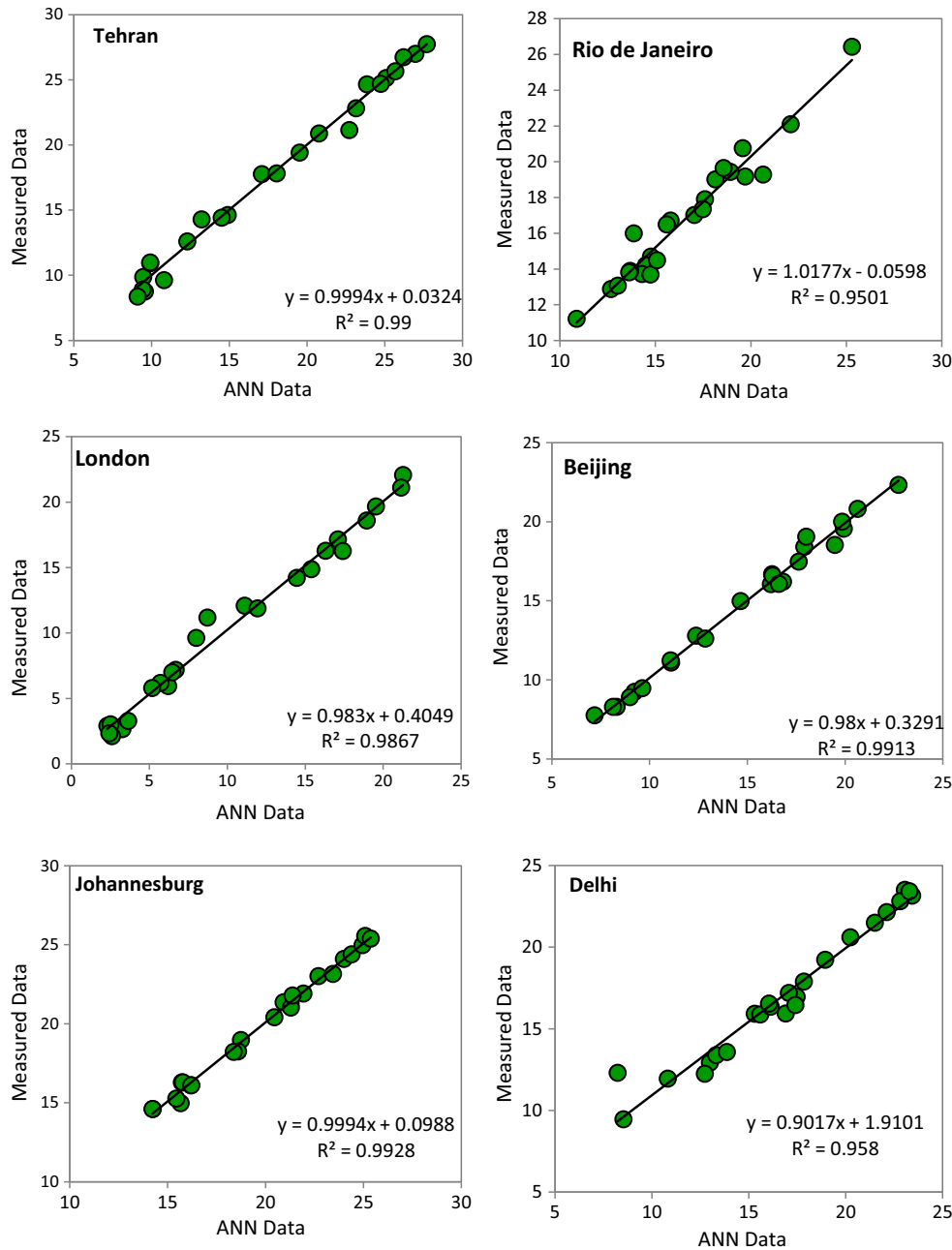
Fig. 17 The diagrams of the correlation between the foretold amounts of ANN-3 model and the experimental data in different cities worldwide
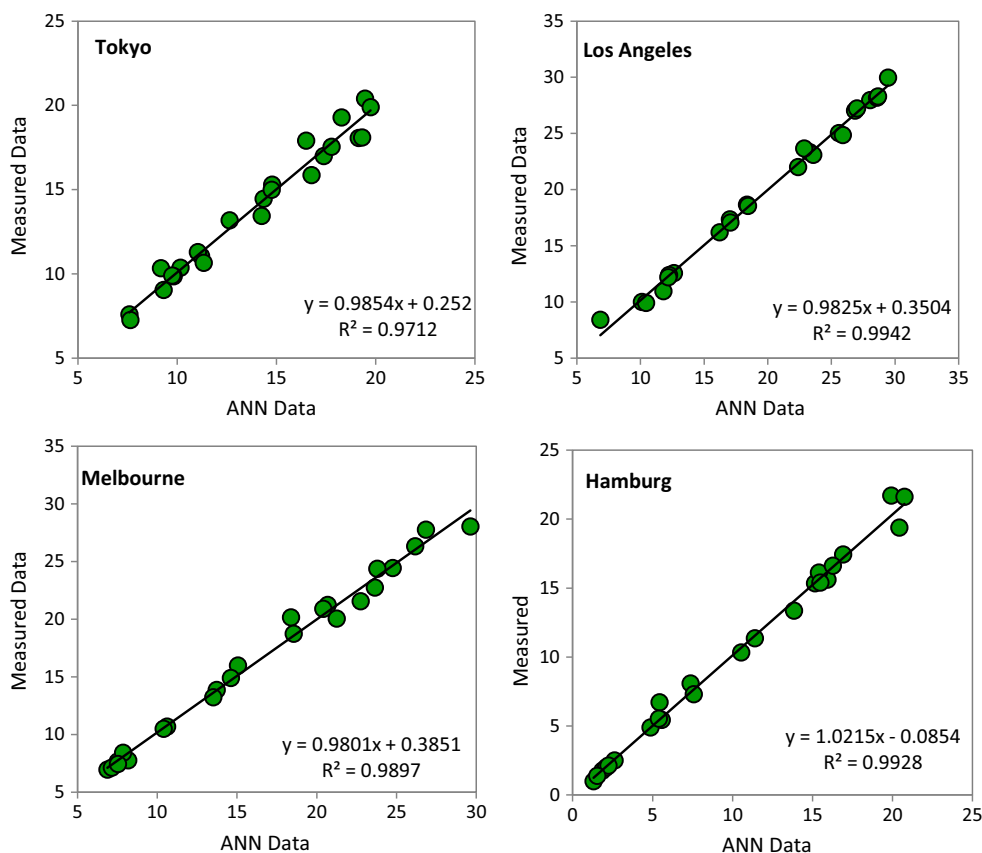
Fig. 17 (continued)

output is mean solar radiation, the last layer's output layer has a single neuron. Finally, by comparing the results, the best model is related to ANN-3.

Considering the statistical indicators introduced before, if the MAPE index is less than 10%, the model would have high accuracy and reliability. The closer it is to zero, the more accurate it would be. According to Table 4, the

Fig. 18 The error amount of ANN-3 model in comparison with real values in different cities worldwide
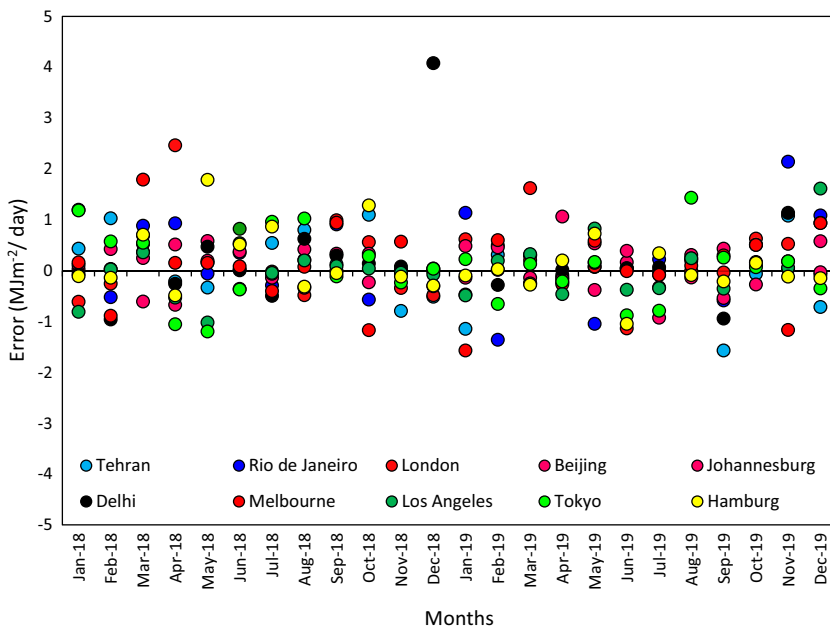
Table 5 The results of calculated statistical indexes of modeling accuracy in Ga-ANN modeling concerning various conditions

