PERGAMON

Energy Conversion and Management 44 (2003) 2519–2530

ENERGY CONVERSION & MANAGEMENT

www.elsevier.com/locate/enconman

Solar resource estimation using artificial neural networks and comparison with other correlation models

K.S. Reddy *, Manish Ranjan

Department of Mechanical Engineering, Regional Engineering College, Warangal 506 004, Andhra Pradesh, India

Received 19 June 2002; accepted 11 December 2002

Abstract

Artificial Neural Network (ANN) based models for estimation of monthly mean daily and hourly values of solar global radiation are presented in this paper. Solar radiation data from 13 stations spread over India around the year have been used for training and testing the ANN. The solar radiation data from 11 locations (six from South India and five from North India) were used for training the neural networks and data from the remaining two locations (one each from South India and North India) were used for testing the estimated values. The results of the ANN model have been compared with other empirical regression models. The solar radiation estimations by ANN are in good agreement with the actual values and are superior to those of other available models. The maximum mean absolute relative deviation of predicted hourly global radiation tested is 4.07%. The results indicate that the ANN model shows promise for evaluating solar global radiation possibilities at the places where monitoring stations are not established.

Keywords: Solar global radiation; Artificial neural networks; Multi-layer feed forward networks; Backpropagation; Mean absolute relative deviation; Percentage relative deviation

1. Introduction

Solar radiation received at the flat surface is most important as far as designing of solar energy systems, transpiration of crops and photosynthesis etc. are concerned. Solar radiation is one of the most vital meteorological factors determining crop productivity. India is situated between 6° N and 32° N latitudes, and hence, most of the locations in India receive abundant solar energy. Its geographic position favours the utilization of solar energy and development of solar energy

* Corresponding author. Tel.: +91-8702-430390; fax: +91-8702-459547.
E-mail address: ksreddy@nitw.ernet.in (K.S. Reddy).

0196-8904/03/$ - see front matter
doi:10.1016/S0196-8904(03)00009-8

Nomenclature

H	daily global radiation (kW/m^2)
H_0	daily extraterrestrial radiation (kW/m^2)
n	daily sunshine duration (h)
D	astronomical day length (h)
T_{max}	maximum temperature (°C)
T_{min}	minimum temperature (°C)
C_w	mean of total cloud cover of day time observations (h)
E_p	sum of squares of error for output layer neurons
k	index to neuron in output layer
p	index to training vector
y_{pk}	output of kth output neuron for pth training vector
o_{pk}	targeted output for kth output neuron for pth training vector
w_{kj}	weight on connection from jth neuron of hidden layer (immediately preceding output layer) to kth output neuron
DGR_p	predicted daily global radiation (kW/m^2)
DGR_a	actually measured daily global radiation (kW/m^2)
MAE	maximum absolute error, defined as maximum of absolute difference between actual value and predicted value

systems. Solar energy system designers require solar radiation data in various forms, depending on the exact nature of the application. Some of the diversified applications of global radiation data are: engineering system design of solar collection and storage, evaluation of performance of solar systems and future data prediction. For efficient conversion and utilization of solar energy, an accurate detailed long term knowledge of monthly mean daily global solar radiation is of prime importance. However, the mean daily solar radiation is not always the most suitable figure to characterize the potential utility of solar energy. These data can be acquired from a network of pyroheliometric and pyronometric measurements.

Various empirical models have been developed for different geographical and meteorological conditions in Saudi Arabia [1–8]. Solar radiation on horizontal and inclined surfaces are reported for India [9,10], Canada [11,12], Abu Dhabi, UAE [13], Lesotho, South Africa [14,15] and many others. An empirical formula using the daily humidity, latitude, altitude, maximum temperature and location relative to the water surface has been presented by Sabbagh et al. [16].

Kimball [17] first suggested that the sunshine fraction may be closely related to daily global radiation. Angstrom [18] proposed a relation on the basis of monthly average daily radiation. Subsequently, Prescott et al. [19] modified the Angstrom equation by including average daily clear sky radiation.

$$H = H_0\left(\alpha + \beta\frac{n}{D}\right) \tag{1}$$

where α and β are correlation constants. The disadvantage of the modified Angstrom equation is that the local effects on atmospheric transmittance of solar radiation are now considered with an additional empirical constant, whereas it was previously considered through the local average daily clear sky radiation.

