

## Research Article

# Learning Processes to Predict the Hourly Global, Direct, and Diffuse Solar Irradiance from Daily Global Radiation with Artificial Neural Networks

Hanae Loutfi,<sup>1</sup> Ahmed Bernatchou,<sup>1</sup> Younès Raoui,<sup>2</sup> and Rachid Tadili<sup>1</sup>

<sup>1</sup>Laboratory of Solar Energy and Environment, Faculty of Sciences, University Mohammed V, B.P. 1014, Rabat, Morocco

<sup>2</sup>Laboratory of Applied Mathematics, Computer Science, Artificial Intelligence and Pattern Recognition, Faculty of Sciences, University Mohammed V, B.P. 1014, Rabat, Morocco

Correspondence should be addressed to Hanae Loutfi; h.loutfi@um5s.net.ma

Received 20 April 2017; Accepted 2 July 2017; Published 11 October 2017

Academic Editor: Angelo Albini

Copyright © 2017 Hanae Loutfi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents three different topologies of feed forward neural network (FFNN) models for generating global, direct, and diffuse hourly solar irradiance in the city of Fez (Morocco). Results from this analysis are crucial for the conception of any solar energy system. Especially, for the concentrating ones, as direct component is seldom measured. For the three models, the main input was the daily global irradiation with other radiometric and meteorological parameters. Three years of hourly data were available for this study. For each solar component's prediction, different combinations of inputs as well as different numbers of hidden neurons were considered. To evaluate these models, the regression coefficient ( $R^2$ ) and normalized root mean square error (nRMSE) were used. The test of these models over unseen data showed a good accuracy and proved their generalization capability (nRMSE = 13.1%, 9.5%, and 8.05% and  $R = 0.98$ , 0.98, and 0.99) for hourly global, hourly direct, and daily direct radiation, respectively. Different comparison analyses confirmed that (FFNN) models surpass other methods of estimation. As such, the proposed models showed a good ability to generate different solar components from daily global radiation which is registered in most radiometric stations.

## 1. Introduction

Morocco is a fossil energy deficient country (95% imports) but has an impressive environmental wealth. It is characterised by an intensive solar irradiation, as it lies in a sunny belt, which favours the utilization of solar energy.

Morocco has launched one of the biggest solar projects: the world's largest concentrated solar power plant (Figure 1), costing about nine billion dollars, with a combined capacity of approximately 2 GW, to be completed by 2020.

The installation of any solar power system requires high-quality solar radiation measurements in order to size and simulate the system's functioning. Lack of long series of data or poor quality data series can combine errors in plant design, sizing, and performance forecasting, the thing that impacts negatively on the investment. Unfortunately,

measures of solar radiation are usually inaccurate and rare over the world [1], especially in Morocco. Due to the measuring device's price, there is only a small number of solar stations. These stations, in addition to being insufficient, generally measure only global radiation. However, the knowledge of direct solar component is necessary for concentrating solar power plant (CSP) or concentrated photovoltaic (CPV) sizing. This component is difficult to measure because it requires the use of a pyrhelimeter equipped with solar tracking system, which is very expensive.

On the other hand, sizing correctly all types of solar system or simulating their performance requires at least, daily or, even better, hourly values of different solar radiation components. Hence, it seems that elaborating relationships between available daily global data and the direct and diffuse solar irradiation ones at different time steps can be beneficial.



FIGURE 1: First part of the Moroccan Solar Plan “Noor1” inaugurated on February 2017.

Previous studies have stated that artificial neural networks (ANNs) are particularly suitable to reach this goal. Indeed, artificial neural networks (ANNs) are powerful when applied to problems whose solutions require knowledge that is difficult to specify but for which there is a huge quantity of examples. The neural network approach does not need to know any information regarding the process that generates the data. Recently, ANN models have been used in solar radiation modelling for many locations with different climates. Pertaining researches have been done in countries such as Greece, Saudi Arabia, Turkey, Egypt, Cyprus, Spain, India, Oman, Algeria, the UK, and Malaysia [2–33]. But no work seems to exist in Morocco. Thus, the purpose of this paper is the generation of horizontal hourly global, direct solar irradiances and daily direct radiation using ANN’s models. Daily global solar radiation, which is the most available component, and different astronomical variables are used as input to the ANN models.