| Model		Function	Neurons	MAPE	RMSE	R^2	MBE
GA-ANN	1	**Tansig-Tansig-Purelin**	**7-10-1**	**3.052**	**0.194**	**0.945**	**0.019**
	2	Tansig-Tansig-Purelin	12-18-1	2.17	0.264	0.88	0.063
	3	Tansig-Logsig-Purelin	9-15-1	9.52	0.202	0.84	-0.025
	4	Logsig-Tansig-Purelin	14-7-1	8.58	0.231	0.914	-0.025
	5	Logsig-Tansig-Purelin	6-12-1	5.71	0.26	0.81	0.026
	6	Logsig-Tansig-Purelin	5-13-1	2.28	0.24	0.86	0.064
	7	Tansig-Purelin	8-1	0.76	0.239	0.86	0.007
	8	Tansig-Purelin	5-1	18.09	0.201	0.86	-0.002
	9	Logsig-Purelin	10-1	11.18	0.265	0.875	0.151
	10	Logsig-Purelin	7-1	1.32	0.254	0.785	-0.049
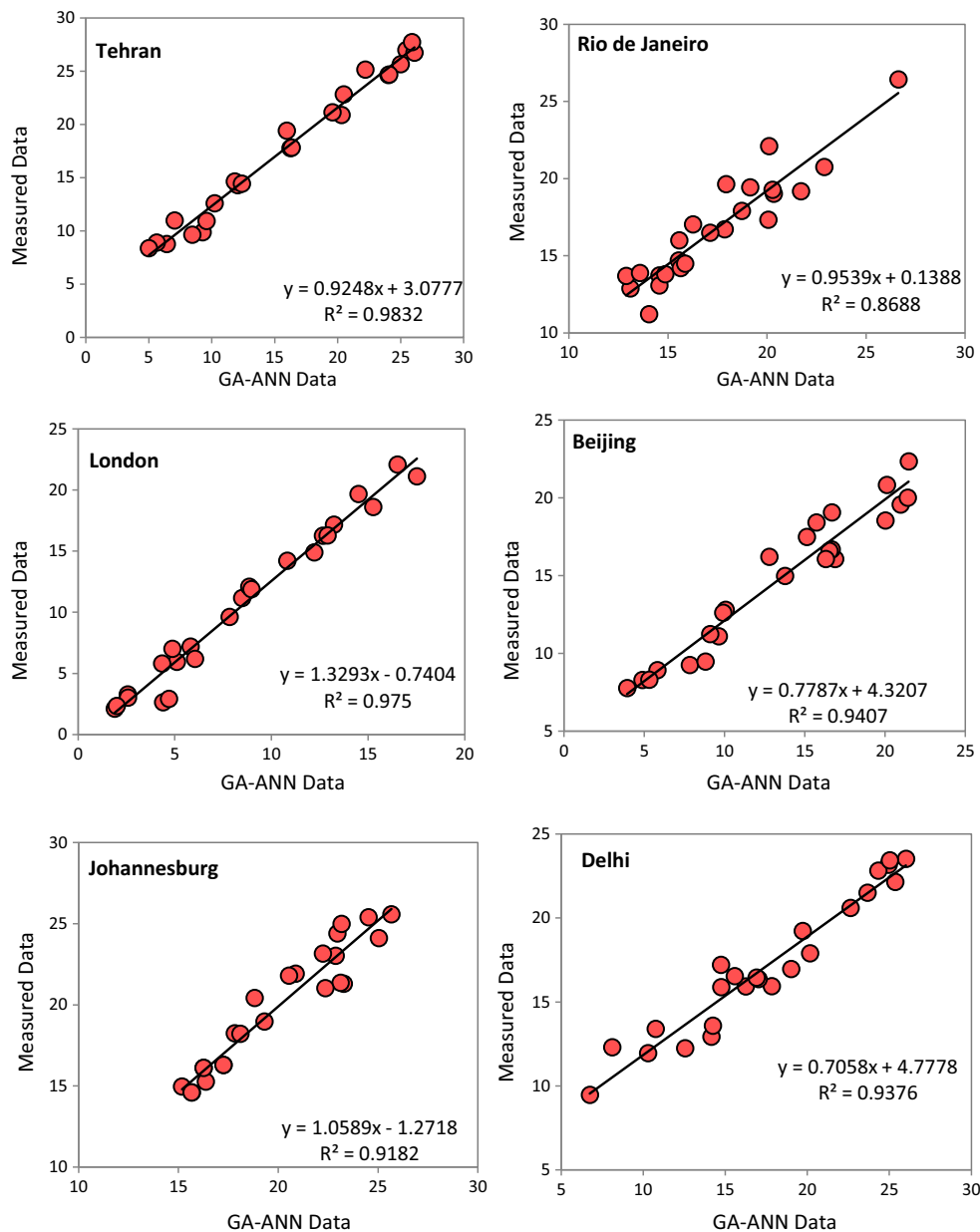
Fig. 19 The diagrams of the correlation between the foretold amounts of GA-ANN-1 model and the experimental data in different cities worldwide
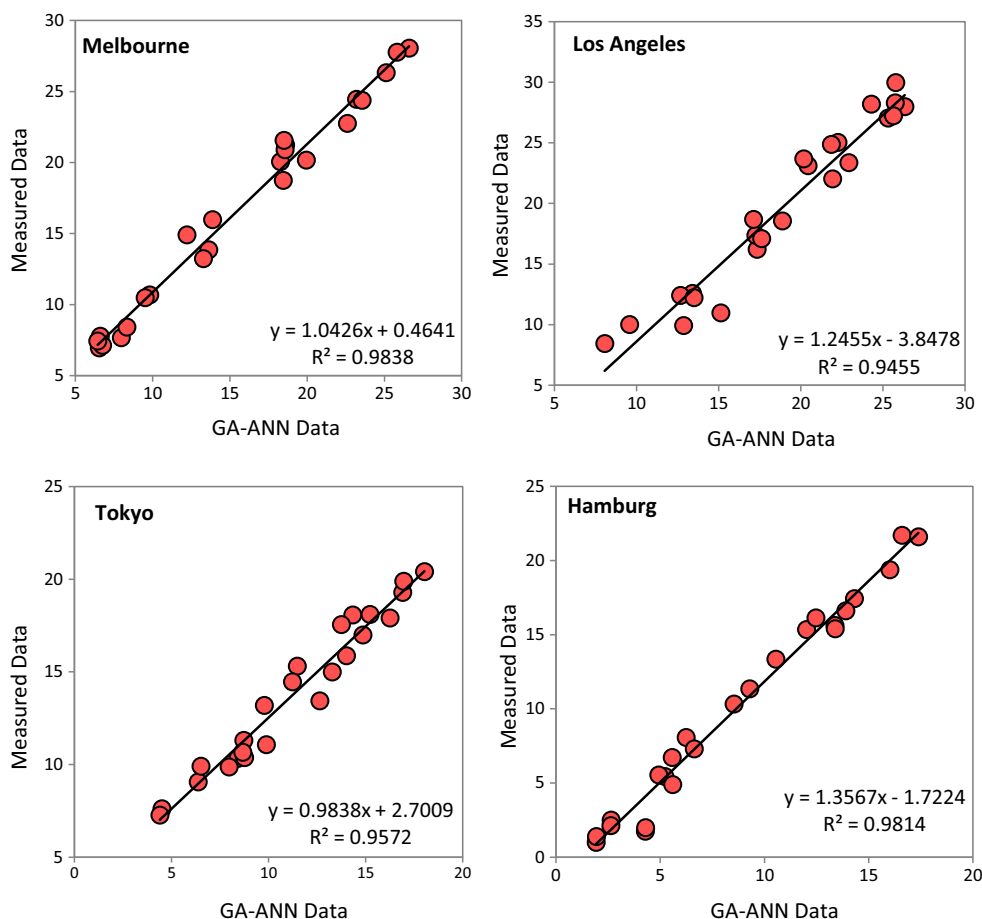
Fig. 19 (continued)

MAPE for ANN-3 is 0.43, the RMSE index is 0.062, the R^2 index is 0.995 and the MBE index is −0.012. The results of indexes can claim, the model has good accuracy and reliability.