Yet another method to estimate the daily global radiation, relating the difference between maximum and minimum temperatures of the day to global radiation, has been proposed by Hargreeves et al. [20], with data available on the Global Telecommunication System (GTS).

$$H = aH_0\sqrt{T_{\max} - T_{\min}} + c \tag{2}$$

where a and c are empirical constants. The estimation accuracy is limited when it is applied to locations in Europe [21].

Supit and Van Kappel [22] proposed an empirical model by considering the cloud cover of the daytime observations.

$$H = H_0\left[a\sqrt{T_{\max} - T_{\min}} + b\sqrt{1 - \frac{C_w}{8}}\right] + c \tag{3}$$

where C_w is the mean of the total cloud cover of the daytime observations, and a, b and c are empirical constants.

Most of the above models lack the detailed knowledge of various parameters, such as the hourly variation of global radiation, latitude, longitude, altitude, month, time, wind speed, air temperature and humidity. Therefore, in this regard, we have investigated the estimation of solar radiation with an extensive knowledge of these parameters. In the present study, ANN based models have been developed for predicting the monthly mean hourly and daily global solar radiation. The trained and tested ANN models show greater accuracy for evaluating solar energy possibilities in regions where a network of monitoring systems have not been established. The predictions from ANN models would enable locating and designing solar energy systems in India and identifying the best of the solar technologies.

2. Artificial neural network (ANN) based model

Artificial neural network techniques are based on some important facts that have been learned by neuroscientists and others about the nervous system, though this acquired knowledge is far from complete. The idea of neurons as structural constituents of the brain was introduced by Ramon Cajal in 1911 [23]. A schematic of a biological neuron is shown in Fig. 1.

Biological neural networks are non-linear, highly parallel information processing systems that are characterised by robustness, fault tolerance and the ability to learn by adapting the connection strength to changes in the surrounding environment. In biological neurons, electrochemical signals (known as stimuli) are received through synapses to the neuron cell. Each synapse has its own weight that determines in what way and to what extent stimuli coming to the neuron through that synapse affects the output of the neuron. The weighted sum of the input stimuli are fed to the nucleus that, in response to this, sends electrical impulses that are transmitted to other neurons or sent to other biological units as actuation signals. The synaptic weights keep modifying during learning. Neurons are interconnected with large numbers of neurons through synapses. Groups of

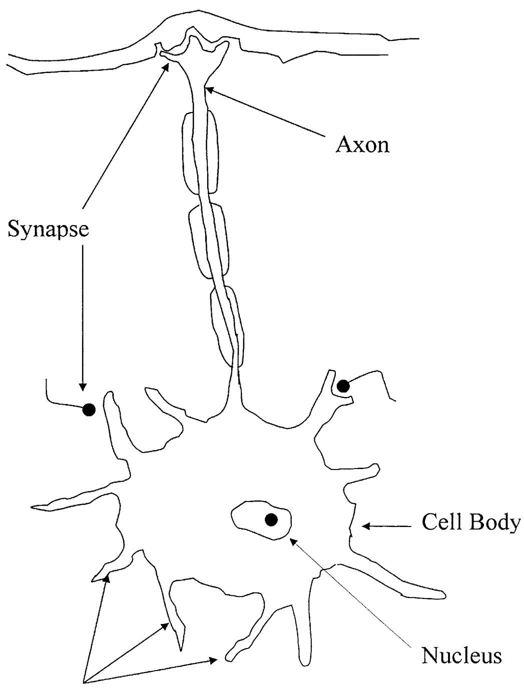

Fig. 1. Schematic of a biological neuron.

neurons are organised into subsystems and the subsystems integrate to form the brain. A simplified model of an ANN is illustrated in Fig. 2.

In the ANN technique, a simulation of a small part of the central nervous system is done wherein stimulation inputs are fed to input neurons (synapses), and these stimuli are altered by weights (synaptic weights). The weighted sum is operated upon by an activation function, and outputs are fed to other neurons in the network. All these neurons are highly interconnected, and the activation values may constitute the final output or may be fed to the next model. These connection weights are modified during training to obtain better and better generalisation and interpolation of training patterns presented to the network during training in order to achieve the desired accuracy by the network. The most suitable architecture and nature of the neurons of the ANN is problem specific.

