www.ietdl.org

Published in IET Renewable Power Generation
Received on 17th February 2014
Revised on 19th June 2014
Accepted on 30th July 2014
doi: 10.1049/iet-rpg.2014.0057

ISSN 1752-1416

Harmonic analysis based hourly solar radiation forecasting model

Mehmet Fidan[1], *Fatih Onur Hocaoğlu*[2,3], *Ömer Nezih Gerek*[1]

[1]*Department of Electrical and Electronics Engineering, Anadolu University, Eskişehir, Turkey*
[2]*Depatment of Electrical Engineering, Afyon Kocatepe University, Afyonkarahisar, Turkey*
[3]*Solar and wind research and application center (GURAM), Afyonkocatepe University, Afyon, Turkey*
E-mail: mfidan@anadolu.edu.tr

Abstract: In this study, a novel two-step hourly solar radiation modelling procedure is developed. In the first step, hourly solar radiation data of a single day is considered and its periodic Fourier series coefficients are calculated. The first nine Fourier series coefficients are considered to accurately render the data. Among these nine coefficients, the yearly variations of the first two coefficients exhibit relatively large magnitudes and therefore are mathematically modelled. In the second step, various NN models are built to predict the solar behaviour using the Fourier model and the past data. The network inputs are the solar radiation data belonging to early sunrise hours of each day, and the model generated values of first two Fourier series coefficients. The NN outputs are the differences between actual and model values of first two Fourier series coefficients, together with the actual values of the following seven Fourier series coefficients. The NN models are trained and tested using different simulation parameters. Comparative results of the models are discussed in detail on hourly solar radiation data obtained from Izmir region of Turkey. It was observed that, once the model is constructed, it is possible to explain the behaviour of the solar radiation data in any day just by using the solar radiation data obtained early in the morning.

1 Introduction

Forecasting and accurate modelling of solar radiation in a particular region remains to be an important engineering problem. Efficiency of solar energy generators, solar heat systems, or even the architectural design applications depends on the robust solar radiation model. In a recent study the importance of monthly, daily and hourly solar radiation for the sizing of solar systems is analysed [1]. In another study the crucial role of the solar radiation model on a solar chimney power plant is explained [2]. The radiation data is naturally stochastic because of atmospheric effects. The eminent randomness makes it difficult to forecast the solar radiation in any hour of the day accurately. Consequently, many researchers deal with the solar radiation forecasting or modelling issue from a wide scale of mathematical and practical approaches. The approaches range from the correlation behaviour (spectral density) of the data, to more adaptive methods, including neural networks (NNs). Typical studies are summarised below. In [3] an Auto-Regressive (AR) model for the data is assumed, and tried to model accordingly. In a similar study, Mellit *et al.*, 2005 extended this model to auto-regressive moving-average (ARMA) [4]. Another study is carried out using Markov models [5]. Alternatively, NN-based models are also considered form modelling [6]. Kaplanis and Kaplani developed a stochastic simulation model for PV sizing [7]. A different stochastic model based on hidden Markov model is developed in [8]. A model is constructed using a novel visualisation method by Hocaoglu *et al.* [9]. In that study, the sequential solar radiation data are rendered in a 2-dimensional (2-D) matrix to utilise image-processing methods such as optimal coefficient linear prediction filters with Neural Networks. Moghaddam and Seifi apply linear predictive filters and NN for the prediction of solar irradiation for smart grid installations [10]. Almorox *et al.* employed air temperature data to predict daily solar radiations in Madrid [11]. There are some more examples for utilisation of NNs in both daily and hourly solar radiation data prediction models [12–14]. In another recent study principal component analysis is applied to model the variability of solar radiation data [15]. Yang *et al.* use Support Vector Machines for short-term solar radiation prediction [16]. An alternative approach for the solar radiation data analysis is the harmonic (Fourier) analysis. In such a study, harmonic analysis is performed over the daily solar radiation data belong to different regions of Turkey [17]. A similar study is also presented for Seeb region [18]. Harmonic analysis is performed for seven different regions in Oman in another study [19].

This particular work contains a novel attempt to utilise NNs over the harmonic (Fourier) data for the prediction of variational spectrum model. Despite the pseudo cyclic behaviour of solar radiation, all the previous prediction models attempt the problem as a time series, in plain time domain. Obviously, if a cyclic behaviour is expected (distorted by an additive noise element – which is the cloudlyness in our case), the robust parametrisation with as

few coefficients as possible becomes the frequency model (A.K.A. the Fourier model). It is also a well-known fact that NNs perform faster and more accurately for input/output parameters that are fewer in number. The time domain parameter reduction corresponds to short time data usage [9, 10], and naturally this causes omission of the cyclic behaviour. Despite the fact that the oscillations are best depicted by the Fourier model, the approach of such parameter reduction for solar radiation modelling was not used before, except the classical model by Kaplani and Kaplanis [7], which uses one sinusoidal component. We know that if the distortion of the cyclic signal causes a deviation from sinusoids, a few more Fourier series coefficients will more accurately define the signal. The experimental results are in accordance with the expected performance; the achieved mean squared error (MSE) between the model output and actual data was reduced to 3.49% of the true energy of the data, which could not be achieved using direct time domain prediction methods. Consequently, it is argued that modelling in Fourier domain together with a 'startup' information (first few hours of the daylight) provides a new insight and a powerful tool for predicting the 'rest' of the lighting throughout the day.

The model illustration is carried out over the data acquired in Izmir region of Turkey between years 2004 and 2005. The harmonic analysis enables us to construct the daily solar radiation data model in a unified and compact form. The idea is inspired from the fact that, the Discrete Fourier Transform of hourly solar variation in a day can be easily modelled by few simple harmonic components superimposed by a hard-to-predict random noise. Therefore the approach deals with the general behaviour of the data using Fourier analysis as explained in Sections 2 and 3. On the other hand it also considers the random behaviour of the data using the non-linear forecasting ability of NNs as explained in Section 5. The incorporation of the two concepts can be summarised as follows: The analysis starts by calculating 9-tap Fourier series (FS) coefficients of hourly solar radiation data for each day in a year. DFT length of 9 is chosen because of the 9-hour lit portion of the day. Experimentally, the training data is selected as the data obtained in 2004. The first two coefficients (among a total of nine) are modelled using the Kaplanis' approach as mentioned in Section 2 [20]. Model generated values are used as the first two inputs of the proposed feed-forward NNs. Apart from these 'fixed' model-generated values, the network contains other inputs such as the solar radiation values obtained in early periods of each day. The NN is trained to predict the 9-tap spectral pattern corresponding to the daily solar radiation. Specifically, the NN produces 'error values' between actual and model generated coefficient value together with the remaining seven FS coefficient values are determined as outputs of the NN. By training the NN for hourly data, the network is expected to converge to a state, where the first two lit hour solar radiation recordings would suffice for the prediction of the rest (using an indirect prediction on the Fourier domain, followed by an inverse Fourier transform). The NN models are tested under different simulation parameters. Reasonably accurate results are obtained and presented in Section 6.