In order to have a more comprehensive study of the neural network's performance and the results of modeling, it is essential to investigate the correlation between experimental and predicted results. Therefore, Fig. 17 shows the correlation between the optimal model's predicted data and the amount

Fig. 20 The error amount of GA-ANN-1 model in comparison with real values in different cities worldwide
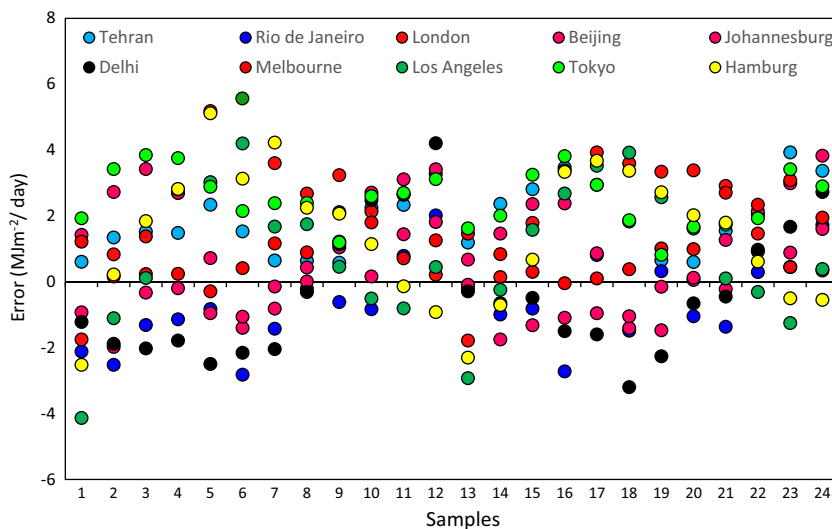
Table 6 The results of calculated statistical indexes of modeling accuracy in ANFIS modeling concerning various conditions.

Model		Cluster	MAPE	RMSE	R^2	MBE
ANFIS	1	3	1.60E−03	1.51	0.979	−2.00E−04
	2	4	4.00E−03	1.21	0.987	6.00E−04
	3	5	1.00E−02	1.13	0.98	−1.00E−03
	4	7	8.00E−04	0.86	0.993	−6.00E−04
	5	10	4.00E−04	0.68	0.995	−8.00E−04
	6	13	5.00E−05	0.58	0.996	−2.50E−05
	7	14	1.00E−04	0.536	0.997	−1.40E−06
	8	**16**	**5.50E−04**	**5.90E−05**	**0.999**	**0.425**
	9	19	2.80E−03	−2.00E−04	0.996	0.556
	10	20	1.80E−03	−1.00E−04	0.997	0.482

of experimental measurement of mean solar radiation for the different cities around the world. According to the diagrams, the correlation coefficient between the experimental and forecasted results is 0.981 on average. The correlation coefficient is close to one, and the prediction results are suitable for each city.

Figure 18 shows the deviation of the mean solar radiation predicted in the ANN-3 model from the measured data separately for the different cities in months between 2018 and 2019. According to the investigation of each measured data, the prediction results had a deviation of about 0.1 MJ/m²/day on average, and the maximum error is 4.07 MJ/m²/day, which belongs to Delhi.

Hybrid GA-ANN method

According to the stages explained in the "Hybrid genetic algorithm and ANN method" section, modeling has been done with the hybrid of genetic algorithm and artificial neural network. In this stage, as same as ANN, a trial and error method has been used to determine the number of neurons, layers and functions of the network that act in every layer. Table 5 shows the accuracy of the model results and statistics indexes and the best result of modeling mentioned in bold.

Since the trial-and-error method was used to predict a fair number of neurons and layers, investigation of statistics indexes is a suitable method for choosing the most accurate model, which has less difference between measured data and predicted data. Considering the indexes figures' investigation, it can be seen that GA-ANN 1 has the most accurate results. The MAPE index is 3.052, and since this amount is close to zero, it shows that the model is accurate. Furthermore, the amounts of RMSE and R^2 are 0.194 and 0.945, respectively. The former is near zero, and the latter is close to one, which means the model is suitable for predicting mean solar radiation. Finally,

considering the MBE index, GA-ANN 1 model predicts the mean solar radiation of about 0.019 MJ/m²/day more than the actual amount.

In Fig. 19, the amount of predicted and measured data of mean solar radiation has been compared. As can be seen, most of the data are near to the trendline, which is linear. This means that predicted data and actual data have proper conformity in comparison to each other. According to the diagrams shown in Fig. 19, the correlation coefficient between the measured data and the predicted data is 0.978 on average.

Figure 20 shows the deviation of the mean solar radiation predicted in the GA-ANN-1 model from the measured data separately for the different cities in months between 2018 and 2019. According to the investigation of each measured data, the prediction results had a deviation of about 1.00 MJ/m²/day on average, and the maximum error is 5.66 MJ/m²/day, which belongs to London.

ANFIS method

In line with the process of preparing data for modeling, at the beginning of modeling, measured data, namely, latitude; longitude; min and max temperatures; relative humidity; wind speed; surface pressure; amount of O_3, NO_2, $PM_{2.5}$ and PM_{10}; dew point; and wet bulb temperature are used as input parameters and mean solar radiation as the output parameter.

In this method, while creating a network, the input data is randomly divided into two sections, 70% of them would be dedicated to training and the remaining 30% to network testing. According to the measured data, which are 2856 data, including 13 types of input data and one type of output, 2000 data were used for training and 856 data for testing.

In this data clustering method, FCM, the trial-and-error method, has been used to determine each cluster's repetition number. Therefore, in each rerun, the number of clusters has been changed, and the statistic indexes for any new model have been calculated. The accuracy and precision of each model and the results of each model's statistical indicators are shown in Table 6 and the best result of modeling mentioned in bold.

According to Table 6, the mean absolute percentage error is 5.50E−04, the root mean square error is 5.90E−05, the determination coefficient is 0.999 and the mean bias error is 0.425. These values indicate high accuracy and reliable model results. Figure 21 shows the correlation between the optimal model's predicted data and the amount of measured data of mean solar radiation for the different cities worldwide. According to the diagrams, the correlation coefficient between the experimental and the anticipated results is 0.988, which means that the model is suitable for predicting mean solar radiation in every city around the world.

Figure 22 shows the deviation of the mean solar radiation predicted in the ANFIS-8 model from the measured data

Fig. 21 The diagrams of the correlation between the foretold amounts of ANFIS-8 model and the experimental data in different cities worldwide