ANNs have been used for range of objectives, such as constraint satisfaction, content addressable memories, control, data compression, diagnostics, forecasting, general mapping, multisensor data fusion, optimization, pattern recognition etc. Difficulties, due to the uncertain nature of solar radiation, in deriving relations that map reasonably the various spatial, temporal and climatic parameters of solar radiation values and the modelling abilities of ANNs have inspired the application of ANN techniques to determine solar radiation. Modelling of solar radiation on a

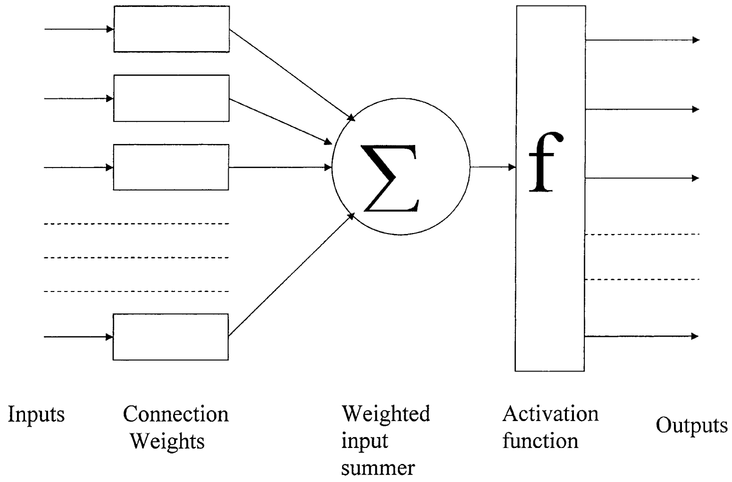

Fig. 2. Block diagram of a simplified model of an artificial neuron.

horizontal surface in the Kingdom of Saudi Arabia has been done by using multi-layer perception networks [24].

3. Multi-layer feed forward (MLFF) network

In MLFF networks, neurons are arranged in layers with connectivity between the neurons of different layers. The layer that receives inputs is called the input layer, and that which gives the output (or output vector) is called the output layer. Other layers, as they do not receive any direct input or contribute to output directly, are called hidden layers. Input signals are propagated in gradually modified form in the forward direction, finally reaching the output layer. The activation function for neurons in a MLFF network can be linear or non-linear. A sigmoid function is a widely used non-linear activation function whose output lies between 0 and 1 and is defined as

$$f(x) = \frac{1}{1 + \mathrm{e}^{-x}} \tag{4}$$

An important characteristic of this function that makes it suitable for use in conjunction with a learning algorithm (the weight modification is done in proportion to the negative gradient of the output) for a MLFF network is that it is differentiable throughout its domain. The error for hidden layers is determined by propagating back the error determined for the output layer; hence the technique is named backpropagation. During learning, the weights of the neurons are optimised according to the Generalized Delta Rule (GDR), which is the learning algorithm for a backpropagation MLFF network. The error that is minimized by the GDR is the sum of the squares of the errors for all the output units, defined as:

$$E_{\mathrm{p}} = \sum_k (y_{pk} - o_{pk})^2 \tag{5}$$

For modification of the weights of the output layer, the direction in which the weights need to be shifted is determined by the negative gradient of E_p with respect to the weight w_{kj}. The adjustments in the weight for each neuron is the product of the error in the neuron's output, the gradient of the neuron's output, the net input given to the neuron and a learning rate parameter. The weight modification for a hidden layer is done in proportion to the gradient of E_p with respect to the hidden layer weights. In this way, each updated weight in a hidden layer is dependent on all the error terms of the output layer. Thus, the errors that could be exactly determined only for the output layer are propagated back to the hidden layers. MLFF learning takes place under supervision, and an important parameter that has a controlling influence is the learning rate constant. It decides the magnitude of changes to the connection weights. A high learning rate constant has the advantage of faster learning, but it may cause the weights to bounce around error minima, thus failing to learn properly. On the other hand, if the learning rate constant is too small, the learning may take a long time because of the slow descent along the error surface, which may be favourable as the network may find a better error minimum and, hence, more accurate learning.