The paper is organized as follows. In the first paragraph, we present a bibliographical review that illustrates the ability of artificial neural networks (ANNs) to elaborate nonlinear relationships between input and output data. Such relations were developed between meteorological parameters and global solar irradiances for different time scales especially monthly, daily, or hourly mean values but rarely for direct solar component. In the second paragraph, we present the proposed feed forward neural networks (FFNNs) and the statistical parameters to evaluate their performances.

In the third part, we give a description of our data measurement station where different meteorological variables are measured. A database choice criteria will be defined. Then, for each solar component, the methodology that was developed to address different issues will be applied and results will be discussed. Finally, conclusions and perspectives for further studies are presented.

## 2. Bibliographical Review

Modelling or predicting solar irradiation methods can be classified in three categories:

- (i) Methods that use empirical relations between clearness index ( $K_t$ ) and sunshine fraction ( $S_f$ );  $K_t$

is defined as a ratio of horizontal global solar irradiation on extraterrestrial irradiation while  $S_f$  is the sunshine duration divided by the theoretical day length. Generally, most of these methods have not been very accurate as they used high time steps or averaged data [26, 34, 35]. Stochastic models have been also applied at different time scales [36–38].

- (ii) The second category concerns models considering irradiative transfer modes, solar radiation, and earth atmosphere exchanges such as Rayleigh diffusion and absorptions by ozone, aerosols, and water vapour [26, 39–41]. Besides the fact that these models are complex, they only estimate solar irradiation in clear sky conditions.
- (iii) The most recent category is based on artificial intelligent methods. Nowadays, different models based on the artificial neural networks (ANNs) were introduced in literature for modelling and prediction of solar radiation data from meteorological and geographical parameters. ANNs are powerful when applied to problems whose solutions require knowledge that is difficult to specify but for which there is an abundance of examples. Indeed, for an ANN approach, we need a long-term data in order to get a better model. The developed model can be used either for forecasting data series, estimating solar irradiation from exogenous meteorological data, or for extrapolating solar irradiation from data measured on other sites [14, 42–44]. In fact, FFNNs can be used for any kind of input to output mapping. A feed forward network with one hidden layer and enough neurons in the hidden layers can fit any finite input-output mapping problem [45].

This study belongs to this last category: While in literature many ANNs have been developed for the prediction of global radiation, estimation of the direct irradiation has been less investigated [46]. Thus, different architectures of the feed forward neural network (FFNN) models were proposed to predict from daily global solar radiation the direct and diffuse solar radiation at different time steps. All components are considered on a horizontal surface.

## 3. Conception of the Artificial Neural Network (ANN) Model

**3.1. Artificial Neural Networks.** Artificial neural networks are information processing systems that are nonalgorithmic and massively parallel. They are composed of layers of parallel units called neurons. These neurons are connected with links called synapses. Each neuron has its own weight computed from the learning from data. They receive inputs over their incoming connections, perform nonlinear operations generally, and output the final results. ANNs have been applied in various aspects of science and engineering [47, 48].

There are two major categories of ANN: feed forward and feedback (recurrent) networks. The main difference between

