Theor. Appl. Climatol. 66, 185–197 (2000)

Theoretical and Applied Climatology

Laboratory of Meteorology, Division of Applied Physics, Department of Physics, University of Athens, Greece

The total solar radiation time series simulation in Athens, using neural networks

G. Mihalakakou, M. Santamouris, and **D. N. Asimakopoulos**

With 10 Figures

Received November 22, 1999
Revised February 17, 2000

Summary

The present study describes a neural network approach for modeling and making short-term predictions on the total solar radiation time series.

The future hourly values of total solar radiation for several years are predicted, by extracting knowledge from their past values, using feedforward backpropagation neural networks. The results are tested using various sets of non training measurements, the findings are very encouraging and the model is found able to simulate the future values of total solar radiation time series based on their past values. “Multi-lag” output predictions are performed using the predicted values to the input database in order to model future total solar radiation values with sufficient accuracy. Furthermore, an autoregressive model is developed for analysing and representing the total solar radiation time series. The predicted values of solar radiation are compared with the observed data series and it was found that the neural network approach leads to better predictions than the AR model.

1. Introduction

In recent years, there has been a growing interest in applying intelligent techniques to time series prediction. The intelligent techniques such as neural networks of fuzzy logic methods can be designed and used for predicting the future values of a non linear dynamic process on the basis of collected historical data (Farmer and Sidorowich, 1987; Teodorescu, 1990; Cichocki and Unbehauen, 1993; Pham and Liu, 1995).

In fact, predicting the future values of a time series is a very old and general problem which has potential applications in several fields working with time series such as atmospheric sciences, economic and engineering applications, climatology, etc. The most important means of predicting the future can be presented as follows:

a. One of the most powerful and accurate approaches of time series prediction is the development of analytic models based on the knowledge of a law underlying the analytic given dynamic process or phenomenon (Weigend et al., 1990). These models are able to give sufficiently accurate estimations, provided that such a law can be discovered and described using, for example, a set of differential equations or parameterised expressions. However, the development of such an analytic model describing a dynamic process is often a very difficult task taking into account that the information about a dynamic phenomenon is often partial and incomplete, the initial and boundary conditions of the problem cannot always be completely and clearly specified and sometimes the analytic model requires a large number of non available input parameters.
b. Another less powerful approach is the discovery of several strong empirical regularities or periodicities during the observation of the

dynamic process (Farmer and Sidorowich, 1988; Cichocki and Unbehauen, 1993). However, the empirical regularities or periodicities are not always evident as they are masked by noise (Chakraborty et al., 1992). Moreover, many real world processes are described by chaotic time series for which long-term predictions are not possible since the uncertainty of the prediction increases exponentially with time.

c. There are various stochastic models for time-series analysis, such as the autoregressive models, which assume linear processes. The autoregressive models are predictive models describing the available data and they represent a realization of the process to be simulated using a suitable set of parameters, so that this set becomes representative of the process itself (Box and Jenkins, 1970). Autoregressive models are attractive because they are simple to use. They come from a large family of time series models called autoregressive, moving-average models, (ARMA), (Leite and Peixoto, 1996; Lalarukh and Jafri, 1997). Autoregressive stochastic models are essentially linear models, simple and understandable but incapable of simulating the nonlinear nature of various dynamic processes of the real world.
d. Fourier techniques have also extensively been used to analyse and predict time series (Carson, 1963; Lamba and Khambete, 1991; El-Shal and Mayhoub, 1995).
e. Although chaos prevents a long-term predictability, a short-time forecasting is possible and very promising results have been obtained by using an intelligent technique such as neural networks of fuzzy logic methods for nonlinear modeling of multivariate time-series (Li et al., 1990). Artificial neural networks are computational models which can be regarded as an attempt to simulate in a simpler way the structure and functions of the human brain. Neural networks belong to the class of "data-driven" approaches, instead of "model-driven" approaches because the analysis and the results depend on the available data (Chakraborty et al., 1992). Relationships between variables, models, laws and predictions are constructed after building a machine which simulates the considered data. The process of constructing such a machine based on available data is addressed by certain algorithms like "perceptron" (Rosenblatt, 1961) or "backpropagation" (Rumelhart et al., 1986). Various researchers proposed learning algorithms for time series prediction and they applied them to feedforward multilayered or recurrent neural networks (Wong, 1991; Hondou and Sawada, 1994; Connor et al., 1994; Eisenstein et al., 1995; Dash et al., 1995; Kalogirou et al., 1997).

In the present study a neural network approach is used for the prediction of future values of total solar radiation time-series for several years. Solar radiation measurements are rather sparse and for this reason theoretical deterministic models and intelligent data driven approaches are developed to predict the available solar radiation. Various atmospheric deterministic models have been developed for the prediction of solar radiation using as inputs several meteorological parameters such as air temperature, relative humidity, sunshine duration and cloudiness. These parameters are usually measured extensively at all the meteorological stations.

In Santamouris et al. (1999) three models, one deterministic atmospheric model and two intelligent data driven models, estimating the total short-wave radiation using as inputs several meteorological parameters, were presented and compared. The atmospheric model is an analytical approach, based on parameterised expressions, which requires as inputs several climatological parameters such as air temperature, relative humidity, sunshine duration, cloudiness, surface albedo etc. This model was tested and found able to give sufficiently accurate estimations provided that all the required input parameters are available. The intelligent data driven approaches were a neural network model and a fuzzy logic method. From the comparison of the two intelligent methods results with the results of the atmospheric model, it was observed that the performance of the two data driven models is very satisfactory especially for the summer period.

The main objectives of the present study are to design a new neural network approach for predicting the future values of total solar radiation time series, to examine its ability to model and forecast the total solar radiation and to investigate its limitations. The applicability of the neural network system in "one-lag" and "multi-lag" output radiation predictions is examined and discussed. Furthermore, the feasibility of the system is investigated by comparing its results with the results of a conventional linear autoregressive time series prediction model. The advantages and disadvantages of the neural network approach as well of the autoregressive model are presented and discussed. The paper is organised as follows: In the first section there is an analytical description of the neural network architecture. The second section contains the description of the data base, the presentation of the neural network methods's predictive ability as well as the results from a comparison between the neural network method and a linear autoregressive model. Finally, the conclusions are given in the last section.

2. Modeling the total solar radiation

Artificial neural networks are computing systems which attempt to simulate the structure and function of biological neurons. Neural networks generally consist of a number of interconnected processing elements or neurons. How the interneuron connections are arranged and the nature of the connections determine the structure of a network. Neural networks can be classified according to their structures into the following two types (Pham and Liu, 1995):

Feedforward networks: In a feedforward network, the neurons are generally grouped into layers. Signals flow always from the input layer through to the output layer via unidirectional connections, the neurons being connected from one layer to the next, but not within the same layer.