Tehran — y = 1x − 0.0003, R² = 0.9865
Rio de Janeiro — y = 1x − 0.0002, R² = 0.9151
London — y = 1.0001x − 0.001, R² = 0.9996
Beijing — y = 0.9934x + 0.0689, R² = 0.9996
Johannesburg — y = 0.995x + 0.0863, R² = 0.9988
Delhi — y = 1.0037x − 0.0537, R² = 0.9997
Tokyo — y = 1.0001x − 0.0016, R² = 0.9883
Los Angeles — y = 1.0004x − 0.0293, R² = 0.9996
Melbourne — y = 0.9938x + 0.0833, R² = 0.9997
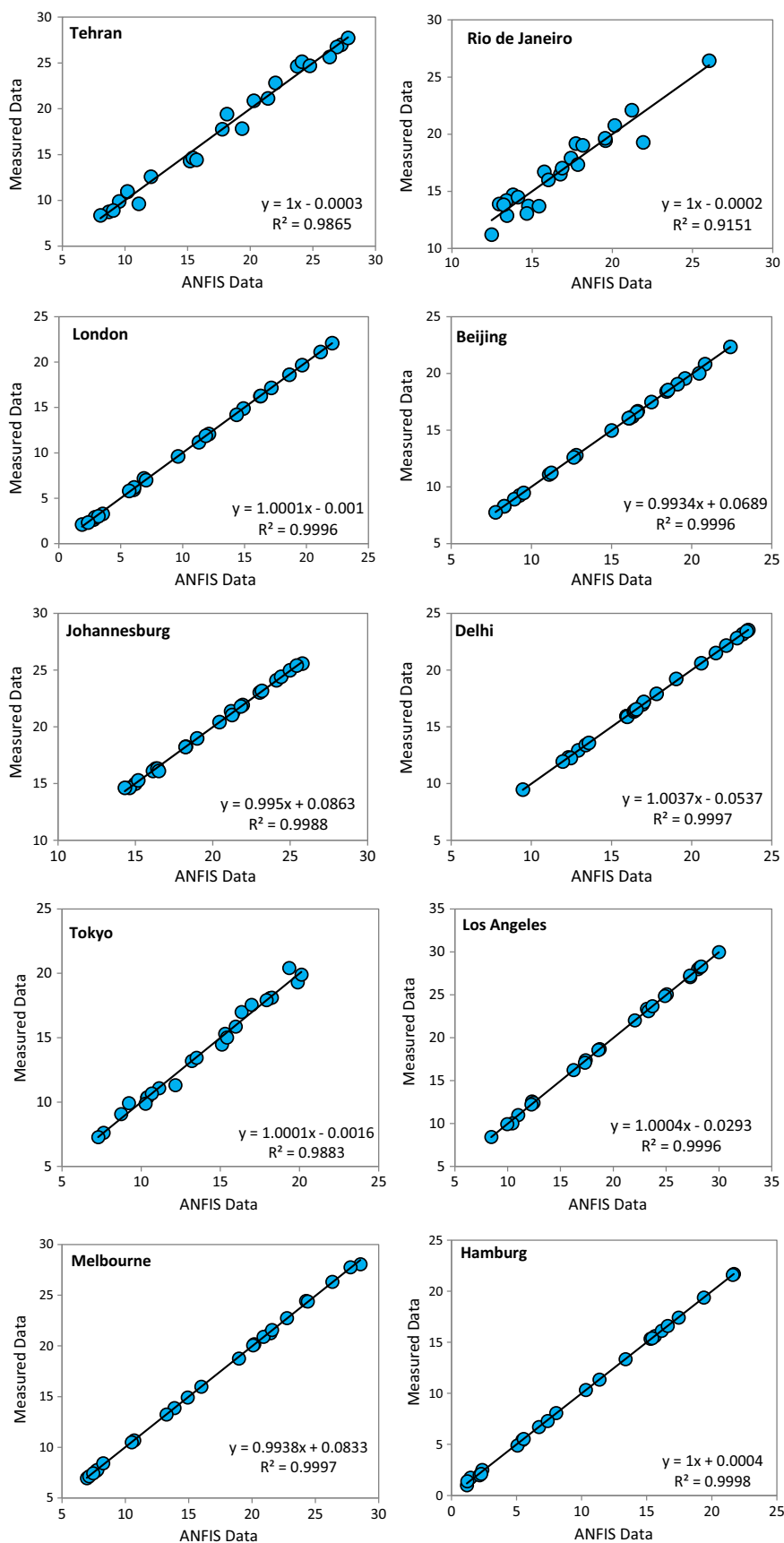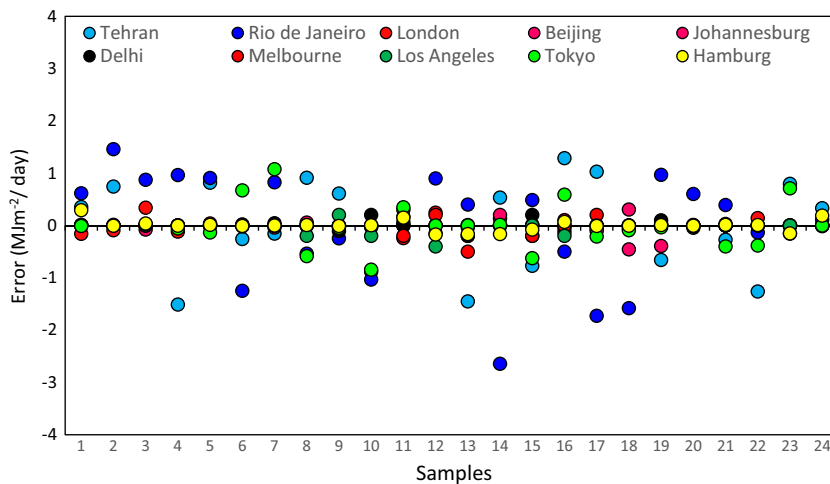Hamburg — y = 1x + 0.0004, R² = 0.9998

Fig. 22 The error amount of
ANFIS-8 model in comparison
with real values in different cities
worldwide

separately for the different cities in terms of the months between 2018 and 2019 investigating each measured data. The predicted results had a deviation of about −0.007 MJ/m²/day on average. The maximum error is −2.64 MJ/m²/day, which belongs to Rio de Janeiro.

It is necessary to compare three models for investigation of the performance of each model. Therefore, in Fig. 23, each model's most accurate results have been compared with all cities' measured data. According to

the results shown in Fig. 23, there is the most correlation between data in ANFIS modeling. It can be seen that the points in this modelling are closer to the bisector line, which means that predicted are almost as high as the measured data. Moreover, each model's calculated R index shows that the ANFIS model, which is 0.995 and the most amount among different models, is the most accurate.

