

Estimation of monthly global solar radiation in the eastern Mediterranean region in Turkey by using artificial neural networks

Muhittin Sahan¹ and Emre Yakut²

¹Osmaniye Korkut Ata University, Faculty of Arts and Science, Department of Physics, Osmaniye, Turkey

²The Faculty of Economics and Business Administration, Management Information Systems, Osmaniye, Turkey

Abstract. In this study, an artificial neural network (ANN) model was used to estimate monthly average global solar radiation on a horizontal surface for selected 5 locations in Mediterranean region for period of 18 years (1993-2010). Meteorological and geographical data were taken from Turkish State Meteorological Service. The ANN architecture designed is a feed-forward back-propagation model with one-hidden layer containing 21 neurons with hyperbolic tangent sigmoid as the transfer function and one output layer utilized a linear transfer function (purelin). The training algorithm used in ANN model was the Levenberg Marquand back propagation algorithm (trainlm). Results obtained from ANN model were compared with measured meteorological values by using statistical methods. A correlation coefficient of 97.97 (~98%) was obtained with root mean square error (RMSE) of 0.852 MJ/m², mean square error (MSE) of 0.725 MJ/m², mean absolute bias error (MABE) 10.659MJ/m², and mean absolute percentage error (MAPE) of 4.8%. Results show good agreement between the estimated and measured values of global solar radiation. We suggest that the developed ANN model can be used to predict solar radiation another location and conditions.

1 Introduction

The Sun, which is a source of heat and light for all living things on Earth, is an average star with a diameter of 1.39×10^6 km, with a distance of the earth of approximately 1.49×10^8 km, and with an effective surface temperature of around 5800K [1]. It is known a source of the most important fusion process in nature. The sun is a natural nuclear fusion reactor in which hydrogen atoms combine to form helium and in the process release a tremendous amount of energy. Solar radiation on the Earth's surface varies from the solar radiation incident on the Earth's atmosphere. Latitude of a location and the time of the year cause variations in solar radiance at the Earth's surface. The spectrum of the solar radiation at the earth's surface has both direct radiation coming straight from the sun and diffuse radiation scattered from the sky and from the surroundings. The total ground radiation combining direct and diffuse radiation is called the solar global radiation.

Global solar radiation data are a fundamental input parameter for solar energy applications. Since solar radiation measurements are not easily available for many region, it is necessary to develop some methods to predict the solar radiation using available meteorological data. Many models such as empirical regression models or artificial neural network (ANN) models have been developed to predict the amount of global solar radiation on horizontal surfaces using various climatic parameters,

such as relative humidity, sunshine duration etc. It is known that artificial neural networks (ANNs) models discussed in the literature are shown to be more suitable to predict solar radiation than other empirical regression models [2,3]. The ANN models are generally presented as systems of interconnected neurons and are widely used as an alternative way to solve many complex problems. The ANN models provide a computationally efficient way of determining an empirical, possibly nonlinear relationship between a number of inputs (such as latitude, longitude, humidity, sunshine duration) and one or more outputs (such as solar energy). ANN models are type of solar prediction models and there have been several articles that have used artificial neural networks for predicting solar radiation [4-7].

2 Data Description and Methodology

A total five metrological stations with long-term available data of global solar radiation were used in this study. The geographical information (longitude, latitude, and altitude) for these stations are listed in Table 1. The stations roughly range from 36°.15 to 37°.26 latitude North and from 35°.19 to 36°.10 longitude East and from 2 to 1344m altitude (From mean sea Level).

Table 1. Geographical features of five stations from Mediterranean region (Turkey).

Location	Latitude (°N)	Longitude (°E)	Altitude (m)
Adana	37.00	35.19	27
Kozan	37.26	35.49	115
Karata	36.34	35.23	22
Antakya	36.15	36.08	100
Iskenderun	36.35	36.10	2

In the present study, we developed a computer program running under MATLAB (R2015a) using Neural Network Fitting Tool (nftool). Artificial Neural Network (ANN) was used to analyse the monthly global solar radiation data belonging to 5 centers of localization (Table 1) from east Mediterranean region in Turkey. Data from 5 stations averaged over eighteen years (1993-2010) were used to train and test the networks. The data sets of 80% were used for ANN training and the remaining 20% data sets were used for ANN testing. Once the data are trained, they can perform predictions at very high speed [7].