Recurrent networks: In a recurrent network, the outputs of some neurons are feedback to the same neurons or to neurons in preceding layers. Therefore, signals can flow in both forward and backward directions.

A multi-layer feedforward neural network is shown in Fig. 1. The network consists of three layers: an input layer, an output layer and an intermediate or hidden layer. The neurons in the input layer act only as buffers for distributing the input signals to the neurons in the hidden layer. The dashed lines in Fig. 1 mean that there are more neurons in each layer than the represented in this figure. Figure 2 shows the basic artificial neuron of the hidden layer.

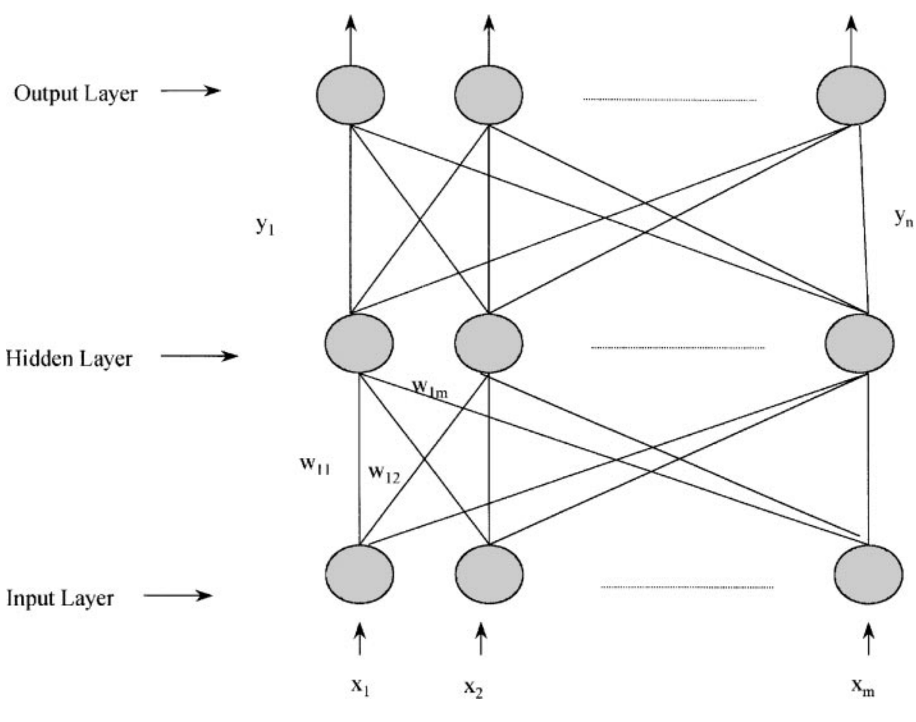

Fig. 1. Architecture of a neural network system

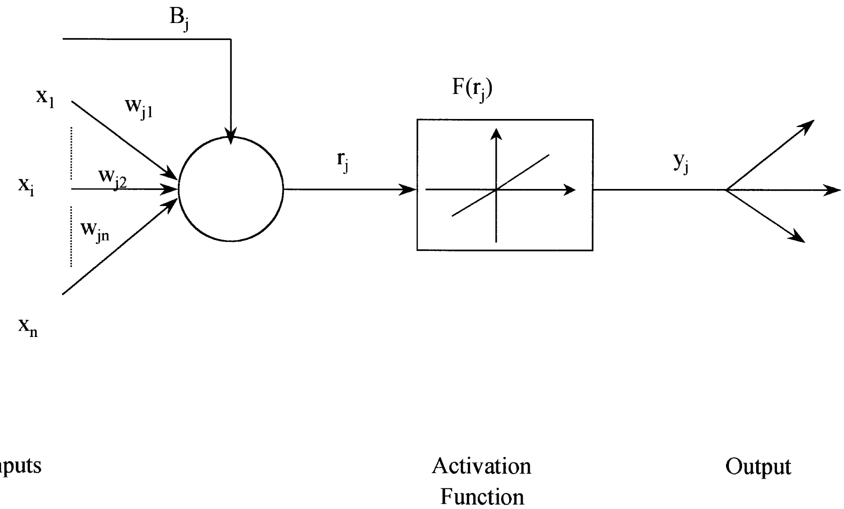

Fig. 2. Presentation of a basic artificial neuron

The time series prediction problem using a neural network approach can be separated into three successive steps or subproblems:

- model building or neural network architecture
- the learning or training process
- the testing or diagnostic checking

In the present study a multi-layer feedforward network based on backpropagation learning procedure is designed for predicting the total solar radiation time series. This type of neural network is extensively used in the time series prediction. The selected neural network architecture consists of one hidden layer of 16 log-sigmoid neurons followed by an output layer of one linear neuron. This scheme was selected after trying several ones as it gave the better convergence.

Learning is achieved using the backpropagation algorithm of Rumelhart et al. (1986) to train the network.

A learning rate of 0.3 and an error goal of 0.5 were selected while the number of epochs varied between 3000 and 4000 in all cases.

3. Results and discussion

3.1 Data base

The time series generated in the present paper are total solar radiation measured on an horizontal surface at the Institute of Meteorology and Physics of the Atmospheric Environment of the National Observatory of Athens. The (IMPAE/NOE) Institute is situated on a hill at the centre of Athens (37.967° N, 23.717° E, altidude = 107 m). Continuous observations of standard meteorological parameters have been performed at this location, the close surroundings of which remain unaltered, since 1864.

Integrated hourly, daily and monthly values of total solar radiation in MJ/m^2 are measured at the Observatory with Kipp-Zonen and Eppley actinometers and pyranometers respectively. Hourly values of total solar radiation for twelve years (1984–1995) and for various months of the year are used for training and testing the network. The data are divided into a training set which provides a fitting approximation function and a testing set which is used to validate the ability of prediction of the already trained network (Weigend et al., 1990). Nine years (1984–1992) were used for training the neural network and three years (1993, 1994 and 1995) for testing the training data. The night-time values of total solar radiation, which probably are zero values, are omitted from the training and testing sets and therefore they are not used in the training and testing processes.

Before training, it is useful to scale the inputs and targets by using a pre-processing procedure. In the present application the used approach for scaling network inputs and targets is to normalise the mean and standard deviation of the training set. This procedure normalises the inputs and targets so that they will have zero mean and unity standard deviation.

3.2 "One lag" predictions

In the "one lag" output prediction the future values of the total solar radiation time series are based only on actual past values. The neural network is designed to predict the next value of the time series given a number of past values of the series. In this case, the network is able to provide only one step-ahead prediction ("one lag" prediction).