4. Methodology

To estimate the global hourly mean radiation from the minimal data available on the position for any location, a multi-layer feed forward network is trained and tested for its ability to generalise and interpolate. The selected ANN structure (Fig. 3) is a feed forward, fully connected hierarchical network consisting of an input layer, two hidden layers and an output layer. The first hidden layer has eight neurons, and the second hidden layer has seven neurons. There is a single output neuron. Iterative backpropagation with the GDR algorithm has been implemented to determine errors for the hidden layer neurons and subsequent weight modification according to the GDR. In order to avoid undesirably long training time (in the event of inability of network to map the presented pattern of training vectors or network being trapped in local error minima, which causes error more than the acceptable limit etc.), a termination criterion has been adopted. This criterion may be either completion of a maximum number of epochs (training cycles) or achievement of the error goal.

The hourly solar radiation data for training and testing is taken from the Meteorological Department. The data is extracted and formatted in accordance with the ANN demands. The data is then normalised to suit the ANN. The formatted data are matrices of size $[10 \times 280]$ and $[10 \times 336]$ for each of the three seasons (summer, rainy and winter) for North India, and South India, respectively. To train the networks, five cities have been considered from North India, viz. Ahmedabad, Calcutta, Mumbai, Nagpur and Jodhpur and six cities across South India, viz. Bangalore, Kodaikanal, Madras, Port Blair, Vishakhapattnam and Poona. The geographical positions of the cities considered are listed in Table 1. Each city has about 168 data sets around the year. The following input parameters have been considered to estimate the radiation for each city: latitude, longitude, altitude, month, time, air temperature, wind speed, relative humidity and rainfall.

Initially, the network was trained with the total available data.The error in mapping was found to be very high, and the network failed to give reliable results. In order to improve the perfor-

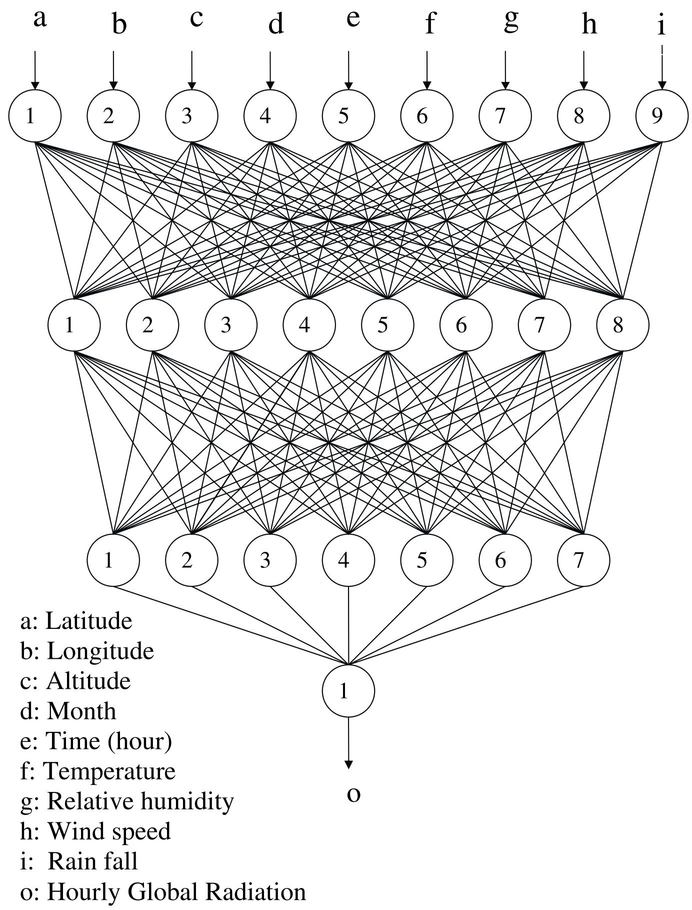

Fig. 3. ANN Architecture used for estimation of hourly global radiation.

mance of the network, it was decided to divide the data. The entire training set was divided based on region (South India and North India) and seasons (summer, rainy and winter), and all these networks were trained and tested individually.