2 Fourier series model for solar radiation

The literature survey about hourly solar radiation of a specific region yields the existence of various mathematical models. The classical model for hourly solar radiation was presented by Kaplanis [20], which is also shown in (1)

$$I(h, n_j) = a(n_j) + b(n_j) \cdot \cos\left(\frac{2\pi h}{24}\right) \tag{1}$$

The classical model has one DC and one cosine component for expressing the hourly behaviour of solar radiation in a day. The coefficients $a(n_j)$ and $b(n_j)$ vary according to the day, n_i. This model is altered in another work as in (2) [21]

$$I(h, n_j) = A(n_j) + B(n_j) \cdot \frac{\mathrm{e}^{-\mu(n_j)x(h)} \cos\left((2\pi h)/24\right)}{\mathrm{e}^{-\mu(n_j)x(h=12)}} \tag{2}$$

This model has an additional exponential term that depends on $\mu(n_j)$, which is the solar beam attenuation coefficient that is modelled using extra-terrestrial radiation and daily global solar radiation.

In this work, the proposed model is inspired by the above classical models. The model in (1) considers only one harmonic term and (2) contains non-linearities. In that aspect, our proposal is an extension of (1) to multiple harmonics, avoiding non-linearities. Naturally, the incorporation of harmonics indicates the utilisation of FS coefficients. It must be noted that some previous researches about solar radiation modelling also include FS theory [17–19]. However, in those works, the Fourier series expansion was used for modelling daily solar radiation (instead of hourly solar radiation data).

As indicated above, the classical model (1) can be thought as a Fourier series with one cosine harmonic. Here, the classical model is expanded using other cosine and sine harmonics (3)

$$I(h, n_j) = a_{n_j,0} + \sum_{i=1}^{N} \left(a_{n_j,i} \cdot \cos\left(\frac{2\pi i h}{24}\right) + b_{n_j,i} \cdot \sin\left(\frac{2\pi i h}{24}\right) \right) \tag{3}$$

The main goal of this expansion is to improve the accuracy of the model. If one-day data is modelled with (3), one-year data can be notated in a matrix, which is constructed from FS coefficients, as shown in (4)

$$\begin{bmatrix} a_{1,0} & a_{1,1} & \cdots & a_{1,N} & b_{1,1} & \cdots & b_{1,N} \\ \vdots & & & & & & \\ a_{n_j,0} & a_{n_j,1} & \cdots & a_{n_j,N} & b_{n_j,1} & \cdots & b_{n_j,N} \\ \vdots & & & & & & \\ a_{365,0} & a_{365,1} & \cdots & a_{365,N} & b_{365,1} & \cdots & b_{365,N} \end{bmatrix} \tag{4}$$

Given the hourly solar radiation for a day as $I(h, n_j)$, where h stands for the hour of day n_j, the FS coefficients, $a_{n_j,0}$, $a_{n_j,i}$ and $b_{n_j,i}$ can be calculated as shown in (5), (6)

and (7), respectively

$$a_{n_j,0} = \frac{\sum_{h=-11}^{12} I(h, n_j)}{24} \tag{5}$$

$$a_{n_j,i} = \frac{\sum_{h=-11}^{12} \left[I(h, n_j) \cos\left((2\pi ih)/24\right) \right]}{\sum_{h=-11}^{12} \left(\cos\left((2\pi ih)/24\right)\right)^2} \tag{6}$$

$$b_{n_j,i} = \frac{\sum_{h=-11}^{12} \left[I(h, n_j) \sin\left((2\pi ih)/24\right) \right]}{\sum_{h=-11}^{12} \left(\sin\left((2\pi ih)/24\right)\right)^2} \tag{7}$$

These parameters will be used in the NN model as input/output data. The recovery of the hourly information is obtained by the inverse Fourier transform.

3 Daily model for Fourier coefficients a_0 and a_1

Starting with the Fourier model of the hourly solar radiation (3), the coefficient, $a_{n_j,0}$, corresponds to the mean (average) of the values for day n_j. Experimentally, it was observed that the relation between $a_{n_j,1}$ and $a_{n_j,0}$ fits to a simple linear model in the form of $a_{n_j,1} = m \quad a_{n_j,0} + n$. Consequently, a coarse inter-coefficient model is obtained as

$$\begin{aligned} a_{n_j,0} &= I(n_j) \\ a_{n_j,1} &= mI(n_j) + n \end{aligned} \tag{8}$$

In Kaplanis *et al.*, $H(n_j)$, which is the daily solar radiation at surface, was defined as

$$H(n_j) = c_1 + c_2 \cos\left(\frac{2\pi n_j}{364} + c_3\right) \tag{9}$$

Experimentally, the equivalence of $I(n_j)$ to the above expression is confirmed, thus we have

$$I(n_j) = c_1 + c_2 \cos\left(\frac{2\pi n_j}{364} + c_3\right) \tag{10}$$

as well. According to the dependence of $a_{n_j,1}$ and $a_{n_j,0}$, we have

$$\begin{aligned} a_{n_j,1} &= m\left(c_1 + c_2 \cos\left(\frac{2\pi n_j}{364} + c_3\right)\right) + n \\ &= d_1 + d_2 \cos\left(\frac{2\pi n_j}{364} + d_3\right) \text{ where } d_3 \cong c_3 \end{aligned} \tag{11}$$

The coefficients c_1, c_2 and c_3, can be calculated as in (12)–(14)

$$c_1 = \frac{\sum_{n_j=-181}^{182} I(n_j)}{364} \tag{12}$$

$$c_2 = \max\left(\frac{\sum_{n_j=-181}^{182} I(n_j) \cos\left(2\pi(n_j + k)\right)}{\sum_{n_j=-181}^{182} \left(\cos\left(2\pi(n_j + k)\right)\right)^2}\right) \quad k = -181, -180, \ldots, 181, 182 \tag{13}$$

$$c_3 = \frac{2\pi}{364} \arg\max_M \left\{ \left(M = \frac{\sum_{n_j=-181}^{182} I(n_j) \cos\left(2\pi(n_j + k)\right)}{\sum_{n_j=-181}^{182} \left(\cos\left(2\pi(n_j + k)\right)\right)^2} \right) \quad k = -181, -180, \ldots 181, 182 \right\} \tag{14}$$

The coefficients d_1, d_2 and d_3 in (11) can also be calculated similarly, using $a_{n_j,1}$ values instead of $I(n_j)$.