The past eight values of the total solar radiation time series are used for predicting the next ninth value by the previously described neural network (Elsner, 1992; Mihalakakou et al., 1998). For example, if the series was represented as $G(t_i)$ where $i = 1, 2, \ldots, N$ and the first set of inputs is $\{G(1), G(2), \ldots, G(8)\}$, then the predicted output would be G_9. Similarly, the second set of inputs is $\{G(2), G(3), \ldots, G(9)\}$ and the predicted output is G_{10}. The selection of the used past values number for the prediction of the next radiation value was decided after various trial runs using five to ten values. The runs indicated that the network efficiency, regarding the obtained outputs, increased when five up to eight past values were used as inputs for the radiation prediction, while for nine or more values, the efficiency remained unchanged. The network efficiency increases as the mean squared error between the measured and predicted data becomes smaller. Training continues over all training pairs for several thousand iterations. The network is asked to predict the next value in the time sequence. The error between the value predicted by the network and the measured one is then measured and propagated backwards along the feedforward connections. The training process is performed on hourly total solar radiation data for the first nine years of the whole set of measurements (1984–1992) and for various months of the year. As learning occurs the sum-squared error decreases. Results from trial runs indicated that adding more hidden layers or nodes did not significantly improve the network's prediction capabilities, rather only slowed the convergence.

Calculations are performed for all the months of the considered data base and the following two time periods are selected for the presentation of results:

- The cold period, which consists of the months of November, December, January and February and March. The months of January and February are regarded as representative of the cold period for the results' presentation.
- The warm period of the year, which consists of the months of May, June, July, August and September. Accordingly, the months of July and August were considered to be the representative months of the warm period for the presentation of results.

The temporal variation of the predicted and measured total solar radiation values for four randomly selected cases containing three continual days of July 1989, three continual days of August 1990, three continual days of January 1990 and three continual days of February 1987 is shown in Fig. 3. The first case consists of the 5th, 6th and 7th of July 1989, the second case of the 18th, 19th and 20th of August 1990, the third of the 2nd, 3rd and 4th of January 1990 and the fourth of the 14th, 15th and 16th of February 1987. In this figure, the continual line indicates the measured total solar radiation values while the cross symbols indicate the model's predictions. As shown, there is a good agreement between the predicted and the measured data. Quite similar performance was observed for the whole training set of data.

Figure 4 shows the comparison of the measured integrated hourly total solar radiation time series values with the neural network predicted ones for two years from the training set of data, (1986, 1990), for the months of July and August and for two other years from the training set of data, (1985 and 1989), for January and February. As it can be seen from Fig. 4, the predicted values perform well with the measured ones. For most cases, the radiation differences are less than $0.20\,\mathrm{MJ/m^2}$ while the root mean squared error between the measured and the estimated values is found equal to $0.18\,\mathrm{MJ/m^2}$ for the month of July, $0.17\,\mathrm{MJ/m^2}$ for the month of August, $0.21\,\mathrm{MJ/m^2}$ for the month of January and $0.23\,\mathrm{MJ/m^2}$. The correlation coefficient between measured and predicted values is, in most cases and for the whole set of training data, better than 0.95. In Fig. 4, the correlation coefficient was 0.97 for the month of July, 0.96 for the month of July, 0.93 for the month of January and 0.94 for the month of February.

The neural network's predictions were checked by comparing its results with the actual values of a testing set of data which consists of the hourly total solar radiation measured values for the years

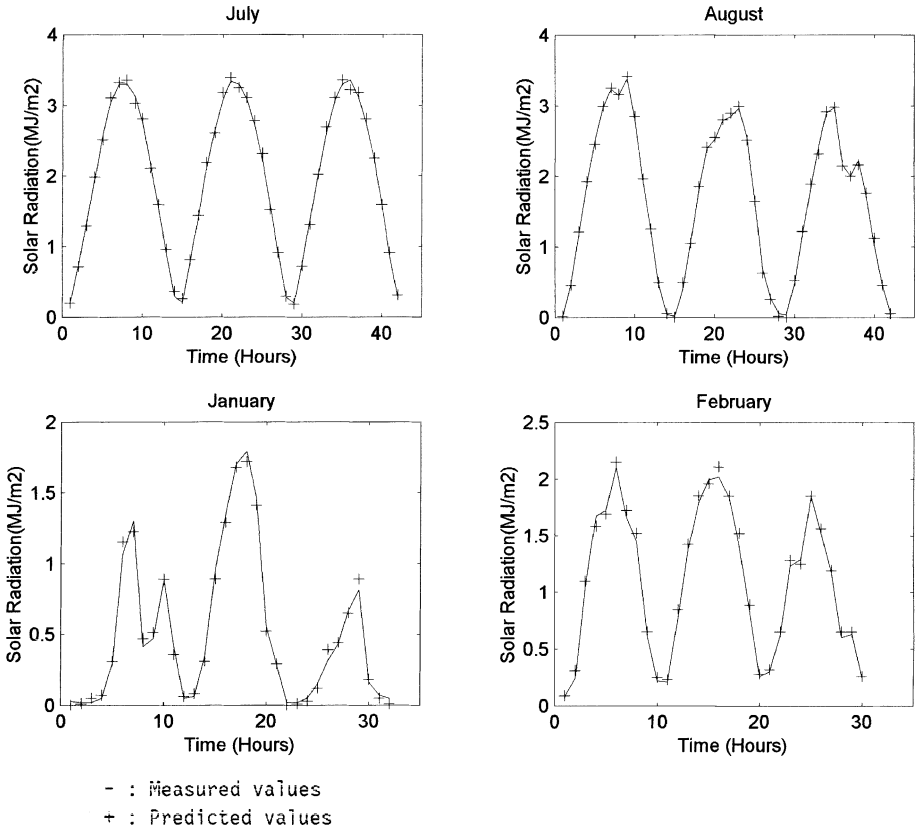

Fig. 3. Temporal variatioin of the predicted with the neural network and of the measured total solar radiation values for one randomly selected case of three continual days of July 1989 (5^{th}, 6^{th}, and 7^{th}), of August 1990 (18^{th}, 19^{th}, and 20^{th}), of January 1990 (2^{nd}, 3^{rd}, and 4^{th}), and of February 1987 (14^{th}, 15^{th}, and 16^{th})