5. Results and discussion

The trained networks were tested for New Delhi (North India) and Mangalore (South India) for each of the three seasons. The mean absolute relative deviations (MARD) of the ANN predicted hourly global radiation are listed in Table 2. MARD is defined as:

$$\mathrm{MARD} = \frac{\sum\left(\left|\frac{\mathrm{Predicted}-\mathrm{Actual}}{\mathrm{Actual}}\right| \times 100\right)}{D} \tag{6}$$

Table 1
Geographical parameters

Location	Latitude (°N)	Longitude (°E)	Altitude (m)
North India			
Ahmedabad	23.07	72.63	55
Calcutta	22.65	88.45	6
Jodhpur	26.30	73.02	224
Mumbai	19.02	72.90	14
Nagpur	21.10	79.50	310
New Delhi	28.58	77.20	216
South India			
Bangalore	12.97	77.58	921
Kodaikanal	10.23	77.47	2339
Madras	13.00	80.18	16
Mangalore	12.91	74.88	102
Port Blair	11.67	92.72	79
Poona	18.53	73.85	559
Vishakhapattnam	17.72	83.23	3

Table 2
Mean absolute relative deviation of ANN estimated monthly mean daily global radiation

Station	Winter	Summer	Rain
New Delhi	2.98	2.24	4.07
Mangalore	2.64	2.37	3.84

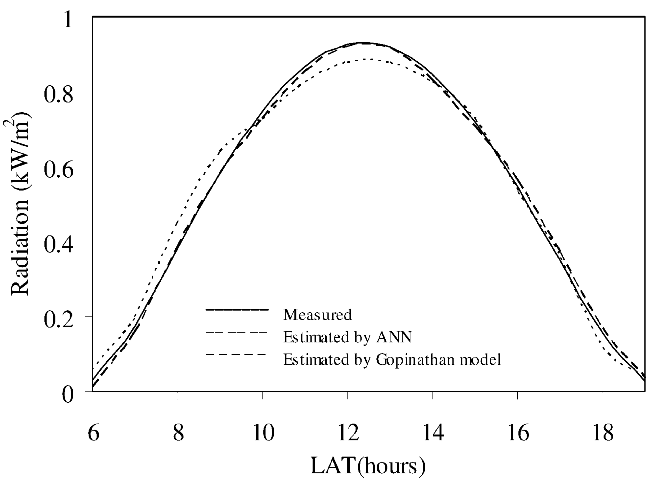

Fig. 4. Comparison of estimated hourly global radiation values.

where D is the number of comparison data. It can be observed that for both cities, the deviations are least in summer and then followed by the winter and rainy season, respectively. The maximum MARD is about 4%.

To test the ability of the ANN, the hourly global radiation estimations of the ANN model have been compared with the Gopinathan [25] model, for a typical summer day at New Delhi. The comparison is shown in Fig. 4. The results show that the present model is superior to the Gopinathan [25] model, which over predicted in the morning of the day and under predicted for mid day hours whereas the predictions from the ANN model are very close to the actual radiation values.

The diurnal variations of global radiation predicted by the ANN model and actual data for New Delhi (North India) and Mangalore (South India) for the three seasons are illustrated in Figs. 5(a)–(c) and 6(a)–(c), respectively. The predicted solar radiation values are very close to the actual values for the summer and winter (Fig. 5(a) and (b)). The maximum absolute error (MAE) values are only 0.021 and 0.016 W/m^2, respectively. A small perceptible deviation is observed for the rainy season (Fig. 5(c)), and the MAE is 0.075 W/m^2. The ANN model for the rainy season has over predicted the radiation values in the middle of the day, whereas the values are in close agreement with actual values in the early hours and late hours of the day. The results for Mangalore add to the confidence in the ability of ANN models to predict radiation values. For the summer, winter and rainy seasons, the MAE values are 0.028, 0.06 and 0.032 W/m^2