A typical ANN consists of an input layer, weights, a hidden layer, a transfer function, and output layer as shown in Fig. 1. Fig. 1 shows a basic diagram of an artificial neuron in a feed forward network in ANN. The input layer and the weight at the inputs consist of n elements (xn,wn). The hidden layers, considered as the second layer, contain a nonlinear transfer function such as tansig or logsig. The output layer has a linear transfer function such as purelin. Each input element is multiplied by the relevant weight and then these products and bias are collected with a summing function. The result of summation is sent to the activation function to obtain output data [8].

In order to estimate the global solar radiation, we used well known Multi-Layer Perceptron (MLP) Network and we selected a multilayer feed-forward back-propagation artificial neuron network (FFBPANN) model for ANN structure. This FFBPANN means that the artificial neurons are organized in layers, and send their signals forward, and then the errors are propagated backwards [7].

In order to determine the ANN network model, various models with different training algorithms and architectures (approximately 100 models) have been employed and tested during the training of our network. Diagnostic linear regression statistic, including R², RMSE, MSE, MBE, MABE, and MAPE (see following paragraph) have been used in order to determine the best ANN model. For this reason, we used the highest values of R² and the lowest values of RMSE, MSE, MBE, MABE, and MAPE. Some of these employed and tested training algorithm models are given in Table 2.

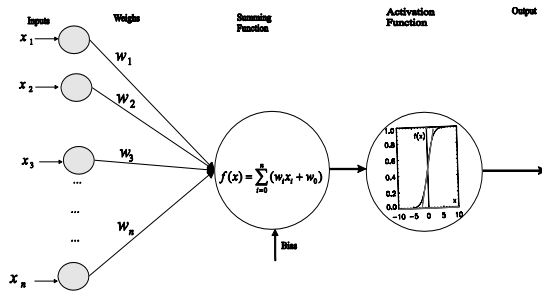


Fig. 1. Caption of the Figure 1. Below the figure. The basic diagram of a neuron in a feed forward network in ANN [8].

Table 2. Various ANN models developed with different training algorithms and architectures. (Training Algorithm: Tr.Alg., Hidden transfer Function: H.Tr.F, Output transfer Function: O.Tr.F, Number of neurons in hidden layer: Num.N.H.L)

Tr.-Alg.	H.Tr.F	O.Tr.F	Num.N.H.L	R ²	RMSE	MSE	MABE	MAPE
trainlm	tansi	purelin	21	0.98	0.85	0.72	0.65	0.04
trainlm	tansi	purelin	5	0.97	0.85	0.72	0.62	0.04
trainlm	tansi	tansig	5	0.97	0.89	0.79	0.69	0.05
trainlm	logsi	tansig	33	0.97	0.85	0.72	0.68	0.05
trainlm	logsi	tansig	30	0.97	0.92	0.84	0.71	0.05
trainlm	tansi	purelin	12	0.97	0.93	0.86	0.71	0.05
trainlm	tansi	purelin	18	0.97	0.94	0.90	0.74	0.05
trainbf	logsi	tansig	20	0.97	0.94	0.88	0.71	0.05
trainlm	logsi	tansig	24	0.97	0.99	0.98	0.70	0.05
trainbf	logsi	tansig	10	0.97	1.01	1.02	0.76	0.05
trainbf	logsi	tansig	35	0.96	1.01	1.03	0.70	0.05
trainsc	logsi	purelin	8	0.95	1.24	1.55	0.95	0.08
trainlm	tansi	purelin	33	0.97	0.90	0.81	0.69	0.05
trainlm	tansi	purelin	39	0.97	0.92	0.85	0.70	0.05
trainlm	tansi	purelin	36	0.97	0.92	0.85	0.70	0.05
trainlm	tansi	tansig	30	0.97	0.94	0.90	0.73	0.05
trainlm	logsi	purelin	12	0.97	0.90	0.81	0.70	0.05
trainbf	logsi	tansig	15	0.97	0.90	0.81	0.70	0.05
trainlm	tansi	tansig	27	0.97	0.93	0.88	0.73	0.05
trainlm	logsi	tansig	18	0.97	0.94	0.88	0.73	0.05
trainlm	tansi	tansig	33	0.97	0.95	0.90	0.73	0.05
trainlm	tansi	purelin	15	0.97	0.94	0.88	0.73	0.05
trainlm	logsi	purelin	10	0.97	0.93	0.87	0.73	0.05
trainlm	logsi	purelin	21	0.97	1.00	1.00	0.75	0.05
trainlm	tansi	tansig	36	0.97	0.92	0.86	0.68	0.04
trainlm	logsi	purelin	39	0.97	0.95	0.91	0.73	0.05