4 Case study for the daily Fourier model

In this work, Izmir, Turkey is selected as the sample region. Using the methodology in Section 3 over the hourly solar radiation values in 2004, the $a_{n_j,0}$ and $a_{n_j,1}$ coefficients were

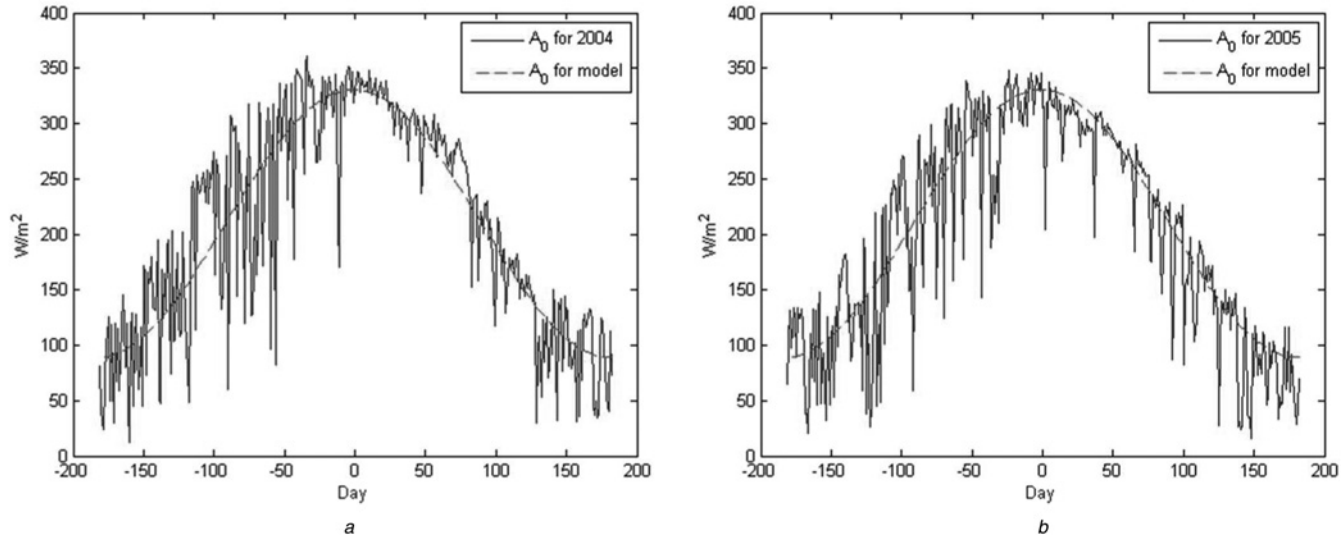

Fig. 1 *Daily calculated (real) $a_{n_j,0}$ values and the corresponding model values*

a Actual and model values of a_0 in year 2004
b Actual and model values of a_0 in year 2005

doi: 10.1049/iet-rpg.2014.0057

obtained as

$$a_{n_j,0} = 209.6827 + 120.6825\cos\left(\frac{2\pi n_j}{364} - 0.0173\right) \quad (15)$$

$$a_{n_j,1} = 317.4188 + 157.3008\cos\left(\frac{2\pi n_j}{364} + 0.0173\right) \quad (16)$$

The coefficients, c_3 and d_3 were naturally found the same, with a value difference within a small error limit of 0.0346 radians.

The daily calculated (real) $a_{n_j,0}$ values and the corresponding model values (derived from (15) and (16)) for the years 2004 and 2005 are shown in Fig. 1.

Fig. 1 indicates that there is a reasonable resemblance between the model and the actual values, which verifies the efficiency of the proposed model. The sample-wise error between actual and model values for $a_{n_j,0}$ in year 2004 and 2005 is shown in Fig. 2.

Here, the average error between the $a_{n_j,0}$ model and $a_{n_j,0}$ calculated from the 2004 radiation data is -7.8082 × 10^{-16}, the root mean square error (RMSE) is 43.61. Comparing this error to the root mean energy of $a_{n_j,0}$ (for the same year, 2004), which is 230.545; the model is visibly capable of accurately estimating the radiation pattern. Similarly, the average error between actual and model generated $a_{n_j,0}$ values for year 2005 is 8.5609 and the RMSE is 42.67, compared with the root mean energy of $a_{n_j,0}$, which is 222.627. Achievement of fair model accuracy for two separate years indicates that the proposed model for the $a_{n_j,0}$ coefficient is pretty robust.

The same comparison suit is carried out over the $a_{n_j,1}$ values for 2004 and 2005. The model-against-actual value comparisons are shown in Figs. 3 and 4.

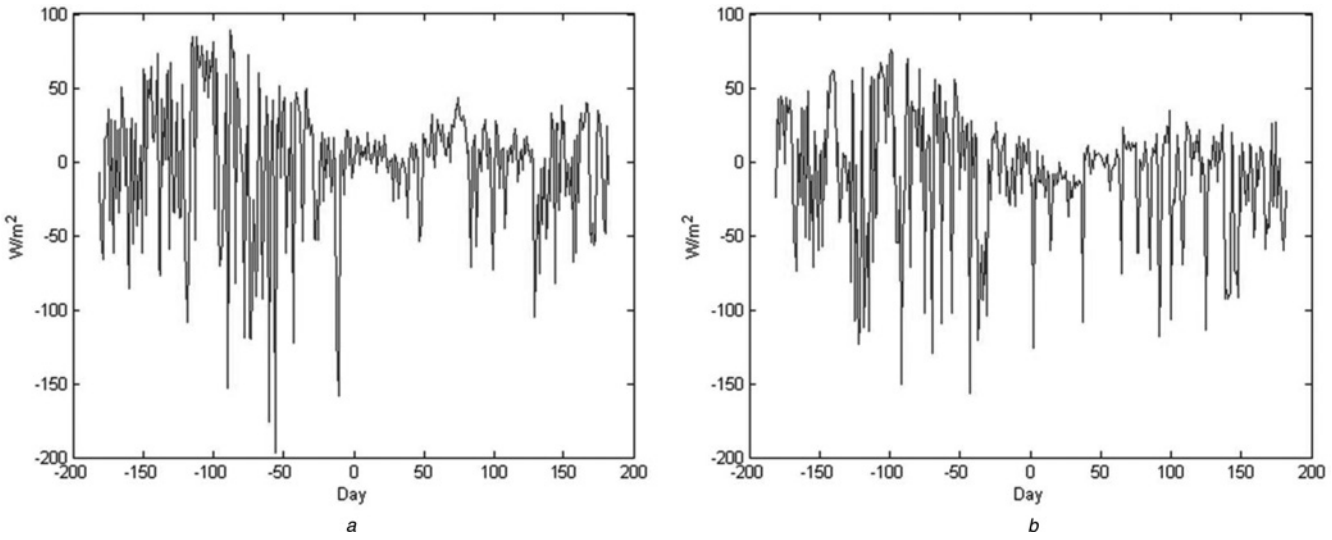

Fig. 2 *Sample-wise error between actual and model values for $a_{n_j,0}$ in year 2004 and 2005*

a Sample difference between actual and model values of a_0 in year 2004
b Sample difference between actual and model values of a_0 in year 2005