Fig. 23 The diagrams of the
correlation between the measured
data and the predicted amounts of
three models: **a** ANN, **b** hybrid
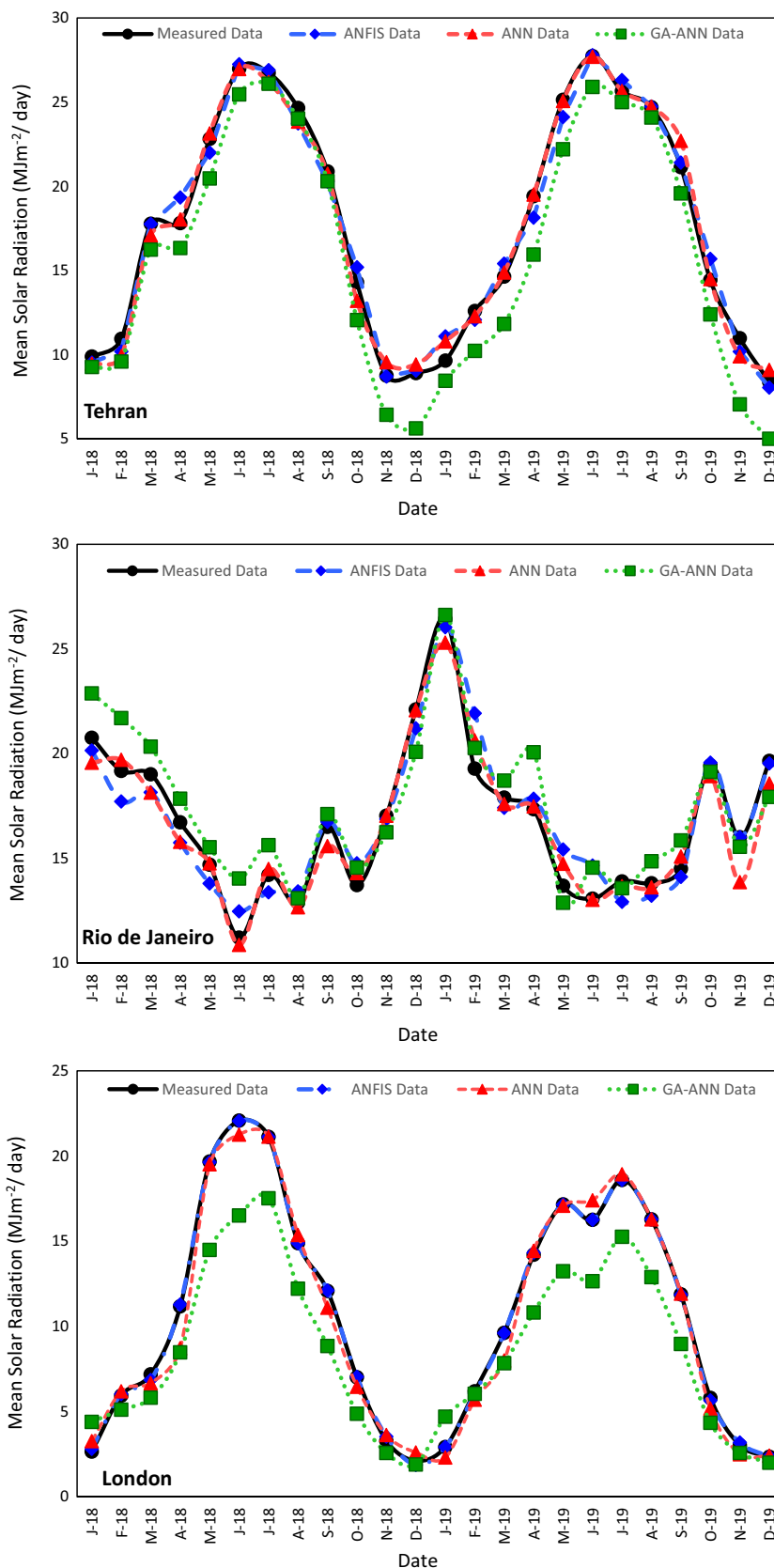ANN and GA and **c** ANFIS
around the world
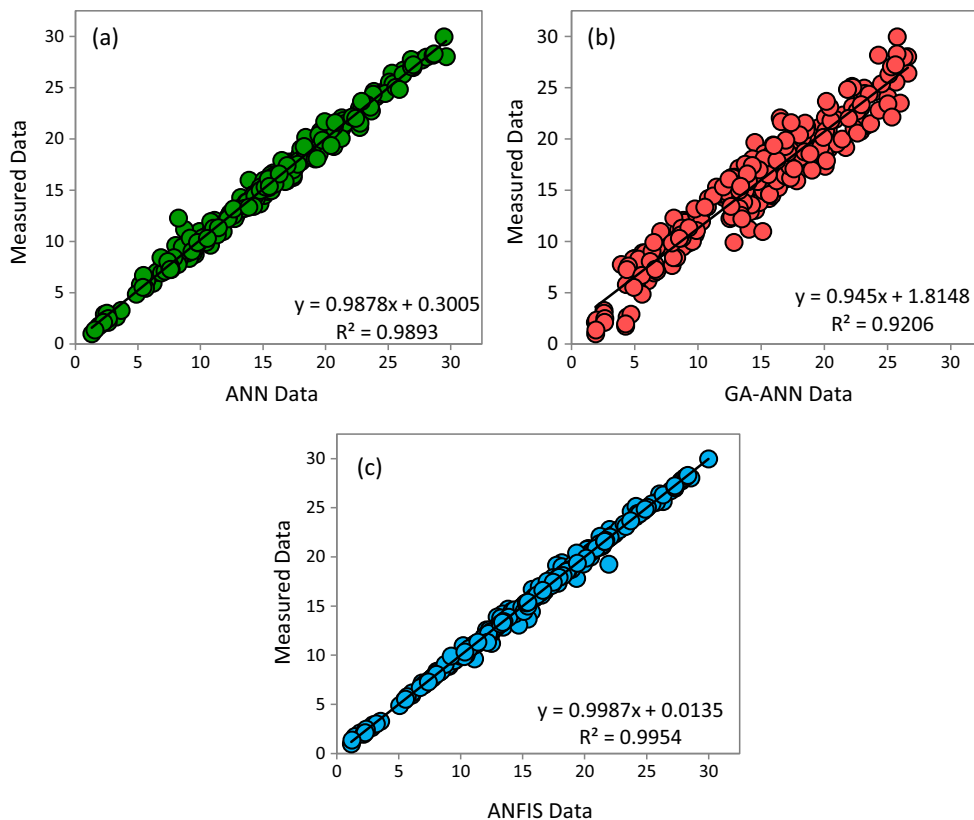
Fig. 24 The investigation of the accuracy of the different modelling methods on prediction of the amount of mean solar radiation in various cities
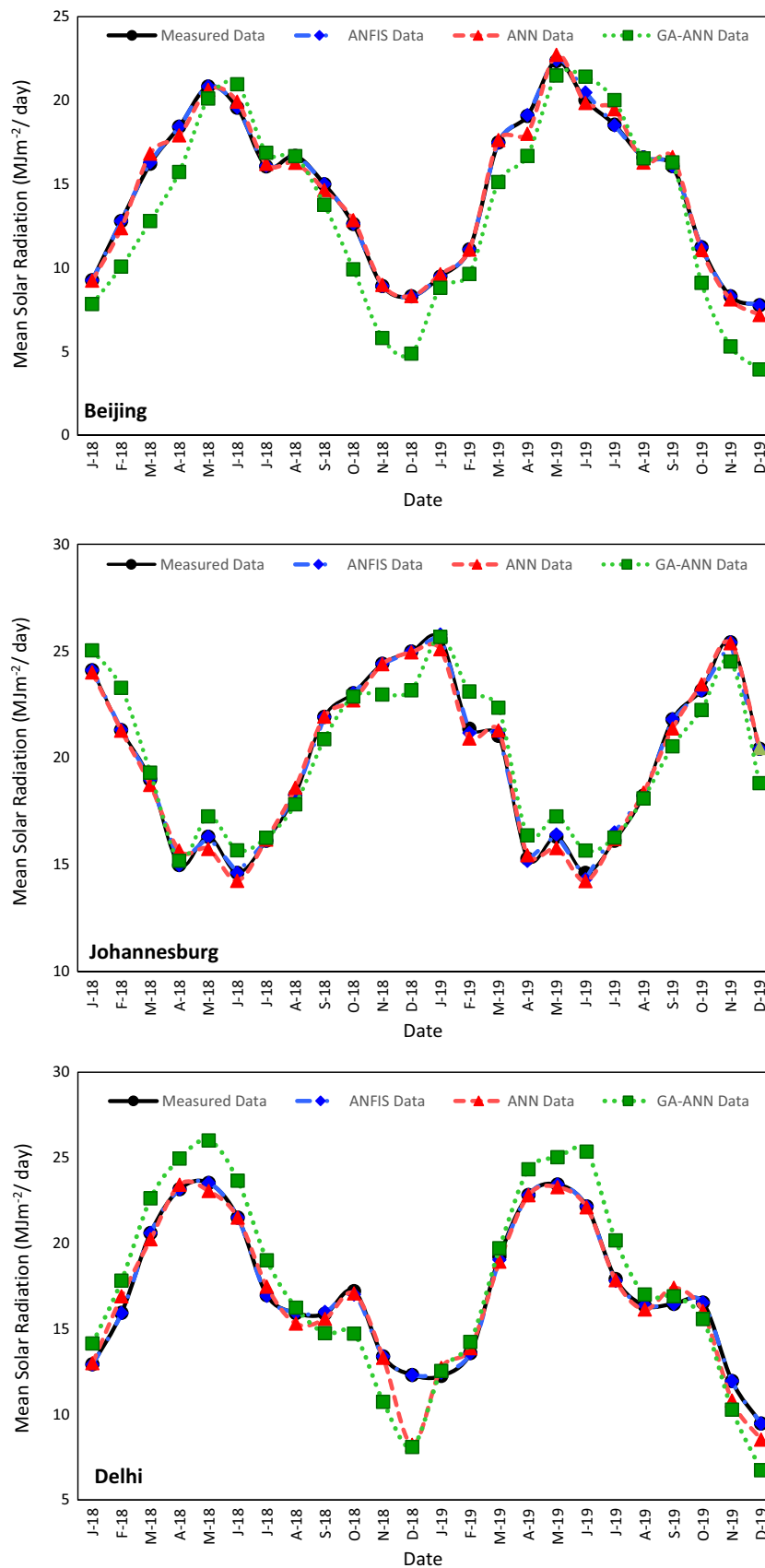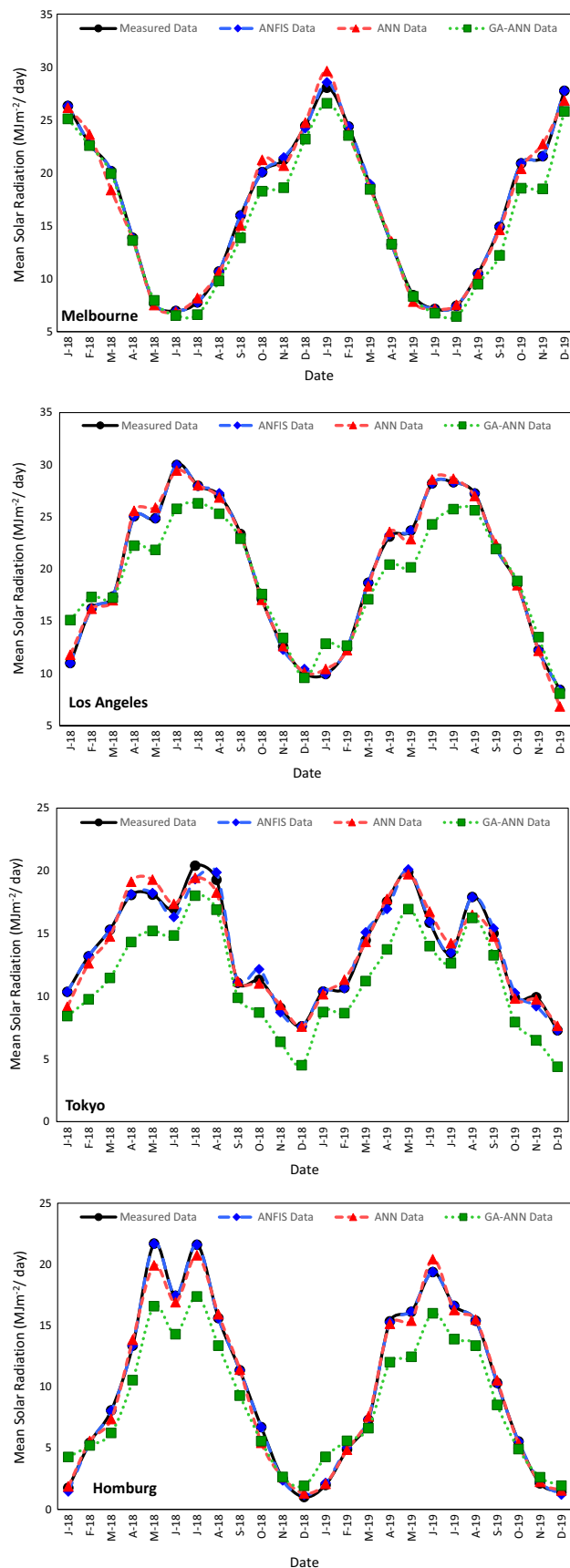
Fig. 24 (continued)