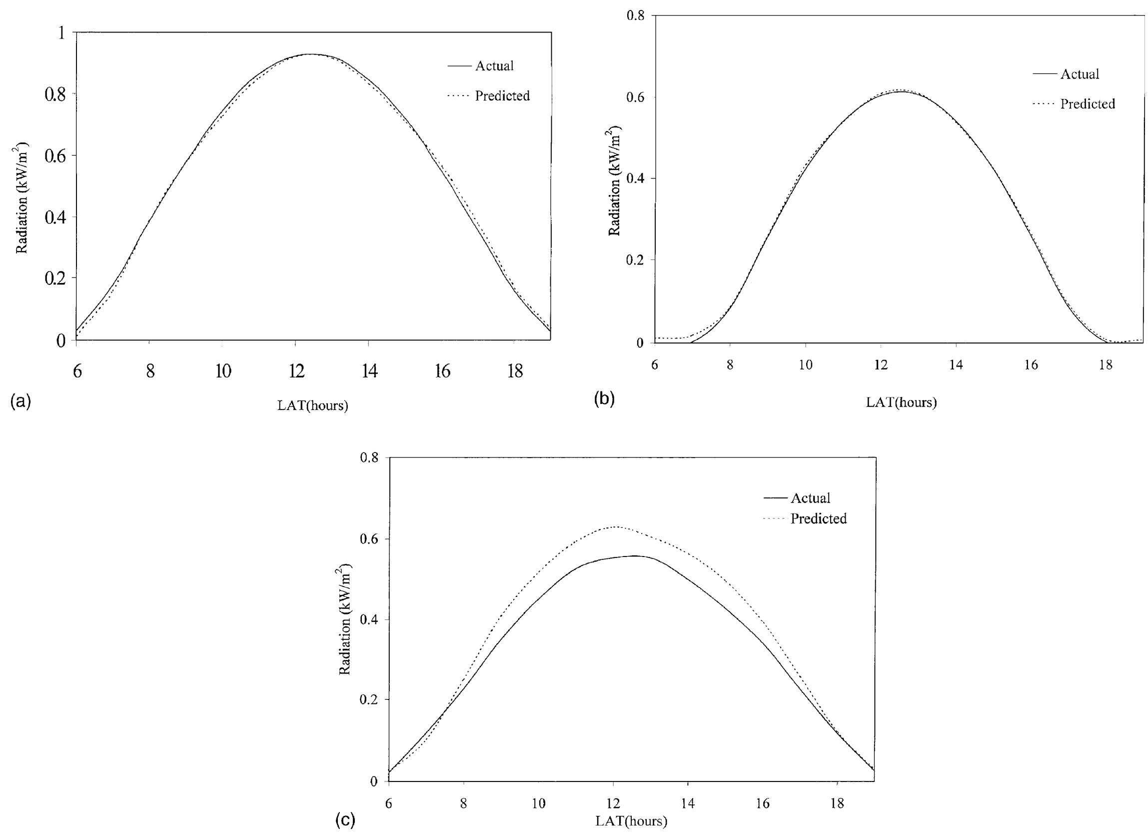

Fig. 5. (a) Variation of hourly global radiation at New Delhi for summer, (b) variation of hourly global radiation at New Delhi for winter, (c) variation of hourly global radiation at New Delhi for rainy season.