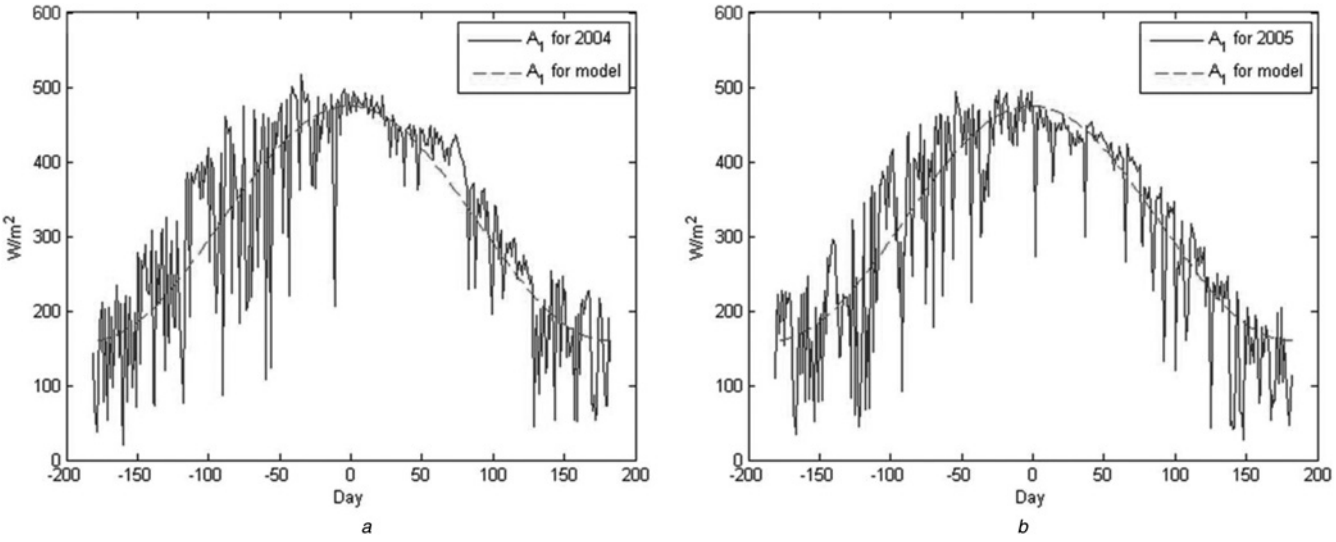

Fig. 3 *Model-against-actual value comparisons*

a Actual and model values of a_1 in year 2004
b Actual and model values of a_1 in year 2005

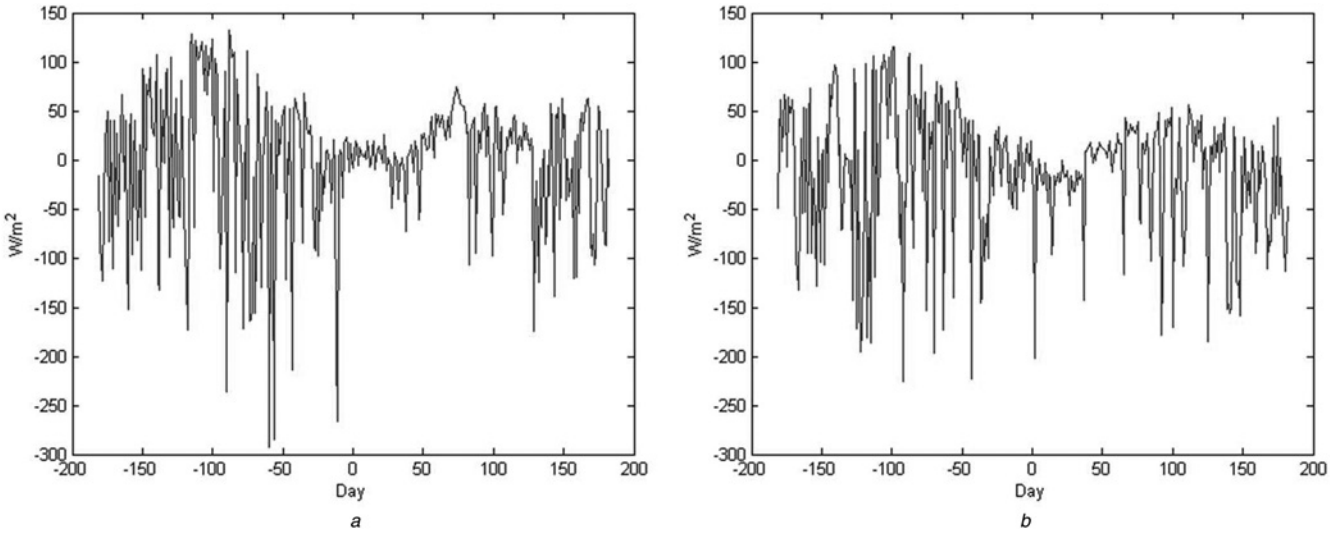

Fig. 4 *Model-against-sample value comparisons*

a Sample difference between actual and model values of a_1 in year 2004
b Sample difference between actual and model values of a_1 in year 2005

The quantitative analysis of the model generated and actual values of the $a_{n_j,1}$ coefficient yield that, similar to the previous coefficient, the average value between the value and the model output is negligible, whereas the RMSE value causes an energy decrease by a factor over 5 (69.21 against 343.380 for year 2004 and 66.88 against 334.036 for year 2005).

5 Neural network model construction

Owing to their learning ability, Neural Networks (NNs) are among the most popular techniques for scientific modelling and analysis of complex processes. The basic structure of a feed-forward NN includes three layers, namely the input layer, the hidden layer and the output layer. The number of neurons in each layer is application dependent. Each neuron in a layer is tied to another neuron in its neighboring layers with weighted connections. This basic structure was inspired by the neuron connectivity of brain cells, hence its name. Each neuron in a layer has an activation function (transfer function) that maps its input parameter values to its output value. In a training stage, the weights are adaptively updated to converge to a set of coefficients that scale the corresponding input value (the value coming from a previous layer's neuron output). The eventual goal is to minimise the error between actual outputs and the model generated outputs. The process is also called 'learning'. Once the final values of the weights with acceptable error are determined, the learning process is ended. At that stage, the NN is ready to use for testing purposes. The test must be performed with data that were not used in the training phase. Using NNs, owing to the non-linear transfer function, it is possible to depict non-linear behaviours. Non-linearity is eminent in solar radiation patterns because of geological motion and random atmospheric effects (such as clouding affect). Therefore non-linear transfer functions are adopted, namely the hyperbolic tangent sigmoid transfer function and the logarithmic sigmoid transfer function. The general form of the hyperbolic tangent sigmoid transfer function and logarithmic sigmoid transfer functions are given in (17) and (18)

$$o = \frac{2}{1 + \mathrm{e}^{-2i}} - 1 \tag{17}$$

$$o = \frac{1}{1 + \mathrm{e}^{-i}} i \tag{18}$$

where o and i are the output and the input of the transfer functions, respectively. Here, the input (i) is obtained as a linear weighted combination of outputs coming from the previous layer's neurons.

There exist several learning algorithms for feed-forward NNs. The performances of the algorithms are discussed in [7] in detail. The arguably most effective learning algorithm in terms of convergence speed and accuracy is the Levenberg–Marquard (LM) learning algorithm, hence it is adopted and used for this study. To assure the efficiency of this algorithm, internal experiments included the 'gradient descent with momentum and adaptive learning rate back propagation' algorithm inside the NN structure. Clearly, there are several more learning algorithms in the literature. However, the main idea presented herein is the 'learning' concept, therefore an exhaustive suit of NN experimentation is left beyond the scope of this work. A brief description of the alternative learning algorithms is given below. Then, in Section 4.3, the proposed NNs are applied to the Fourier coefficients that were explained in Section 3.