Fig. 24 (continued)

As shown in Fig. 24, there is a good correspondence between predicted data by the ANFIS-8 model and the measured data compared with other models, which are ANN-3 and ANN-GA-1. Generally, there is a slight difference between the figures for predicted data on the one hand and the figures for actual data. Among different cities, Hamburg has the most conformity, while Rio de Janeiro has the least.

Moreover, this study's modeling results show that input parameters' role in predicting solar radiation is vitally important. Using parameters such as latitude, longitude, min and max temperatures, relative humidity, wind speed, surface pressure, amount of air pollutants (O_3, NO_2, $PM_{2.5}$, PM_{10}), dew frost point and wet bulb temperature as input parameters can result in a decline in mean absolute percentage error (MAPE) compared to other studies which other researchers in the prior did according to Table 7.

Conclusion

Solar energy has gained considerable attention as clean and sustainable energy. Solar radiation is a vital parameter in modeling, designing and operating solar energy systems, and accurate solar radiation estimation is the primary concern in solar energy applications. Although because of the limited number of meteorological stations, solar radiation data are unavailable for each desired location. These days, empirical and artificial intelligence models have been developed to predict solar irradiance data at every point of interest. These techniques utilise the correlation between solar radiation and different atmospheric factors derived from climatological and geographical data.

In this study, mean monthly global solar radiation has been estimated by ANN, a hybrid of ANN-GA and ANFIS. The investigation results showed that using a fuzzy c-means clustering type model with 16 clusters was more accurate than other models, ANN and GA-ANN. This was determined by the determination coefficient, mean absolute percentage error, root mean square error and mean bias error, which were 0.999, 5.50E−04, 5.90E−05 and 0.425, respectively.

Finally, since the measurement empirically of global solar radiation needs specific equipment, it also may be expensive and impossible in some areas. Therefore, according to the modelling results, particularly the ANFIS method, the use of artificial intelligence for predicting mean monthly global solar radiation by considering input parameters is recommended by authors to replace the empirical method. Moreover, the authors would like to encourage other researchers to develop the learning process using metaheuristic algorithms, such as ant colony, particle swarm, to training input data.

Table 7 Comparison between the results of this study with other prior researches done by other researchers

Source	Network type	Location of station	MAPE (%)	Input parameters
Al-Shamisi et al. (2014)	ANN	UAE	4.61	Mean relative humidity, maximum temperature and mean wind speed, sunshine
Assas et al. (2014)	ANN	UAE	4.86	Relative humidity, wind speed, the maximum temperature and sunshine hours
Faceira et al. (2015)	ANN	Portugal	5.1	Past values of solar radiation, time, latitude, longitude and date
Çelik et al. (2016)	ANN	Turkey	5	Latitude, longitude, altitude, month, monthly mean temperature and monthly mean sunshine duration
Vakili et al. (2017c)	ANN	Iran	1.50	Relative humidity, wind speed, $PM_{2.5}$ and PM_{10}, max and min temperature,
Torabi et al. (2019)	CBA-ANN-SVM	Iran	1.342	Relative humidity, sunshine duration, maximum and minimum temperatures and daily clearness index
Rocha et al. (2019a)	BFGS	Brazil	20.83	Day of the year, maximum temperature, minimum temperature, irradiance, precipitation, cloudiness, extraterrestrial radiation, relative humidity, evaporation and wind speed
Ağbulut et al. (2021)	ANN	Turkey	17.82	Daily minimum and maximum ambient temperature, cloud cover, daily extraterrestrial solar radiation, day length
Taki et al. (2021)	RBF	Iran	5.1	Day length, daily average relative humidity, daily maximum temperature, daily minimum temperature, daily average temperature, daily total rainfall
Present study (ANN-3)	ANN	China, India, Iran, Japan, Germany, England, South Africa, Brazil, the USA and Australia	0.43	Latitude, longitude, min and max temperatures, relative humidity, wind speed, surface pressure, amount of air pollutants (O_3, NO_2, $PM_{2.5}$, PM_{10}), dew frost point and wet bulb temperature
Present study (GA-AN-N-1)	GA-ANN	China, India, Iran, Japan, Germany, England, South Africa, Brazil, the USA and Australia	3.052	Latitude, longitude, min and max temperatures, relative humidity, wind speed, surface pressure, amount of air pollutants (O_3, NO_2, $PM_{2.5}$, PM_{10}), dew frost point and wet bulb temperature
Present study (ANFIS-8)	ANFIS	China, India, Iran, Japan, Germany, England, South Africa, Brazil, the USA and Australia	$5.50E{-}04$	Latitude, longitude, min and max temperatures, relative humidity, wind speed, surface pressure, amount of air pollutants (O_3, NO_2, $PM_{2.5}$, PM_{10}), dew frost point and wet bulb temperature

Author contribution Sh. Riahi: writing, revising and editing, modeling and software

E. Abedini: investigating, collecting and analysing data, validation

M. Vakili: collecting data, revising, software and modeling, observing and managing the group

M. Riahi: software and modeling

Declarations

Ethical approval No ethical approval is foreseen since the research described by this paper only concerns the analysis of meteorological data made public by the institutions that collected them. No participant has been involved, not directly or indirectly, as a subject of the research itself.

Consent to participate No participant has been involved, not directly nor indirectly, in the research itself.

Consent for publication All the authors mentioned in the manuscript have agreed to authorship, read, and approved the manuscript and given consent for submission and subsequent publication.

Competing interests The authors declare no competing interests.