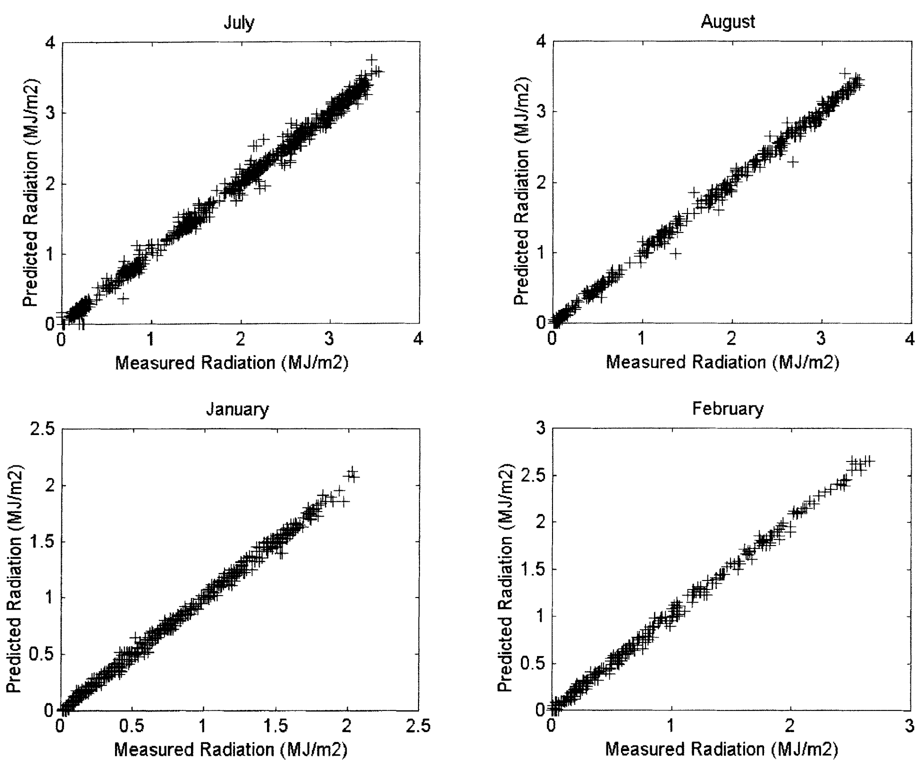

Fig. 4. Comparison of the measured with the neural network predicted total solar radiation values for two years from the training set of data for the months of July, August, January and February

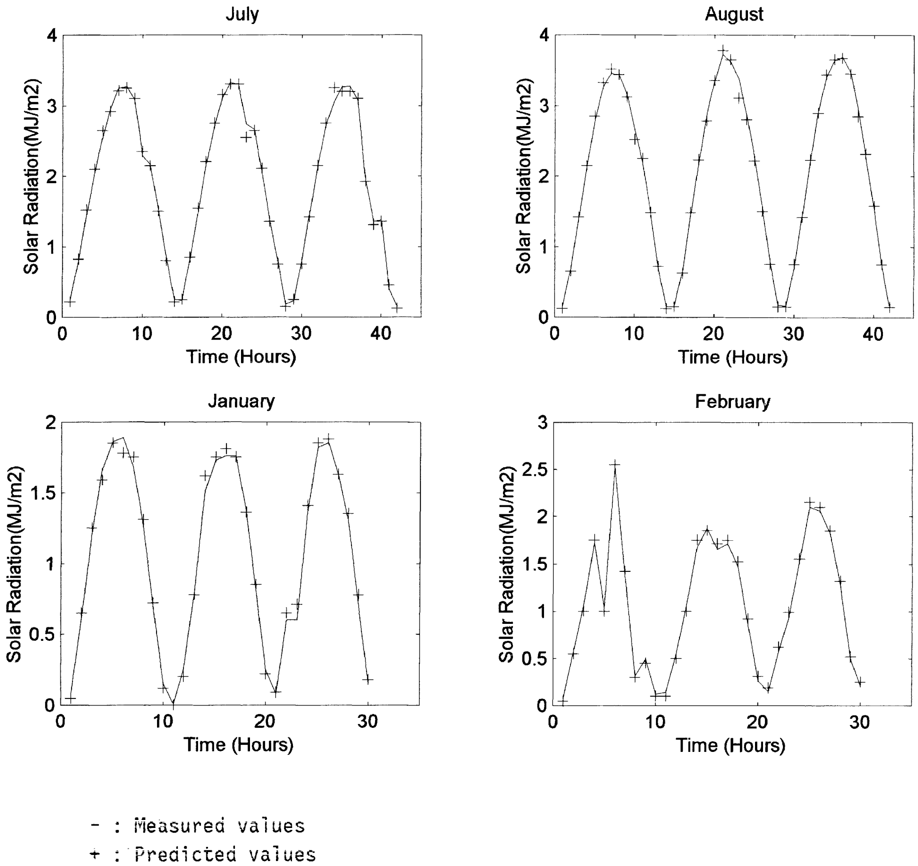

Fig. 5. Temporal variation of the predicted total solar radiation values and of the testing set measurements for three continual days of July 1993 (1^{st}, 2^{nd}, and 3^{rd}), of August 1994 (13^{th}, 14^{th}, and 15^{th}), of January 1995 (2^{nd}, 3^{rd}, 4^{th}), and of February 1995 (9^{th}, 10^{th}, 11^{th})

1993, 1994 and 1995. The temporal variation of the predicted from the network total solar radiation values and of the testing set measured values for three randomly selected continual days of July 1993 (1^{st}, 2^{nd}, 3^{rd}), of August 1994 (13^{th}, 14^{th}, 15^{th}), of January 1995 (2^{nd}, 3^{rd}, 4^{th}), and of February 1995 (9^{th}, 10^{th}, 11^{th}) are presented in Fig. 5. Again, the continual line indicates the measured total solar radiation values while the cross symbols indicate the model's predictions. As it can be seen the neural network predicted values perform well on the testing set of measurements. Quite similar performance has been observed for the whole set of the testing data. Figure 6 shows the comparison between the predicted hourly values and the measured values of total solar radiation for the testing set of data. The months of July and August of 1994 as well the months of January and February 1995 are used in Fig. 6 for the presentation of the results. The mean squared errors are found equal to $0.22\,\mathrm{MJ/m^2}$ for July, $0.19\,\mathrm{MJ/m^2}$ for August, $0.25\,\mathrm{MJ/m^2}$ for January and $0.23\,\mathrm{MJ/m^2}$ for February. The correlation coefficients are 0.94 for July, 0.92 for August, 0.90 for January and 0.91 for the month of February. Similar performance is observed for the whole set of the testing data. The present results are very encouraging and the neural network approach is found able to simulate and predict the future values of total solar radiation time series with sufficient accuracy.