5.1 Basic back propagation algorithm (Gradient descent algorithm)

The basic back propagation training algorithm works by changing the values of the weights with a fixed length vector in the direction of the negative gradient. Therefore it is also called the gradient descent algorithm. Error function should be calculated from (19)

$$E = \frac{1}{2} \sum_{i=1}^{n} (t_i - o_i)^2 \tag{19}$$

where t_i is the target value (desired output) and o_i is the actual

doi: 10.1049/iet-rpg.2014.0057

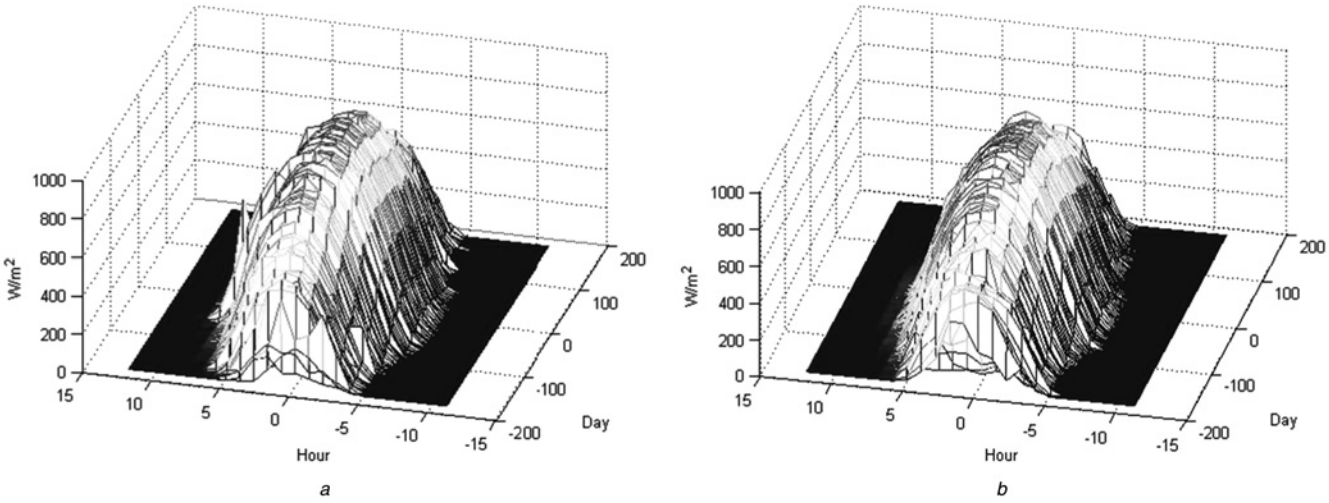

Fig. 5 *Parametric models of $a_{n_j,1}$ and $a_{n_j,0}$ coefficients are constructed from the hourly solar radiation data obtained at 10 m high ground level in Izmir region*

a Hourly solar radiation data obtained from Izmir region in 2004
b Hourly solar radiation data obtained from Izmir region in 2005

output and n is the number of input (or output) data dimension. It should be noted that (19) is valid for the networks that have one output vector. Consequently, the update function U_n is defined by

$$U_n = -g_n = -\eta \frac{\partial E}{\partial w_{ij,n}} \tag{20}$$

where w_{ij} are the weights of connections between neurons (or nodes) i and j and η is the, hence called, learning rate. A disadvantage of back propagation algorithm is its slow convergence. To accelerate the speed of convergence, utilisation of adaptive learning rates is usually necessary.

5.2 LM algorithm

A popular alternative to the basic gradient descent algorithm is the LM algorithm. It is particularly popular in application-based studies. In this algorithm, the update function, U_n, can be calculated as

$$U_n = -[\boldsymbol{J}^{\mathrm{T}} \times \boldsymbol{J} + \mu \boldsymbol{I}]^{-1} \times \boldsymbol{J}^{\mathrm{T}} \times \boldsymbol{e} \tag{21}$$

where $\boldsymbol{J}$ is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases and $\boldsymbol{e}$ is a vector of network errors. The parameter μ is a scalar number and $\boldsymbol{I}$ is the identity matrix. The update becomes identical to the basic back propagation (with a small step size) whenever the μ parameter is large. Typically, μ value should be decreased after each successful step and should be increased only when a tentative step increases the error term (or performance function). Therefore the performance function is guaranteed to reduce or get bounded at each iteration (Hagan and Menhaj, 1994). Owing to the speedy convergence to minima ability, LM algorithm is preferred in this study and the basic back propagation alternative is merely compared.

5.3 Neural network application to the Fourier series model of solar radiation

As explained in Section 3, parametric models of $a_{n_j,1}$ and $a_{n_j,0}$ coefficients are constructed from the hourly solar radiation data obtained at 10 m high ground level in Izmir region during year 2004. These data are compactly presented in Fig. 5*a*. The Fourier model is obtained from the solar radiation data of 2004. The NN adaptation (training) is achieved using the first months of 2005, and the testing is performed over the rest of the months in 2005. Solar radiation data of 2005 is analogously presented in Fig. 5*b*.

NN model is used for predicting the Fourier series coefficients corresponding to the current day early in the morning, hence that the whole day's solar radiation values can be forecasted from these predicted coefficients. Three separate input / output sets are formed and tested for the NN, as listed in Table 1. Given the hourly solar radiation for a day as $I(h, n_j)$, where h stands for the hour of day n_j, the FS coefficients, $a_{n_j,0}$, $a_{n_j,i}$ and $b_{n_j,i}$ can be calculated as shown in (5), (6) and (7), respectively.

Table 1 Input/output descriptions of the sets used for training the NN

		Set-1	Set-2	Set-3
network inputs	I1	$I(n_j,-6)$	$I(n_j,-6)$	$I(n_j,-7)$
	I2	$I(n_j,-4)$	$I(n_j,-5)$	$I(n_j,-6)$
	I3	$I(n_j,-2)$	$I(n_j,-4)$	$I(n_j,-5)$
	I4	$I(n_j, 0)$	$I(n_j,-3)$	$I(n_j,-4)$
	I5		$\hat{a}_{n_j,0}$	
	I6		$\hat{a}_{n_j,1}$	
network outputs	O1		$a_{n_j,0} - \hat{a}_{n_j,0}$	
	O2		$a_{n_j,1} - \hat{a}_{n_j,1}$	
	O3		$a_{n_j,2}$	
	O4		$a_{n_j,3}$	
	O5		$a_{n_j,4}$	
	O6		$b_{n_j,1}$	
	O7		$b_{n_j,2}$	
	O8		$b_{n_j,3}$	
	O9		$b_{n_j,4}$	

In order to avoid confusion of model generated coefficient estimates and actual coefficients, the model generated values are indicated by a hat ($\hat{a}_{n_j,0}$ and $\hat{a}_{n_j,1}$).