References

Ağbulut Ü, Gürel AE, Biçen Y (2021) Prediction of daily global solar radiation using different machine learning algorithms: evaluation and comparison. Renew Sust Energ Rev 135:110114. https://doi.org/10.1016/j.rser.2020.110114

Alsina EF, Bortolini M, Gamberi M, Regattieri A (2016) Artificial neural network optimisation for monthly average daily global solar radiation prediction. Energy Convers Manag 120:320–329. https://doi.org/10.1016/j.enconman.2016.04.101

Asl SFZ, Karami A, Ashari G et al (2011) Daily global solar radiation modeling using multilayer perceptron (MLP) neural networks. World Acad Sci Eng Technol 79:740–742

Assas O, Bouzgou H, Fetah S, et al (2014) Use of the artificial neural network and meteorological data for predicting daily global solar radiation in Djelfa, Algeria. In: 2014 International Conference on Composite Materials and Renewable Energy Applications, ICCMREA 2014. IEEE Computer Society

Assi AH, Al-Shamisi MH, Hejase HAN, Haddad A (2013) Prediction of global solar radiation in UAE using artificial neural networks. In: Proceedings of 2013 International Conference on Renewable Energy Research and Applications, ICRERA 2013. IEEE Computer Society 196–200

Bagherzadeh SA, Sulgani MT, Nikkhah V, Bahrami M, Karimipour A, Jiang Y (2019) Minimise pressure drop and maximise heat transfer coefficient by the new proposed multi-objective optimisation/ statistical model composed of "ANN + Genetic Algorithm" based on empirical data of CuO/paraffin nanofluid in a pipe. Phys A Stat Mech its Appl 527:121056. https://doi.org/10.1016/j.physa.2019.121056

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

Bamehr S, Sabetghadam S (2021) Estimation of global solar radiation data based on satellite-derived atmospheric parameters over the urban area of Mashhad, Iran. Environ Sci Pollut Res 28:7167–7179. https://doi.org/10.1007/s11356-020-11003-8

Baser F, Demirhan H (2017) A fuzzy regression with support vector machine approach to the estimation of horizontal global solar radiation. Energy 123:229–240

Behrang MA, Assareh E, Ghanbarzadeh A, Noghrehabadi AR (2010) The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Sol Energy 84:1468–1480. https://doi.org/10.1016/j.solener.2010.05.009

Belaid S, Mellit A (2016) Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate. Energy Convers Manag 118:105–118

Belmahdi B, Louzazni M, El Bouardi A (2020) A hybrid ARIMA–ANN method to forecast daily global solar radiation in three different cities in Morocco. Eur Phys J Plus 135:1–23

Bouzgou H, Gueymard CA (2017) Minimum redundancy–maximum relevance with extreme learning machines for global solar radiation forecasting: toward an optimised dimensionality reduction for solar time series. Sol Energy 158:595–609

Çelik Ö, Teke A, Yildirim HB (2016) The optimised artificial neural network model with Levenberg-Marquardt algorithm for global solar radiation estimation in Eastern Mediterranean Region of Turkey. J Clean Prod 116:1–12. https://doi.org/10.1016/j.jclepro.2015.12.082

Chen J-L, Li G-S, Wu S-J (2013) Assessing the potential of support vector machine for estimating daily solar radiation using sunshine duration. Energy Convers Manag 75:311–318

Deo RC, Şahin M (2017) Forecasting long-term global solar radiation with an ANN algorithm coupled with satellite-derived (MODIS) land surface temperature (LST) for regional locations in Queensland. Renew Sust Energ Rev 72:828–848

Dhakal S, Gautam Y, Bhattarai A (2020) Evaluation of temperature-based empirical models and machine learning techniques to estimate daily global solar radiation at Biratnagar Airport, Nepal. Adv Meteorol

Ebrahimi-Moghadam A, Moghadam AJ (2019) Optimal design of geometrical parameters and flow characteristics for Al2O3/water nanofluid inside corrugated heat exchangers by using entropy generation minimisation and genetic algorithm methods. Appl Therm Eng 149:889–898. https://doi.org/10.1016/j.applthermaleng.2018.12.068

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

Faceira J, Afonso P, Salgado P (2015) Prediction of solar radiation using artificial neural networks. In: Lecture notes in electrical engineering. Springer Verlag 397–406

Fan J, Wang X, Wu L, Zhang F, Bai H, Lu X, Xiang Y (2018a) New combined models for estimating daily global solar radiation based on sunshine duration in humid regions: a case study in South China. Energy Convers Manag 156:618–625

Fan J, Wang X, Wu L, Zhou H, Zhang F, Yu X, Lu X, Xiang Y (2018b) Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: a case study in China. Energy Convers Manag 164:102–111

Gairaa K, Khellaf A, Messlem Y, Chellali F (2016) Estimation of the daily global solar radiation based on Box–Jenkins and ANN models: a combined approach. Renew Sust Energ Rev 57:238–249. https://doi.org/10.1016/j.rser.2015.12.111

Gani A, Mohammadi K, Shamshirband S, et al (2016) Day of the year-based prediction of horizontal global solar radiation by a neural network auto-regressive model. Theor Appl Climatol 125

Garud KS, Jayaraj S, Lee M (2020) A review on modeling of solar photovoltaic systems using artificial neural networks, fuzzy logic, genetic algorithm and hybrid models. Int J Energy Res

Ghimire S, Deo RC, Downs NJ, 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. J Clean Prod 216:288–310

Guermoui M, Gairaa K, Rabehi A, Djafer D, Benkaciali S (2018) Estimation of the daily global solar radiation based on the Gaussian process regression methodology in the Saharan climate. Eur Phys J Plus 133:211

Guermoui M, Melgani F, Gairaa K, Mekhalfi ML (2020) A comprehensive review of hybrid models for solar radiation forecasting. J Clean Prod 258:120357

Haykin S, Network N (2004) A comprehensive foundation. Neural Netw 2:41

Hou K, Shao G, Wang H et al (2018) Research on practical power system stability analysis algorithm based on modified SVM. Prot Control Mod Power Syst 3:1–7

Huang J, Troccoli A, Coppin P (2014) An analytical comparison of four approaches to modelling the daily variability of solar irradiance using meteorological records. Renew Energy 72:195–202

Jahani B, Mohammadi B (2019) A comparison between the application of empirical and ANN methods for estimation of daily global solar radiation in Iran. Theor Appl Climatol 137:1257–1269

Jang JR (1993a) ANFIS: adaptive-network-based fuzzy inference system, 23 (3)

Jang JR (1993b) ANFIS: adaptive-network-based fuzzy inference system. 23

Jumin E, Basaruddin FB, Yusoff YBM, et al (2021) Solar radiation prediction using boosted decision tree regression model: a case study in Malaysia. Environ Sci Pollut Res 1–13. https://doi.org/10.1007/s11356-021-12435-6

Kaba K, Sarıgül M, Avcı M, Kandırmaz HM (2018) Estimation of daily global solar radiation using deep learning model. Energy 162:126–135

Keshtegar B, Mert C, Kisi O (2018) Comparison of four heuristic regression techniques in solar radiation modeling: Kriging method vs RSM, MARS and M5 model tree. Renew Sust Energ Rev 81:330–341

Khanlari A, Sözen A, Afshari F, Şirin C, Tuncer AD, Gungor A (2020a) Drying municipal sewage sludge with v-groove triple-pass and quadruple-pass solar air heaters along with testing of a solar absorber drying chamber. Sci Total Environ 709:136198

Khanlari A, Sözen A, Şirin C, Tuncer AD, Gungor A (2020b) Performance enhancement of a greenhouse dryer: analysis of a cost-effective alternative solar air heater. J Clean Prod 251:119672

Khatib T, Mohamed A, Sopian K (2012) A review of solar energy modeling techniques. Renew Sust Energ Rev 16:2864–2869

Khosravi A, Nunes RO, Assad MEH, Machado L (2018) Comparison of artificial intelligence methods in estimation of daily global solar radiation. J Clean Prod 194:342–358. https://doi.org/10.1016/j.jclepro.2018.05.147

Kisi O, Parmar KS (2016) Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution. J Hydrol 534:104–112

Kumar R, Aggarwal RK, Sharma JD (2015) Comparison of regression and artificial neural network models for estimation of global solar radiations. Renew Sust Energ Rev 52:1294–1299. https://doi.org/10.1016/j.rser.2015.08.021

Liu Y, Zhou Y, Chen Y, Wang D, Wang Y, Zhu Y (2020) Comparison of support vector machine and copula-based nonlinear quantile regression for estimating the daily diffuse solar radiation: a case study in China. Renew Energy 146:1101–1112

Lotfinejad MM, Hafezi R, Khanali M, Hosseini S, Mehrpooya M, Shamshirband S (2018) A comparative assessment of predicting daily solar radiation using bat neural network (BNN), generalised regression neural network (GRNN), and neuro-fuzzy (NF) system: a case study. Energies 11:1188