The best models developed according to training algorithms and transfer functions, and number of the neurons in the hidden layer have been selected. The “trainlm” algorithm was used in order to train the ANN network and “tansig”, and “purelin” activation functions (see Fig. 2) were used in the hidden layer and in the output layer, respectively. Our ANN architecture consists of three layers called as input layer, hidden layer and output layer. In these layers, we used 7 input neurons, 21 hidden neurons which are the collection of neurons, and a single output neuron (namely, 7-21-1) to be able to generate a modest instance for training and then testing our data. Figure 2 depicts the activation function used in the known Multi-Layer Perceptron forecaster network.

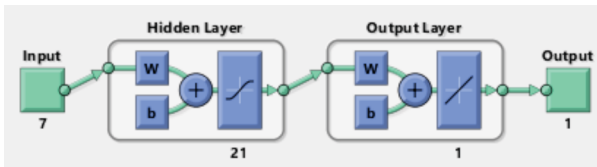


Fig. 2. Activation function used in Multi-Layer Perceptron network.

There are several statistical measures available to evaluate the association between estimated and measured values [9]. The famous of them can be given as the correlation coefficient of determination (R^2), the RMSE (Root Mean Square Error), the Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), and MABE (Mean Absolute Bias Error). These statistical indicators are defined by the following mathematical equations:

$$\begin{aligned}
 R &= \frac{\sum_{i=1}^n (x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x}_i)^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y}_i)^2}}, & MAPE &= \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{x_i} \right| \times 100 \\
 RMSE &= \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}, & MSE &= \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \\
 MABE &= \frac{1}{n} \sum_{i=1}^n |x_i - y_i|
 \end{aligned} \quad (1)$$

where y_i and x_i parameters represent the estimated and the measured (actual) values, respectively and n is the number of data points. The correlation coefficient of determination (R^2) is used to determine the linear relationship between the estimated and measured values. The R^2 varying from 0 to 1 is adopted to measure the fit of the model on estimated data. A model is more efficient when R^2 is closer to 1 [10]. The value of RMSE provides information on the short term performance of the correlations and is always positive and zero in the ideal case.

3 Results and Discussions

In this study, we applied ANN techniques to predict monthly global solar radiation using meteorological and geographical data from five stations from eastern Mediterranean region (Turkey) as Adana, Kozan, Karatas, Antakya, and Iskenderun (see Table 1). Meteorological and geographical data include the monthly mean values of the global solar radiation on horizontal surface, air temperature, relative humidity, sunshine duration, latitude, longitude, month, and altitude. All meteorological and geographical data between 1993-2010 (18 years) were obtained from Turkish State Meteorological Service (TSMS).

The actual performance of the ANN models during training and testing procedure are obtained using the basis of statistical error tests given at Eq. 1. It is found that the most successful network was single layer network with 21 neurons in hidden layer as mentioned in pervious section. Fig. 3 shows typical ANN used for estimating of solar radiation. As seen from Fig. 3, while four geographical parameters (month, latitude, longitude,

and altitude) and three meteorological parameters (monthly average air temperature, relative humidity, and sunshine duration) are used as input layer, solar radiation values are obtained from the output layer (see Fig. 3).