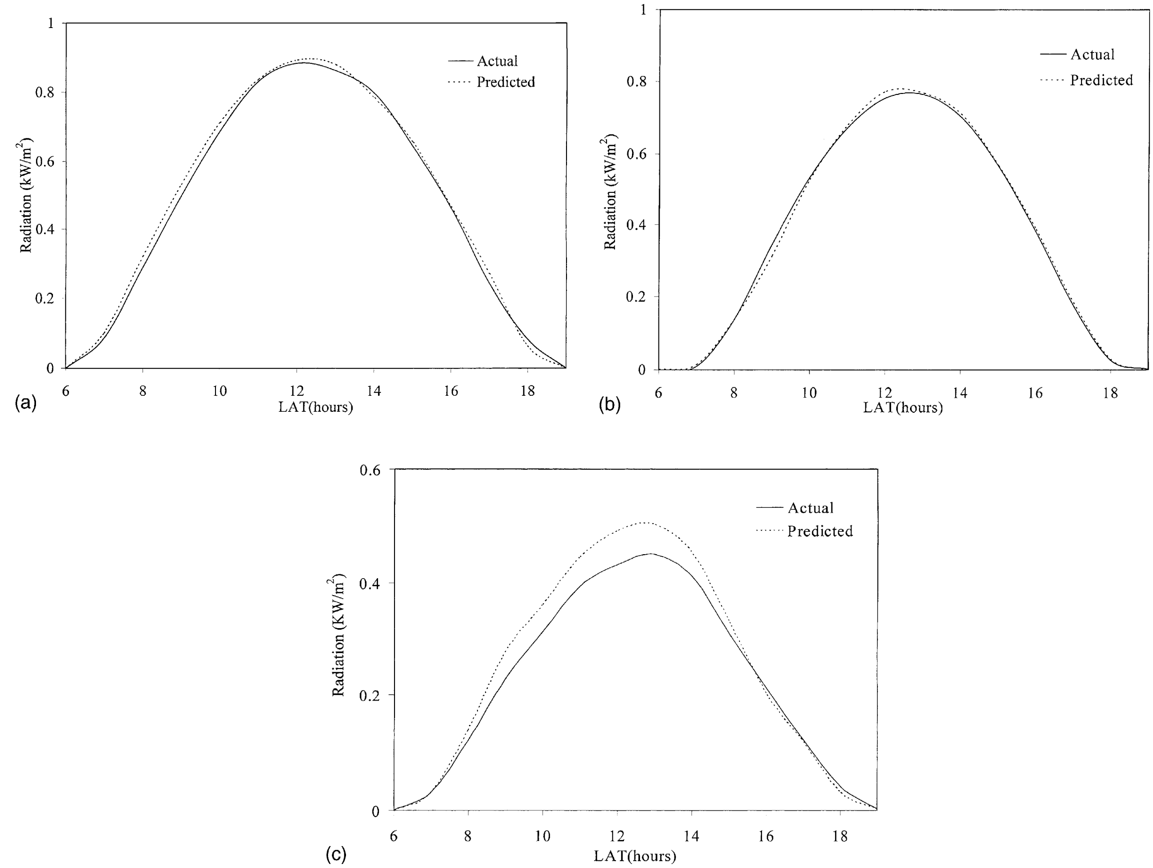

Fig. 6. (a) Variation of hourly global radiation at Mangalore for summer, (b) variation of hourly global radiation at Mangalore for winter, (c) variation of hourly global radiation at Mangalore for rainy season.

respectively. The small deviation from measured hourly global radiation in the rainy season is seemingly because of the higher degree of climatic uncertainty.

Estimations of daily global radiation (DGR) by the ANN model are compared with other empirical models. Predicted DGR values, along with their percentage relative deviation (PRD), are listed in Table 3. The percentage relative deviation is defined as

$$\mathrm{PRD} = \frac{\mathrm{DGR_p} - \mathrm{DGR_a}}{\mathrm{DGR_a}} 100 \tag{7}$$

It is evident that the predictions from the ANN models are better as compared to the other regression models. The percentage relative deviations are limited to 1.8% and 1.9% for summer and winter, respectively, whereas the predictions from the other empirical models are deviating significantly. Because of the climatic uncertainty in rainy seasons, the deviations for the ANN models for Mangalore and New Delhi are 10.2% and 12.5%, respectively, which shows the good prediction ability of the ANN models even when fuzzy information is presented.

Table 3
Comparison of daily global radiation predicted from ANN and various regression models with actual value

Station	Season	Monthly averaged daily global radiation (kW/m^2)								
		Actual DGR	ANN		Angstrom		Hargreeves		Supit	
			DGR	PRD (abs)	DGR	PRD (abs)	DGR	PRD (abs)	DGR	PRD (abs)
Mangalore	Summer	6.368	6.482	1.79	5.434	14.67	5.127	19.49	5.564	12.63
	Rain	3.068	3.381	10.20	4.415	43.9	3.577	16.59	3.665	19.46
	Winter	5.072	5.096	0.47	4.521	10.86	5.358	5.64	4.947	2.46
New Delhi	Summer	7.277	7.279	0.03	6.526	10.32	7.526	3.42	7.665	5.33
	Rain	4.436	4.99	12.5	5.247	18.28	5.651	27.39	5.665	27.71
	Winter	3.825	3.897	1.88	4.679	22.33	3.813	0.31	3.778	1.23