3.3 "Multi-lag" predictions

In order to achieve a prediction several steps into the future, ("multi-lag" prediction), the predicted output is fed back to the input for the next prediction and the other input data are shifted back one unit and so on. For this purpose, a value is predicted one step into the future and then this predicted value is used as one of the lagged inputs for the next prediction two time steps into the future. Similarly, the predicted value at this second time step, as well as the previous time step are used a lagged inputs for the next prediction three time steps into the future. This is the "multi-lag" prediction where the predicted values are appended to the network input database

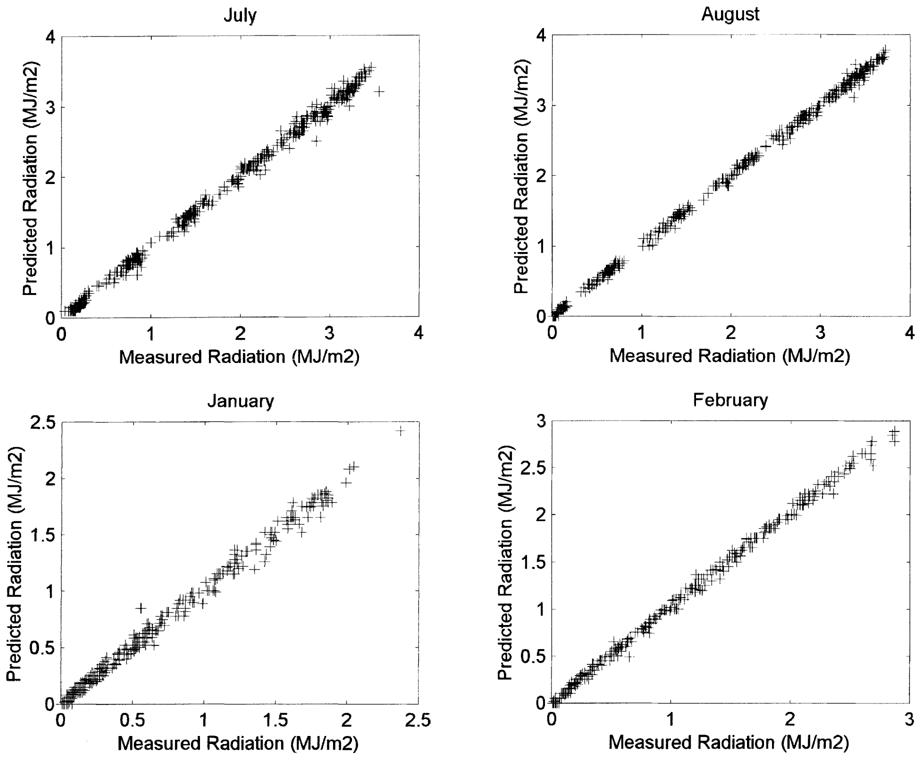

Fig. 6. Testing of the neural network using the actual radiation values for July 1994, for August 1994, for January 1995 and for February 1995

and used to predict future values. For instance, the network is used to predict the seventh value G_7 from the measured total solar raditaion values $G(1), \ldots, G(6)$, then the next neural network prediction G_8 is made using as inputs $G(2), \ldots, G(6), G_7$, and the subsequent network prediction G_9 is made using the radiation values $G(3)$, $G(4)$, $G(5)$, $G(6)$, G_7 and G_8. In the beginning, six measured solar radiation values were used as inputs for the first prediction as in the "one-lag" prediction, and the "multi-lag" prediction results were not satisfactory. Furthermore, the number of measured values used as inputs for the prediction of the first output is increased in order to improve the performance of the method and the six input values became seven, eight, etc. Thus, it is estimated that the more radiation values are appended as inputs for the prediction of the first output, the better the longer-term predictions can be made. Six to thirteen past solar radiation values are used as inputs for the prediction of the first output. This work has been done for nearly every month of the whole set of testing data (1993, 1994 and 1995) and it was found that using the radiation measured values for thirteen hours, it is possible to predict with sufficient accuracy, the total solar radiation values ten to twenty days in advance for all the summer months.

In Fig. 7 the temporal variation of the measured and predicted total solar radiation values for a randomly selected case of 14 days of July 1995 (18^{th}–31^{st} of July 1995), using 13 hourly radiation measurements as inputs for the first output prediction, is presented. As shown from this figure, for the first 10 days time period the predicted values are in close agreement with the measured data. Figure 8 shows the temporal variation of the relative error (%RE) between measured and predicted total solar radiation values for the fourteen days of July 1995 shown in Fig. 7.

$$(\%\mathrm{RE}) = [(\mathrm{R}_{\mathrm{meas}} - \mathrm{R}_{\mathrm{pred}})/\mathrm{R}_{\mathrm{meas}}]^{*}100$$

where, $\mathrm{R}_{\mathrm{meas}}$ and $\mathrm{R}_{\mathrm{pred}}$ are the measured and predicted from the neural network model total solar radiation values respectively. For the first ten days, 90% of the (%RE) values fall between −10% and 15%. For the eleventh day, 90% of the (%RE) values fall between −23% and 31%. For the twelfth day 90% of the (%RE) values fall between −49% and 53% while for the thirteenth day between −80% and −69% and for the fourteenth day the 90% of the (%RE) values fall between −123% and 70%, (Fig. 8).