Comparison between measured and estimated monthly global solar radiation values for all selected regions from Mediterranean region in Turkey were given in Fig. 4. Data points (with symbol filled circle) given in Fig. 4 represent measured and estimated monthly global solar radiation data. A regression line was fitted to monthly estimated and measured global solar energy values in order to compare with each other as seen in Fig. 4. Regression equation and R^2 values are labeled in the upper left corner of Fig. 4. y and x parameters in regression equation represent measured and estimated solar energy, respectively. As seen from Fig. 4, the estimated solar radiation from the ANN model were correlated with the measured values, giving coefficient of determination (R^2) of 98%. This means that there is a good fitting between the estimated using ANN model and measured monthly mean global solar irradiation.

Monthly global solar radiation estimations from the ANN model were compared with the actual data by using simple error analysis; R^2 , RMSE, MSE, MABE, and MAPE to the monthly mean value for the five variables under study. Results for each location were given in Table 3. Table 3 contains results of R^2 , RMSE, MAPE MSE, and MABE values for selected five locations. As seen from Table 3, the linear correlation coefficients (R^2) are high for all the variables changing between a minimum of 97.1% at Adana and a maximum of 99.5% at Karatas. This implies that there are statistically significant relationships between estimated actual measured global solar radiation. The lowest RMSE value is found for location of Kozan (0.506 MJ/m²) and the maximum for Karatas (1.110 MJ/m²). The MAPE estimate for the five locations ranges between 2.8% for location of Kozan and 6.4% for location of Karatas with an overall average of 0.048%. The highest and lowest MSE values were calculated as 1.231MJ m⁻² for Karatas and 0.256MJ m⁻² for Kozan. The highest and lowest MABE values were calculated as 0.941MJ m⁻² for Karatas and 0.369MJ m⁻² for Kozan.

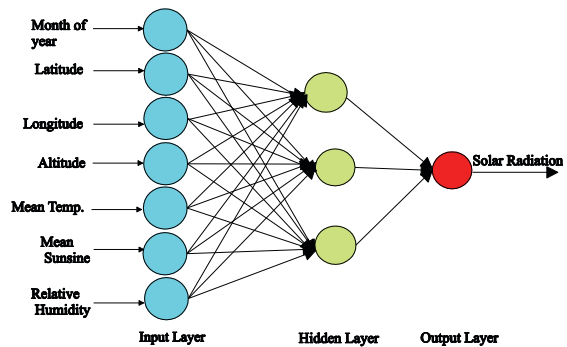


Fig. 3. Activation function used in Multi-Layer Perceptron network.

Results obtained from ANN model were compared with measured meteorological values by using statistical

methods. Fig. 5 presents monthly comparison between estimated from ANN and measured values of solar radiation for five locations (Adana, Kozan, Karatas, Antakya, and Iskenderun) given in Table 3. Location names are labeled in the upper left corner of each plot. Measured and estimated values are represented with open and filled circles and are labeled in the upper left corner of each plot. As seen from Fig. 5 and Table 3, the best R^2 were formed for locations of Karatas (0.995%) and Kozan (0.994%).

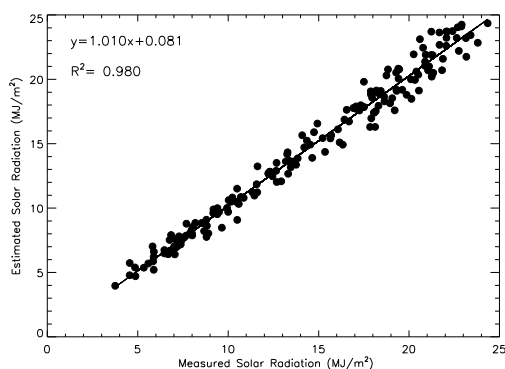


Fig. 4. Comparison between measured and estimated monthly global solar radiation values of the correlation equation for all regions from Mediterranean region in Turkey.