Training inputs and outputs were constructed from solar radiation data of selected 288 days from year 2005 (which is approximately 80% of the year). Test inputs and outputs

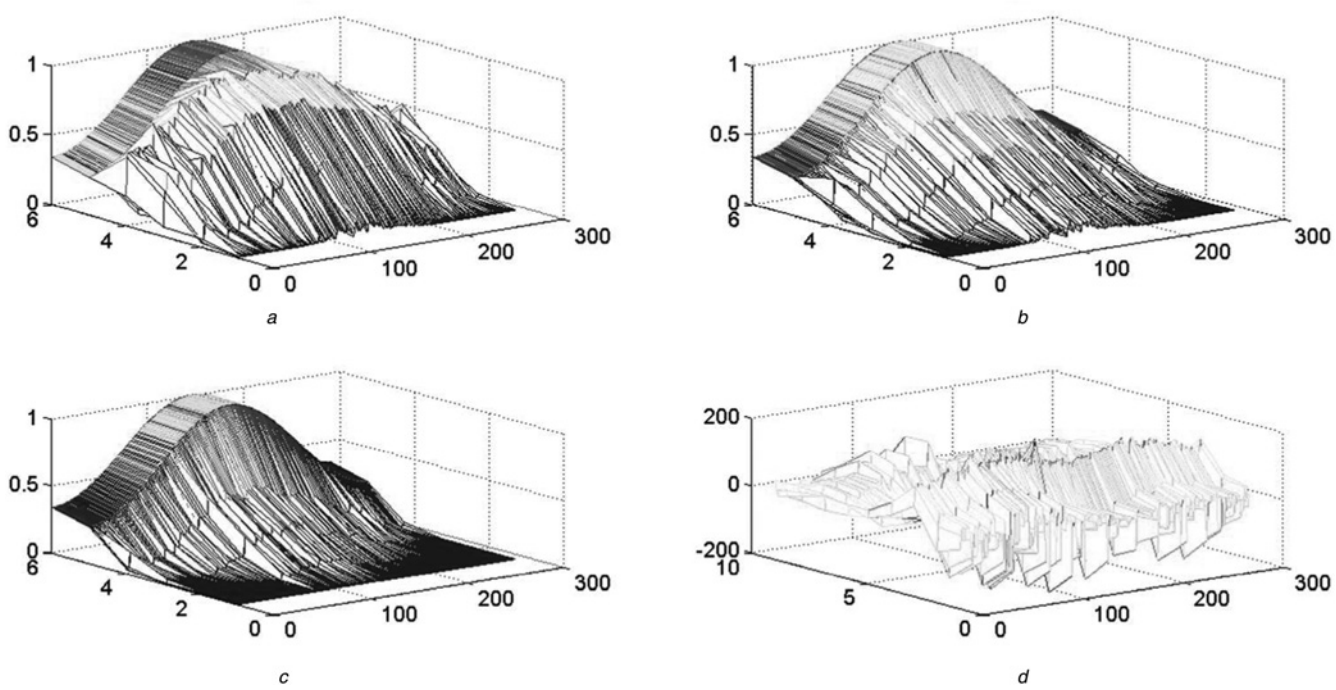

Fig. 6 *Training inputs of sets and desired outputs*

a Training inputs of Set-1
b Training inputs of Set-2
c Training inputs of Set-3
d Training desired outputs of all sets

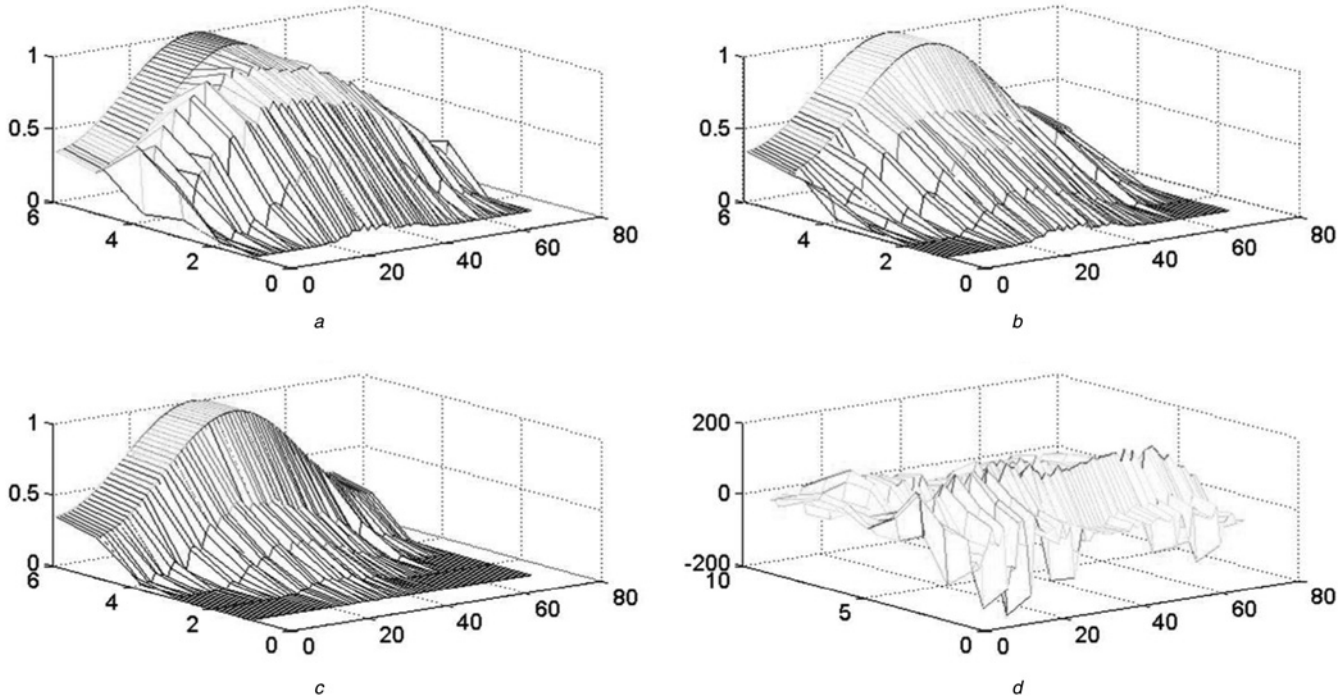

Fig. 7 *Training inputs of sets and desired outputs*

a Test inputs of Set-1
b Test inputs of Set-2
c Test inputs of Set-3
d Test desired outputs of all sets

doi: 10.1049/iet-rpg.2014.0057

were selected as the solar radiation data of the rest (72 days) from year 2005 (which makes approximately 20% of the year). The first four inputs were normalised to the interval [0, 1] (by dividing all input values by the corresponding maximum value in the year). The resultant training input sets and desired training output set are shown in Fig. 6. The corresponding test input sets and desired test output sets are shown in Fig. 7.

The performance results of the NN models are presented in Section 6.

6 Results and discussion

To test the performances of the NN models according to the model parameters, various learning algorithms with different transfer functions are applied while learning. It is observed that, the models learned with LM and Accelerated back propagation with momentum (AB) algorithms and the models with Sigmoid and Logsigmoig (Logsig) transfer functions give reasonable RMSE values (that are better than the other conducted tests). Therefore results corresponding to these methods and parameters are discussed here. Once, input/output pairs for each model are constructed, the data are separated into two parts being test data (including randomly selected 20% days of the whole year) and train data (remaining days). All models are trained for 100 epochs and tested with the (exclusive) test data. Using the Fourier coefficients (outputs of NN models with error terms) obtained from each model, solar radiation values are reconstructed through inverse Fourier transform of the corresponding day vector. Consequently, prediction values of solar radiations are obtained. The RMSE values between measured and predicted data are calculated and Tabulated in Table 2. The ratio between MSE of predictions and MS energy of original solar energy data are shown in Table 3 as in percentage and compared with classical model which is mentioned in the research of Kaplanis *et al.* with (2) [21]. It is clear from this table that the model incorporating LM algorithm together with Tan-sigmoid activation function is the best among the suit. The significant reduction of the output energy proves that initial solar radiation values carries significant knowledge about the behaviour of remaining solar radiation data, and it is possible to accurately predict the remaining data using the proposed approach. To illustrate the accuracy of the proposed method, the mesh plots of the reconstructed train data (predicted data) and the mesh plot of residual error are given in Figs. 8 and 9, respectively.