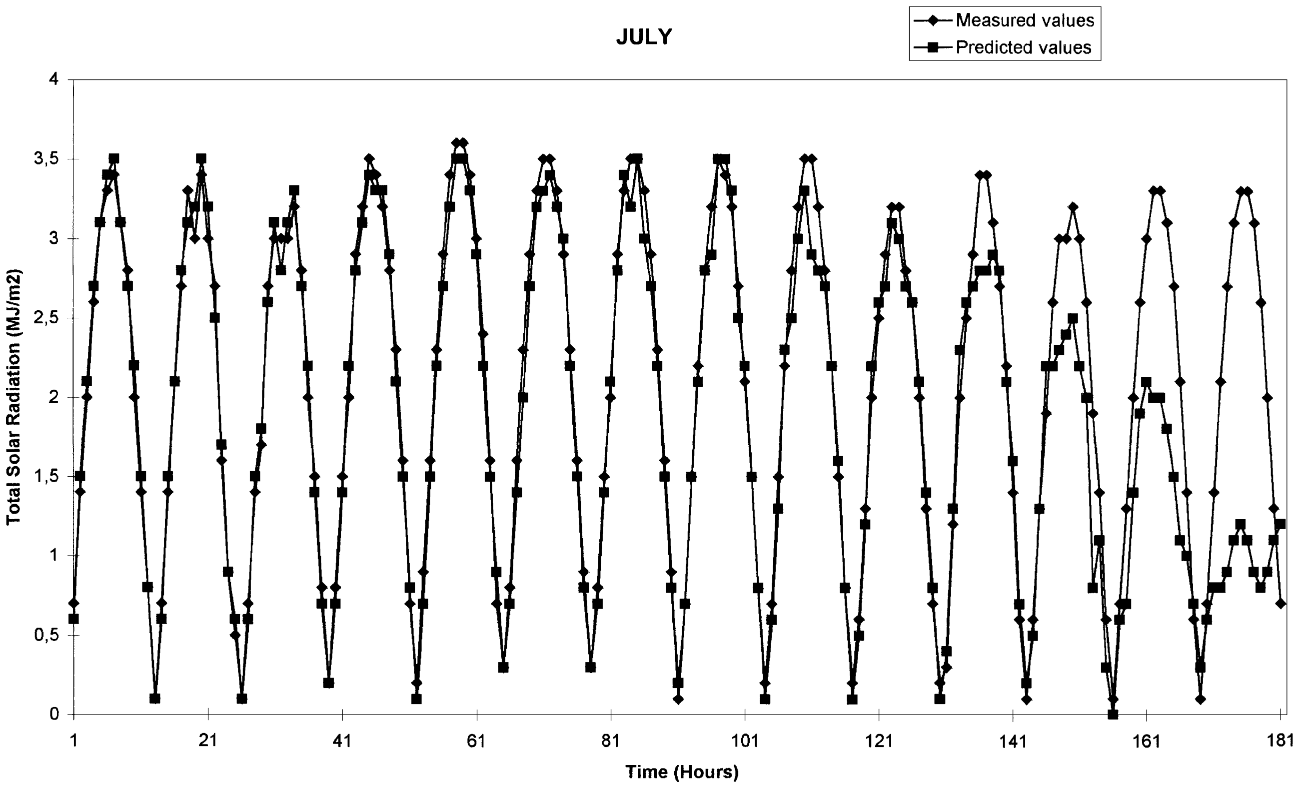

Fig. 7. Temporal variation of the measured and predicted solar radiation values for 14 days of July 1995, (18^{th}–31^{st}), using 13 hourly measurements as inputs for the prediction of the first output

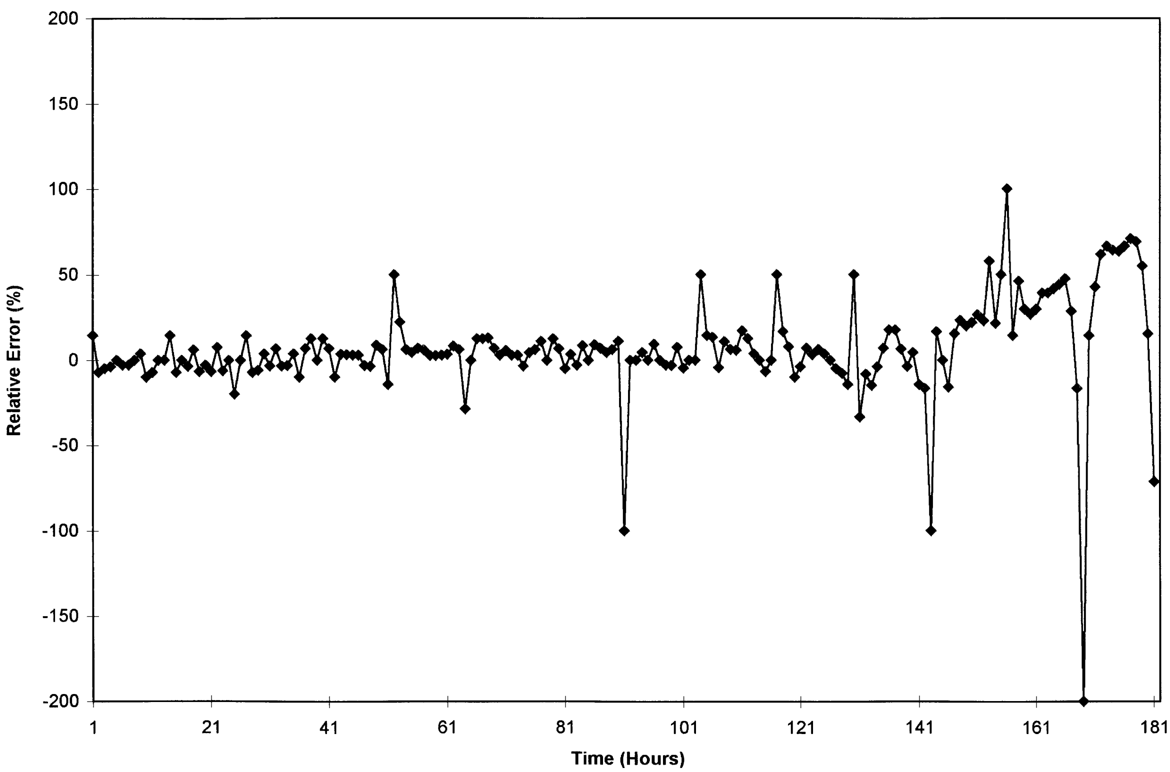

Fig. 8. Temporal variation of the relative error (%) between measured and predicted total solar radiation values for the 14 days of July 1995, (18^{th}–31^{st})

Table 1. Temporal variation of the root mean squared error (%) between the measured and the predicted total solar radiation values for the 14 days of July 1995, (18^{th}–31^{st})

Day	Root mean squared error
1	6.05
2	8.31
3	7.30
4	15.64
5	11.07
6	7.76
7	28.60
8	14.42
9	16.63
10	16.05
11	31.19
12	38.97
13	65.08
14	79.02

Table 1 presents the (%) root mean squared errors calculated between the measured and the predicted total solar radiation values for the 14 days of July 1995 shown in Fig. 7. As shown, the predicted results are better for the first ten days while the predictive ability of the network drops since the error (uncertainty of the prediction) increases exponentially with time. Similar results have been achieved for August 1993, (13^{th}–29^{th}), using 13 hourly radiation measurements as inputs for the first output prediction. In this case, for the first 14 days time period the predicted values are in close agreement with the measured data.

Using the present neural network system it is not always possible to predict the total solar radiation for fifteen days in any case and under any climatic conditions. However, various climatological parameters such as the total solar radiation or the air temperature can be predicted for small periods with very promising results using a nonlinear method such as neural network approaches. In the present research the predicted period is mainly concentrated in the summer months. Summer months are characterised as warm and very dry in the Mediterranean area, and they consist of clear and sunny days, usually without weather phenomena. So, solar radiation time series can be modeled and quite longer time predictions can be achieved for this period of the year.

For winter months, the results are not similarly encouraging. Figure 9 shows the temporal variation of the measured and "multi-lag" predicted total solar radiation values for 8 days of January 1995 (7^{th}–14^{th}), using 9 hourly radiation measurements as inputs for the first output prediction. As shown, for the seventh, eighth, ninth and thirteenth and fourteenth of January the predicted values perform well with the measured data. These days can be regarded as clear days. The results are not satisfactory for the tenth, eleventh, and twelfth of January which are considered to be cloudy days. This is explained by the fact that the network predictions are based on the past 9 solar radiation values and cannot predict the solar radiation in various random weather conditions such as cloudiness or storms, phenomena very representative of the cold period of the year.