Man K, Tang K, Kwong S (1999) Genetic algorithms: Concept and design

Marzo A, Trigo-Gonzalez M, Alonso-Montesinos J, Martínez-Durbán M, López G, Ferrada P, Fuentealba E, Cortés M, Batlles FJ (2017) Daily global solar radiation estimation in desert areas using daily extreme temperatures and extraterrestrial radiation. Renew Energy 113:303–311

Marzouq M, Bounoua Z, El Fadili H et al (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

Meenal R, Selvakumar AI (2018) Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters. Renew Energy 121:324–343

Mehdizadeh S, Behmanesh J, Khalili K (2016) Comparison of artificial intelligence methods and empirical equations to estimate daily solar radiation. J Atmos Solar-Terrestrial Phys 146:215–227

Mohammadi K, Shamshirband S, Anisi MH, Alam KA, Petković D (2015a) Support vector regression based prediction of global solar radiation on a horizontal surface. Energy Convers Manag 91:433–441

Mohammadi K, Shamshirband S, Tong CW, Arif M, Petković D, Ch S (2015b) A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation. Energy Convers Manag 92:162–171

Mohandes MA (2012) Modeling global solar radiation using Particle Swarm Optimization (PSO). Sol Energy 86:3137–3145

Moreno A, Gilabert MA, Martínez B (2011) Mapping daily global solar irradiation over Spain: a comparative study of selected approaches. Sol Energy 85:2072–2084

Olatomiwa L, Mekhilef S, Shamshirband S, Mohammadi K, Petković D, Sudheer C (2015) A support vector machine–firefly algorithm-based model for global solar radiation prediction. Sol Energy 115:632–644

Qin W, Wang L, Lin A, Zhang M, Xia X, Hu B, Niu Z (2018) Comparison of deterministic and data-driven models for solar radiation estimation in China. Renew Sust Energ Rev 81:579–594

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-Terrestrial Phys 155:62–70

Rabehi A, Guermoui M, Lalmi D (2020) Hybrid models for global solar radiation prediction: a case study. Int J Ambient Energy 41:31–40

Ramli MAM, Twaha S, Al-Turki YA (2015) Investigating the performance of support vector machine and artificial neural networks in predicting solar radiation on a tilted surface: Saudi Arabia case study. Energy Convers Manag 105:442–452. https://doi.org/10.1016/j.enconman.2015.07.083

Rocha PAC, Fernandes JL, Modolo AB, Lima RJP, da Silva MEV, Bezerra CAD (2019a) Estimation of daily, weekly and monthly global solar radiation using ANNs and a long data set: a case study of Fortaleza, in Brazilian Northeast region. Int J Energy Environ Eng 10:319–334. https://doi.org/10.1007/s40095-019-0313-0

Rocha PAC, Fernandes JL, Modolo AB, Lima RJP, da Silva MEV, Bezerra CAD (2019b) Estimation of daily, weekly and monthly global solar radiation using ANNs and a long data set: a case study of Fortaleza, in Brazilian Northeast region. Int J Energy Environ Eng 10:319–334

Rohani A, Taki M, Abdollahpour M (2018) A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (part: I). Renew Energy 115:411–422

Salcedo-Sanz S, Casanova-Mateo C, Pastor-Sánchez A, Sánchez-Girón M (2014) Daily global solar radiation prediction based on a hybrid Coral Reefs Optimization–Extreme Learning Machine approach. Sol Energy 105:91–98

Sumithira TR, Kumar AN (2012) Prediction of monthly global solar radiation using adaptive neuro fuzzy inference system (ANFIS) technique over the State of Tamilnadu (India): a comparative study. Appl Sol Energy 48:140–145

Tahani M, Vakili M, Khosrojerdi S (2016) Experimental evaluation and ANN modeling of thermal conductivity of graphene oxide nanoplatelets/deionised water nanofluid. Int Commun Heat Mass Transf 76:358–365. https://doi.org/10.1016/j.icheatmasstransfer.2016.06.003

Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. IEEE Trans Syst Man Cybern SMC-15:116–132. https://doi.org/10.1109/TSMC.1985.6313399

Taki M, Rohani A, Yildizhan H (2021) Application of machine learning for solar radiation modeling. Theor Appl Climatol 143:1–15. https://doi.org/10.1007/s00704-020-03484-x

Tilson LV, Excell PS, Green RJ (1988) A generalisation of the fuzzy c-Means clustering algorithm. Remote sensing Proc IGARSS '88 Symp Edinburgh 3(10):1783–1784. https://doi.org/10.1109/igarss.1988.569600

Torabi M, Mosavi A, Ozturk P, et al (2019) A hybrid machine learning approach for daily prediction of solar radiation. In: Lecture notes in networks and systems. Springer 266–274

Ustun I, Karakus C, Yagli H (2020) Empirical models for estimating the daily and monthly global solar radiation for Mediterranean and Central Anatolia region of Turkey. Int J Glob Warm 20:249–275

Üstün İ, Üneş F, Mert İ, Karakuş C (2020) A comparative study of estimating solar radiation using machine learning approaches: DL, SMGRT, and ANFIS. Energy Sources, Part A Recover Util Environ Eff 1–24

Vakili M, Hosseinalipour SM, Delfani S, Khosrojerdi S, Karami M (2016) Experimental investigation of graphene nanoplatelets nanofluid-based volumetric solar collector for domestic hot water systems. Sol Energy 131:119–130. https://doi.org/10.1016/j.solener.2016.02.034

Vakili M, Karami M, Delfani S, Khosrojerdi S, Kalhor K (2017a) Experimental investigation and modeling of thermal conductivity of CuO–water/EG nanofluid by FFBP-ANN and multiple

regressions. J Therm Anal Calorim 129:629–637. https://doi.org/10.1007/s10973-017-6217-4

Vakili M, Khosrojerdi S, Aghajannezhad P, Yahyaei M (2017b) A hybrid artificial neural network-genetic algorithm modeling approach for viscosity estimation of graphene nanoplatelets nanofluid using experimental data. Int Commun Heat Mass Transf 82:40–48. https://doi.org/10.1016/j.icheatmasstransfer.2017.02.003

Vakili M, Sabbagh-Yazdi SR, Khosrojerdi S, Kalhor K (2017c) 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

Vakili M, Yahyaei M, Ramsay J, Aghajannezhad P, Paknezhad B (2021) Adaptive neuro-fuzzy inference system modeling to predict the performance of graphene nanoplatelets nanofluid-based direct absorption solar collector based on experimental study. Renew Energy 163:807–824. https://doi.org/10.1016/j.renene.2020.08.134

Wang L, Kisi O, Zounemat-Kermani M, Salazar GA, Zhu Z, Gong W (2016) Solar radiation prediction using different techniques: model evaluation and comparison. Renew Sust Energ Rev 61:384–397

Wang H, Cai R, Zhou B, Aziz S, Qin B, Voropai N, Gan L, Barakhtenko E (2020) Solar irradiance forecasting based on direct explainable neural network. Energy Convers Manag 226:113487

Yadav AK, Chandel SS (2014) Solar radiation prediction using Artificial Neural Network techniques: a review. Renew Sust Energ Rev 33:772–781. https://doi.org/10.1016/j.rser.2013.08.055

Yahyaei M, Vakili M, Paknezhad B (2020) Artificial brain structure-based modeling to predict the photo-thermal conversion performance of graphene nanoplatelets nanofluid using experimental data. J Therm Anal Calorim. https://doi.org/10.1007/s10973-020-10198-9

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 Sust Energ Rev 82:1528–1537

Zang H, Xu Q, Bian H (2012) Generation of typical solar radiation data for different climates of China. Energy 38:236–248

Zhang Y, Cui N, Feng Y, Gong D, Hu X (2019) Comparison of BP, PSO-BP and statistical models for predicting daily global solar radiation in arid Northwest China. Comput Electron Agric 164:104905

Zhang T, Lv C, Ma F, et al (2020) A photovoltaic power forecasting model based on dendritic neuron networks with the aid of wavelet transform. Neurocomputing

Zhou Y, Liu Y, Wang D, et al (2020) A novel combined multi-task learning and Gaussian process regression model for the prediction of multi-timescale and multi-component of solar radiation. J Clean Prod 124710

Zou L, Wang L, Xia L, Lin A, Hu B, Zhu H (2017) Prediction and comparison of solar radiation using improved empirical models and Adaptive Neuro-Fuzzy Inference Systems. Renew Energy 106:343–353