6. Conclusions

Application of the artificial neural network technique for modelling the spatial and temporal variation of global solar radiation has been reported. The results of validation and comparative study indicate that the neural network method is more suitable to predict the solar radiation than various proposed classical regression models. Although the training data lacked wide and uniform geographical coverage and represented only one year from each of the 11 locations, the study confirms the ability of the ANN to predict solar radiation values closely. Inclusion of more identified parameters and data would further improve the models' mapping ability as the ANN adaptations methods depend on learning from examples. These ANN models are more versatile and can be used to predict radiation for any region provided comprehensive meteorological data is available.

References

[1] Mohandes M, Balghonaim A, Kassas M, Rehman S, Halawani TO. Use of radial basis functions for estimating monthly mean daily solar radiation. Solar Energy 2000;68(2):161.
[2] Ogelman H, Ecevit A, Tasdemiroglu E. A new method for estimating solar radiation from bright sunshine data. Solar Energy 1984;33(6):619.
[3] Goh TN. Statistical study of solar radiation in formation in an equatorial region (Singapore). Solar Energy 1979;22:105.
[4] Rietveld MR. A new method for estimation the regression coefficients in the formula relating solar radiation to sunshine. Agric Meteor 1978;19:243.
[5] Barbaro S, Coolino S, Leone C, Sinagra E. Global solar radiation in Italy. Solar Energy 1978;20:431.
[6] Reddy SJ. An empirical method for the estimation of the total solar radiation. Solar Energy 1971;13:289.
[7] Swatman RK, Ogunladeg O. Correlation of solar radiation with common parameters in Toronto, Canada. Solar Energy 1971;13:345.
[8] Angstrom A. On the computation of global solar radiation from records of sunshine. Arkiv Geophysik 1956;3(23):551.
[9] Mani A, Chacko O. Solar radiation climate of India. Solar Energy 1973;14:139.
[10] Mani A. Handbook of solar radiation data. Allied Publishers; 1981.

[11] Maure D, Galanis N. Solar radiation data for Quebec. Solar Energy 1979;23:309.
[12] Hay JE. Calculation of monthly mean solar radiation for horizontal and inclined surfaces. Solar Energy 1979;23(4):301.
[13] El-Nashar AM. Solar radiation characteristics in Abu Dhabi. Solar Energy 1991;47(1):49.
[14] Gopinathan KK. Solar radiation on inclined surfaces. Solar Energy 1990;45(1):19.
[15] Gopinathan KK. Solar radiation on variously oriented sloping surfaces. Solar Energy 1991;47(3):173.
[16] Sabbagh JA, Sayigh AAM, El-Salam EMA. Estimation of total solar radiation from meteorological data. Solar Energy 1977;19:307.
[17] Kimball HH. Variations in the total and luminous solar radiation with geographical position in the United States. Mon Weather Rev 1919;47:769.
[18] Angstrom A. Solar and terrestrial radiation. QJR Meteorol Soc 1924;50:125.
[19] Prescott JA. Evaporation from a water surface in relation to solar radiation. Trans R Sec South Australia 1940;64:148.
[20] Hargreeves GL, Hargreeves GH, Riley P. Irrigation water requirement for the Senegal river basin. J Irrigation and Drainage Eng ASCE 1985;111:265.
[21] Choisnel E, de Villele O, Lacroze F. Une aproache uniformesee du calcul de levaporation potentielle por lensemble des pays de la communaute europee nne. Publication of the EU, Luxembourg, 1992.
[22] Supit I, Van Kappel RR. A simple method to estimate global radiation. Solar Energy 1998;63:147.
[23] Haykin S. Neural Networks: A Comprehensive Foundation. second ed. Macmillan College Publishing; 1998.
[24] Mohandis M, Rehman S, Halwani T. Estimation of global solar radiation using artificial neural networks. Sixth Arab Int. Solar Energy Conference, Muscat-Sultanate of Oman, 1998.
[25] Gopinathan KK. Estimation of hourly global and diffuse solar radiation from hourly sunshine duration. Solar Energy 1992;48(1):3.