3.4 Comparison of the neural network model results with the results of a linear autoregressive system

In order to validate the predictive ability of the above described neural network model, its results were compared with the corresponding results of a linear autoregressive model. A general linear model for system identification can be described by the following equations (Ljung, 1987; Soderstrom and Stoica, 1989):

$$y_p(k) = -\sum_{i=1}^{n_a} a_i y(k-i) + \sum_{j=1}^{n_b} b_j u(k-j) + \sum_{l=1}^{n_c} c_l e(k-l) \tag{1}$$

$(k = 0, 1, 2, \ldots)$

where $u(k)$ is the scalar input signal, $y(k)$ is the scalar output signal, $e(k)$ is the disturbance (or model error) and $y_p(k) := y(k) - e(k)$ is the output of the model at the time instant $t_k = k\tau$ (τ is the sampling period, k is the time index or the cycle number). In this general model the present output signal $y(k)$ is estimated as a linear combination of the past inputs and outputs and the past samples of noises. This model is well known as an ARMAX model (autoregressive moving average with an exogenous signal, (Cichocki and Unbehauen, 1993).

The standard autoregressive model (AR-model) used in the present study is a special case of the ARMAX model and is obtained when $n_b =$

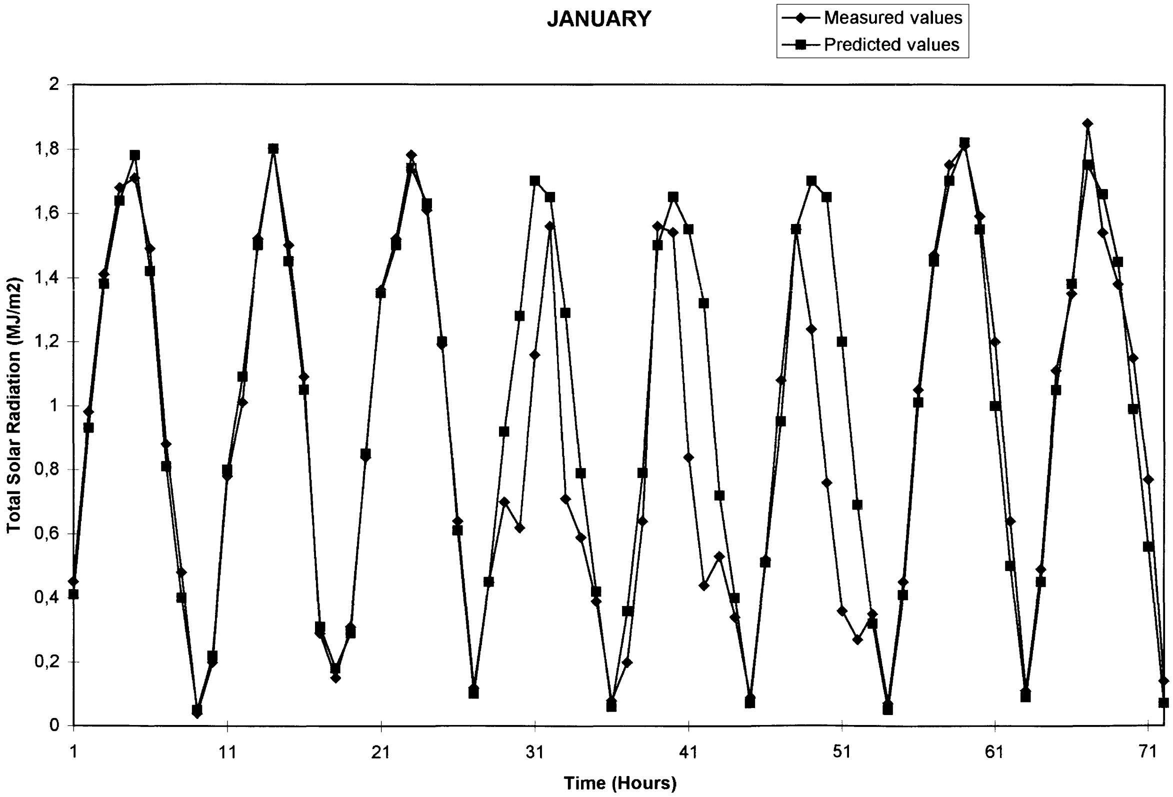

Fig. 9. Temporal variation of the measured and predicted solar radiation values for 8 days of January 1995, (7th–14th), using 9 hourly measurements as inputs for the prediction of the first output

$n_c = 0$. The main characteristic of this model is that the predicted value $y_p(k)$ of the next step is given in terms of a linear combination of a fixed number of past values of the time series. Therefore for the AR model, the Eq. (1) is written as follows:

$$y_p(k) = -\sum_{i=1}^{n_a} a_i y(k-1) \tag{2}$$

In this model only past values of the output are used to predict its present value $y(k)$. The same data base as in the neural network approach are also used for fitting the solar radiation data with the AR method and again the zero night-time values of total solar radiation are omitted. The predicted results are compared with the testing set of measured values. Correlation coefficients between measured and predicted data have been used for the presentation of results.

The pearsonian linear correlation coefficient between two variables (series) X and Y has been used in the present study, usually denoted by r_{xy} or simply r. It expresses a numerical measure of linear relationship between the two variables and is defined as the ratio of the covariance between X and Y, $\mathrm{Cov}(x, y)$, to the product of the standard deviations of X and Y.

$$\begin{aligned} r &= \mathrm{Cov}(x, y)/\sigma_x\sigma_y \\ \mathrm{Cov}(x, y) &= 1/n\Sigma(x-\bar{x})(y-\bar{y}) \end{aligned} \tag{3}$$

If (x_1, y_1), $(x_2, y_2), \ldots, (x_n, y_n)$ are n pairs of observations of the variables X and Y, then

$$\sigma_x = [1/n\Sigma(x-\bar{x})^2]^{1/2}$$

$$\sigma_y = [1/2\Sigma(y-\bar{y})^2]^{1/2}$$

In the present application X and Y variables are the measured and predicted, by the neural network and by the AR model, total solar radiation values.

Figure 10 shows the temporal variation of the correlation coefficients between measured and predicted total solar radiation values using the neural network approach and the AR model for "multi-lag" predictions and for the same fourteen days of July 1995 (18th–31st) used in Fig. 7 for the "multi-lag" predictions with the neural network system. As it can be seen the neural network