Table 3. Comparison of the statistical errors between measured and estimated monthly global solar radiation for five stations.

Location	R^2 (%)	RMSE (MJ/m^2)	MSE (MJ/m^2)	MABE (MJ/m^2)	MAPE (%)
Adana	0.971	1.007	1.013	0.814	0.056
Kozan	0.994	0.506	0.256	0.369	0.028
Karatas	0.995	1.110	1.231	0.941	0.064
Antakya	0.991	0.601	0.361	0.472	0.045
Iskenderun	0.992	0.874	0.764	0.696	0.047

4 Conclusion

Meteorological Service (TSMS). Results obtained from ANN model were compared with measured meteorological values by using statistical methods. These results show a good agreement between the measured by ANN model and estimated monthly mean global solar radiation values. We propose that the ANN model can be applied to estimate global solar radiation for any region having meteorological and geographical parameters.

In this study, an artificial neural network based model for prediction of solar energy has been developed. Meteorological data of five meteorological stations in Mediterranean region (Turkey) for period of 18 years (1993–2010) were taken from Turkish State Meteorological Service (TSMS).

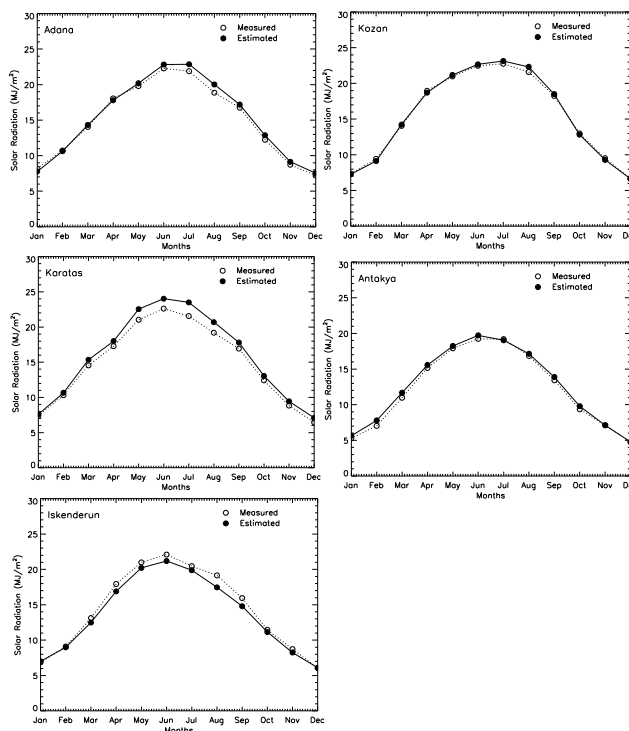


Fig. 5. Monthly mean variation of estimated (by ANN model) and measured values of global solar radiation for five different stations selected from Mediterranean region in Turkey.

References

1. J.A. Duffie and W.A. Beckman, "Solar Engineering of Thermal Processes, 2th Edition", John Wiley and Sons Inc., New York, (1991)
2. D.A. Fadare, Applied Energy, **86** (2009)
3. Y. Jiang, Energy Policy, **36** (2008)
4. M.A. Abdul Azeez, Scholars Research Library, Archives of Applied Science Research, **3**, 2 (2011)
5. A. Moreno, M.A. Gilabert, B. Martinez, Solar Energy, **85** (2011)
6. E. A Ahmed, M. E. Adam, Journal of Clean Energy Technologies, **1**, 2 (2013)
7. S. Edalati, M. Ameri, M. Iranmanesh, Journal of Renewable Energy and Environment, **2**, 1 (2015)
8. D. Graupe, Advanced Series in Circuits and Systems, **6** (2007)
9. E. Yakut, B. Elmas and S. Yavuz, Suleyman Demirel University The Journal of Faculty of Economics and Administrative Sciences, **19**, 1 (2014)
10. S.A. Kalogriou, Appl Energy **67** (2000)