As can be seen from Fig. 9, the residual errors are randomly distributed around zero. This observation, together with a reduction in the error energy indicates that accurate forecasting results are achieved. In order to better illustrate the accuracy of the model on test data, both the measured and the predicted data are shown in Fig. 10*a* using an

Table 2 Performances of the models with respect to RMSE values

Data	RMSE value (reconstructed training data) Watt/m^2	RMSE value (reconstructed test data) Watt/m^2
set-1 (LM, Logsig.)	73.7612	79.8309
set-1 (LM, Tansig.)	43.8285	64.3058
set-1 (AB, Logsig)	117.8501	109.7586
set-1 (AB, Tansig)	110.2149	107.7455
set-2 (LM, Logsig.)	95.9457	109.4935
set-2 (LM, Tansig.)	99.6466	122.7264
set-2 (AB, Logsig)	115.6439	116.5477
set-2 (AB, Tansig)	111.6834	115.3723
set-3 (LM, Logsig.)	94.3508	110.8309
set-3 (LM, Tansig.)	102.3956	135.2779
set-3 (AB, Logsig)	125.3950	124.8640
set-3 (AB, Tansig)	151.2101	159.6120
classical model	140.9759	140.9980

Table 3 Comparison of MSE/MS energy

Data	(MSE/MS energy) × 100% (for training data)	(MSE/MS energy) × 100% (for test data)
set-1 (LM, Logsig.)	4.39	5.38
set-1 (LM, Tansig.)	1.54	3.49
set-1 (AB, Logsig)	11.04	10.04
set-1 (AB, Tansig)	9.80	9.36
set-2 (LM, Logsig.)	7.42	10.11
set-2 (LM, Tansig.)	8.01	12.73
set-2 (AB, Logsig)	10.79	11.48
set-2 (AB, Tansig)	10.07	11.25
set-3 (LM, Logsig.)	7.18	10.38
set-3 (LM, Tansig.)	8.46	15.46
set-3 (AB, Logsig)	12.69	13.17
set-3 (AB, Tansig)	18.45	21.53
classical model	16.04	16.80

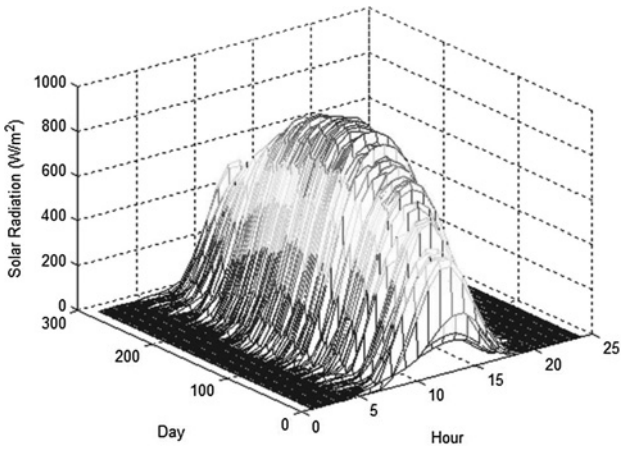

Fig. 8 *Mesh Plot of Predicted training data obtained from S1(LM, Tansig)*

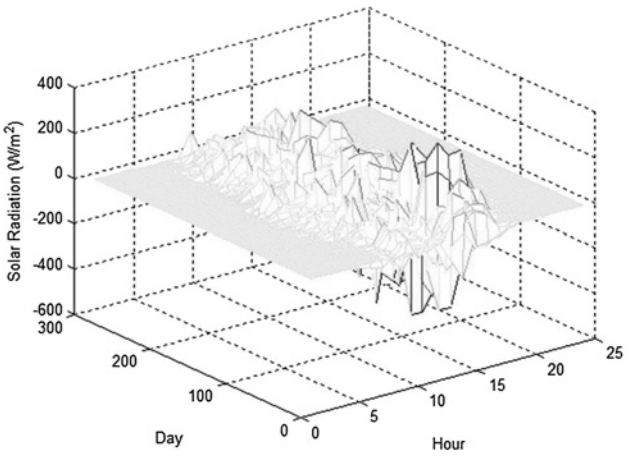

Fig. 9 *Mesh plot of error between predicted training data obtained from Set-1 (LM, Tansig) and measured training data*

overlay-plot. For visualisation purposes, a 'zoom section' from the overlay-plot of Fig. 10*a* is illustrated in Fig. 10*b*.

7 Conclusions

In this study, a novel solar radiation prediction model is developed. Hourly solar radiation data are considered and the first nine Fourier series (FS) coefficients of each day are calculated. Since the most of the energy is contained in the first two terms, these two coefficients are selected to construct an analytical model. Following the stage of this Fourier analysis, using the error between the actual values of the coefficients and their model values, various NN structures are proposed to depict the error. The inputs of NN models were the solar radiation values obtained from morning hours and model generated values of the first two FS coefficients. The first two outputs of the NNs were the error values between the initial two FS coefficients and their model generated values. Remaining seven outputs were chosen as rest of the FS coefficients. The performances of various NN models are discussed and compared. The experimentation of the proposed method is conducted on hourly data obtained from Izmir region during years 2004 and 2005. Using the data measured in 2004, the general FS behaviour of the data is modelled. As a first experimental output, the accuracy of the FS modelling approach is tested on the data measured in 2005. Following the Fourier model, the residues are also modelled using NNs to further reduce the prediction error. Obviously, several other harmonic and non-harmonic parameterisation techniques could be thought of. However, the models used in the previous literature also depend on sinusoidal harmonics (in fact, only one harmonic), therefore it remains reasonable (and novel) to incorporate multiple harmonics,