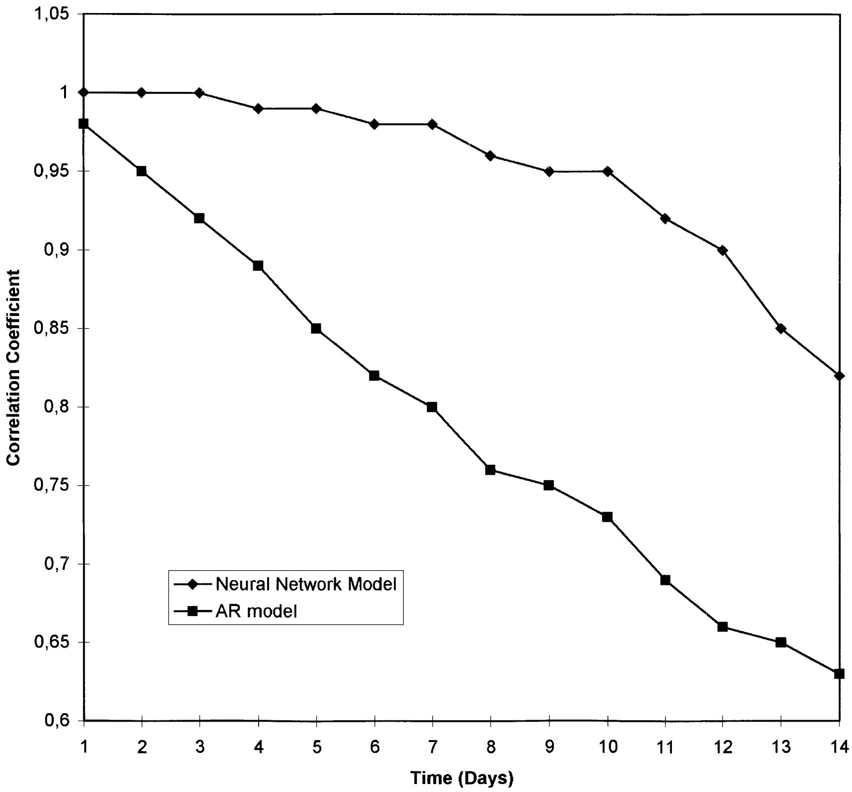

Fig. 10. Temporal variation of the correlation coefficient calculated between measured and predicted solar radiation values using an AR model and the neural network approach for fourteen days of July 1995 (18^{th}–31^{st})

makes significantly better predictions than those of AR model as prediction time increases. Similar results are taken from the comparison of the whole set of AR predictions with the corresponding neural network predicted values. It must be noted that AR model is essentially a linear model and then incapable to capture the nonlinear nature of the time series. The main advantage of the neural network model is that it enables the user to approximate or reconstruct any nonlinear activation function and therefore such a model is sufficiently flexible.

4. Conclusions

Remarkable success has been achieved in training the networks to learn the hourly total solar radiation values and to make accurate predictions of future vlaues. In more detail:

1. For "one-lag" predictions where the prediction of future values was based only on past measured values, it was found that the neural network approach is able to predict with remarkable success the total solar radiation values.
2. "Multi-lag" output predictions were performed using the predicted values to the input database in order to model future solar radiation. From the calculations, it was observed that it is possible to predict with sufficient accuracy, the total solar radiation values ten to twenty days in advance fr the warm period of the year. However, for the cold period of the year the predictions were not so promising as for the warm period because the results of a data-driven method such as the neural network approach depend strongly on the training sets of data and usually it is very difficult to make predictions during this period of the year which is characterised by the high frequency of different weather phenomena.
3. The results of the neural network model were compared with the corresponding results of a linear autoregressive model (AR) and it was found that the neural network approach leads to better predictions than the AR model.

References

Box GEP, Jenkins GM (1970) Time series analysis: forecasting and control. Holden-Day: Merrifield, VA

Carson JE (1963) Analysis of soil and air temperatures by Fourier techniques. J Geophys Res 68: 2217–2232

Chakraborty K, Mehrotra K, Mohan CK, Ranka S (1992) Forecasting the behavior of multivariate time series using neural networks. Neural Networks 5: 961–970

Cichocki A, Unbehauen R (1993) Neural networks for optimization and signal processing. John Wiley & Sons: Stuttgart

Connor JT, Martin DR, Atlas LE (1994) Recurrent neural networks and robust time series prediction. Institute of Electrical and Electronics Engineers (IEEE). Transaction on Neural Networks 5: 240–253

Dash PK, Ramakrishna G, Liew AC, Rahman S (1995) Fuzzy neural networks for time-series forecasting of electric load. IEE Proc-Gener Transm Distrib 142: 535–544

Eisenstein E, Kanter I, Kessler DA, Kinzel W (1995) Generation and Prediction of time series by a neural network. Phys Rev Lett 74: 6–9

El-Shal AO, Mayhoub AB (1995) Estimating solar radiation as a function of air temperature using Fourier series. Theor Appl Climatol 54: 153–159

Elsner JB (1992) Predicting time series using a neural network as a method of distiguishing chaos from noise. J Phys A: Math Gen 25: 843–850

Farmer JD, Sidorowich JJ (1987) Predicting chaotic time series. Phys Rev Lett 59: 845–849

Farmer JD, Sidorowich JJ (1988) Exploiting chaosi to predict the future and reduce noise. Technical Report LA-UR-88-901, Los Alamos National Laboratory

Hondou T, Sawada Y (1994) Analysis of learning processes of chaotic time series by neural networks. Prog Theor Phys 91: 397–402

Kalogirou S, Neocleous C, Michaelides S, Schizas C (1997) A time series reconstruction of precipitation records using artificial neural networks. Proceedings of EUFIT'97, Aachen, Germany, 2409–2413

Lalarukh K, Jafri YZ (1997) Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. Sol Energy 61: 23–32

Lamba BS, Khambete NN (1991) Analysis of soil temperature at various depths by Fourier technique. J Mausam 42: 269–274

Leite SM, Peixoto JP (1996) The autoregressive model of climatological time series: An application to the longuest time series in Portugal. Int J Climatol 16: 1165–1173

Li M, Mehrotra K, Mohan CK, Ranka S (1990) Sunspot numbers forecasting using neural networks. Proceedings of the IEEE Symposium on Intelligent Control 1: 524–529

Ljung L (1987) System identification: theory for the user. New York: Prentice-Hall

Mihalakakou G, Santamouris M, Asimakopoulos D (1998) Modeling the ambient air temperature time series using neural networks. J Geophys Res 103: 509–517

Pham DT, Liu X (1995) Neural networks for identification, prediction and control. Springer: Berlin

Rosenblatt F (1961) Principle of neurodynamics. Spartan Press: Washington

Rumelhart DE, Hinton GE, Williams RL (1986) Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL (ed) Parallel distributed processing. MIT Press: Cambridge, pp 318–362

Santamouris M, Mihalakakou G, Psiloglou B, Eftaxias G, Asimakopoulos D (1999) Modeling the global solar radiation on the earth surface using atmospheric deterministic and intelligent data driven techniques. J Climate 12: 3105–3116

Soderstrom T, Stoica P (1989) System identification. Prentice-Hall: New York

Teodorescu D (1990) Time series-information and prediction. Biol Cybern 63: 477–485

Weigend AS, Huberman BA, Rumelhart DE (1990) Predicting the future: a connectionist approach. Int J Neural Systems 1: 193–209

Wong FS (1991) Time series forecasting using backpropagation neural networks. Neurocomputing 2: 147–159

Authors' address: G. Mihalakakou, M. Santamounis and D. N. Asimakopoulos, Laboratory of Meteorology, Division of Applied Physics, Department of Physics, University of Athens, GR-15784 Athens, Greece.