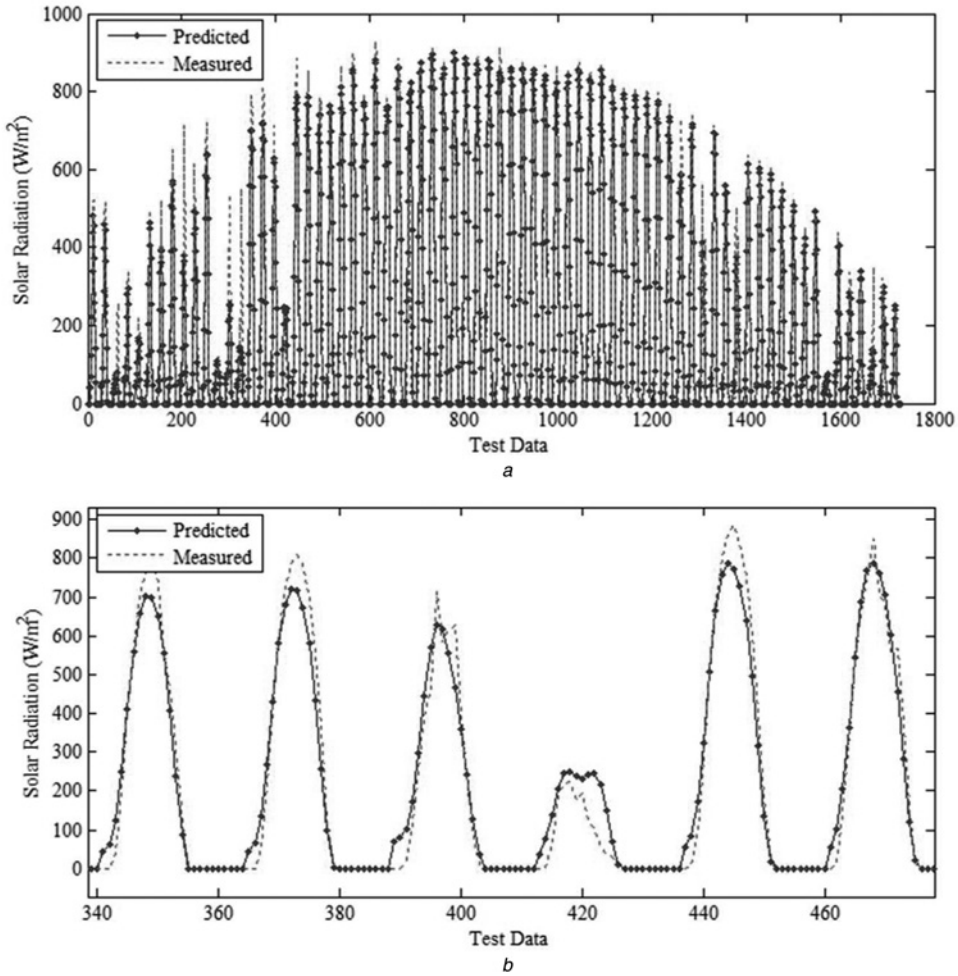

Fig. 10 *Visualisation purposes, a 'zoom section' from the overlay-plot*

a Measured and predicted test data
b Zoomed version of Fig. 10*a* from an arbitrary section

doi: 10.1049/iet-rpg.2014.0057

hence the Fourier theory. Furthermore, the modelling of the harmonics is simpler than the modelling of the time series, itself. Clearly, the proposed model can be extended and applied in several other regions throughout the world, because of the fact that the rotating earth produces cyclic solar illumination patterns that are distorted by the cloud (except the polar regions, where solar energy is rarely an alternative).

Another novelty of this work is the incorporation 'model' and 'initial readings' in an adaptive system (the NN) to predict the rest of the daily data. The prediction output is still in the FS domain, but the time signal can be easily achieved through inverse Fourier transform. The proposed methodology not only provides fair prediction results, but also gives an insight about the stochastic behaviour of the underlying stochastic process. Parametric (instead of FS) variations and different system (NN and other) choices can be considered as possible future directions of this study.

8 Acknowledgment

The authors acknowledge the Turkish State Meteorological Service (DMI) for the supply of hourly solar radiation data.

9 References

1 Khare, A., Rangnekar, S.: 'Optimal sizing of a grid integrated solar photovoltaic system', *IET Renew. Power Gener.*, 2014, **8**, (1), pp. 67–75
2 Guo, P., Li, J., Wang, Y.: 'Numerical simulations of solar chimney power plant with radiation model', *Renew. Energy*, 2014, **62**, pp. 24–30
3 Maafi, A., Adane, A.: 'A two state markovian model of global irradiation suitable for photovoltaic conversion', *Solar Wind Technol.*, 1989, **6**, pp. 247–252
4 Mellit, A., Benghanem, M., Hadj Arab, A., Guessoum, A.: 'A simplified model for generating sequences of global solar radiation data for isolated sites: using artificial neural network and a library of markov transition matrices approach', *Sol. Energy*, 2005, **79**, pp. 469–482
5 Aguiar, J., Collares-Perrira, M., Conde, S.P.: 'Simple procedure for generating of daily radiation values using library of markov transition matrices', *Sol. Energy*, 1988, **49**, pp. 229–279
6 Cao, J.C., Cao, S.H.: 'Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis', *Energy*, 2006, **31**, pp. 3435–3445
7 Kaplani, E., Kaplanis, S.: 'A stochastic simulation model for reliable PV system sizing providing for solar radiation fluctuations', *Appl. Energy*, 2012, **97**, pp. 970–981
8 Hocaoğlu, F.O.: 'Stochastic approach for daily solar radiation modeling', *Sol. Energy*, 2011, **85**, (2), pp. 278–287
9 Hocaoğlu, F.O., Gerek, Ö.N., Kurban, M.: 'Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural networks', *Sol. Energy*, 2008, **82**, pp. 714–726
10 Moghaddam, A.A., Seifi, A.R.: 'Study of forecasting renewable energies in smart grids using linear predictive filters and neural networks', *IET Renew. Power Gener.*, 2011, **5**, (6), pp. 470–480
11 Almorox, J., Hontoria, C., Benito, M.: 'Models for obtaining daily global solar radiation with measured air temperature data in Madrid (Spain)', *Appl. Energy*, 2011, **88**, pp. 1703–1709
12 Alam, S., Kaushik, S.C., Garg, S.N.: 'Assessment of diffuse solar energy under general sky condition using artificial neural network', *Appl. Energy*, 2009, **86**, pp. 554–564
13 Hocaoğlu, F.O., Gerek, Ö.N., Kurban, M.: 'A novel 2-D model approach for the prediction of hourly solar radiation', *LNCS Springer-Verlag*, 2007, **4507**, pp. 741–749
14 Mellit, A., Benghanem, M., Kalogirou, S.A.: 'An adaptive wavelet-network model for forecasting daily total solar-radiation', *Appl. Energy*, 2006, **83**, pp. 705–722
15 Manuel, Z., Pau, M.: 'Modeling the variability of solar radiation data among weather stations by means of principal components analysis', *Appl. Energy*, 2011, **88**, pp. 2775–2784
16 Yang, X., Jiang, F., Liu, H.: 'Short-term solar radiation prediction based on SVM with similar data'. 2nd IET Renewable Power Generation Conf., Beijing, China, September 2013, pp. 1–4
17 Sirdaş, S.: 'Daily wind speed harmonic analysis for Marmara region in Turkey', *Energy Convers. Manage.*, 2005, **46**, pp. 1267–1277
18 Dorvlo, A.S.S.: 'Fourier analysis of meteorological data for Seeb', *Energy Convers. Manage.*, 2000, **41**, pp. 1283–1291
19 Dorvlo, A.S.S., Ampratwum, D.B.: 'Harmonic Analysis of global irradiation', *Renew. Energy*, 2000, **20**, pp. 435–443
20 Kaplanis, S.N.: 'New methodologies to estimate the hourly global solar radiation: Comparisons with existing models', *Renewa. Energy*, 2006, **31**, pp. 781–790
21 Kaplanis, S., Kaplani, E.: 'A model to predict expected mean and stochastic hourly global solar radiation I(h, n_j) values', *Renew. Energy*, 2007, **32**, pp. 1414–1425