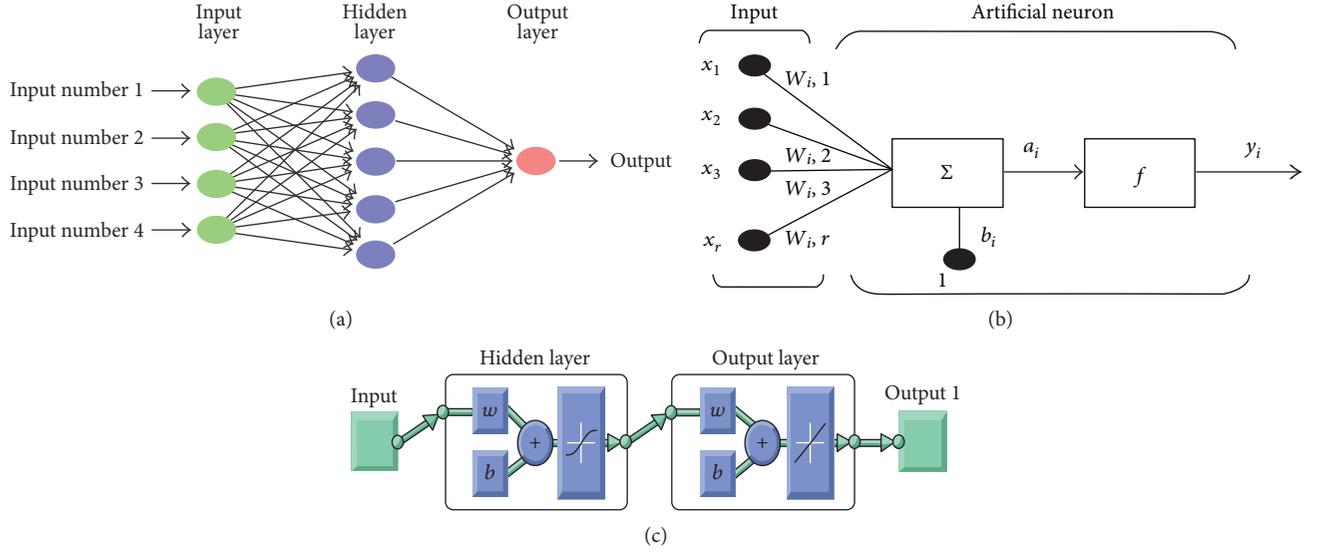


FIGURE 2: (a) Feed forward neural network. (b) Architecture of an artificial neuron. (c) Diagram of the used FFNN in MATLAB.

these categories is the existence of one or more loops in recurrent models, while feed forward networks are organized into layers connected strictly in one direction from the first layer to the last one [12].

**3.2. Feed Forward Neural Network (FFNN).** FFNNs are the most commonly used type of multilayer neural networks. A schematic diagram of the basic architecture is shown in Figure 2(a). Each neuron in the hidden layer (Figure 2(b)) sums up its inputs  $x_k$  after weighting them with the strengths of the respective connections  $w_{ik}$  from the input layer and calculates its output  $y_i$  as follows:

$$y_i = f \left( \sum_{k=1}^N w_{ik} x_k \right). \quad (1)$$

$f$  is a transfer function that can be a sigmoid, hyperbolic tangent, or radial basis function. The final output in the last layer is computed similarly.

The neural network adopted in our study is a multilayer feed forward backpropagation network. The backpropagation is the workhorse of learning with ANNs. It consists to adjust the weights of the neurons by minimizing the measure of the difference between the measured data and the predicted data obtained by (1). Inside the training, the neurons' weights are being updated by reinjection of the error inside the network.

Our network was designed and trained using MATLAB's code and MATLAB's neural network toolbox. A simplified schematic diagram of this network is shown in Figure 2(c); the main characteristics of this model are to be mentioned as follows:

- (i) There is one hidden layer (the user can change the number of hidden neurons).
- (ii) The transfer function adopted is sigmoidal while the output node has a linear activation function.

- (iii) The training algorithm is backpropagation based on a Levenberg-Marquardt (LM) minimization method which is the most commonly used [12, 48].

- (iv) The learning procedure is controlled by a cross-validation technique based on a random division of the initial set of data in 3 subsets (training, validation process control, and testing).

**3.3. Model Evaluation.** To evaluate the quality of estimation and ANN performances, several parameters can be used such as the following [49].

Root mean square error (RMSE) and normalized RMSE (nRMSE) are expressed as follows:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{i=1}^N (y_i - x_i)^2}{N}}, \\ \text{nRMSE} &= \frac{\left[ \sqrt{\frac{\sum_{i=1}^N (y_i - x_i)^2}{N}} \right]}{\bar{x}}. \end{aligned} \quad (2)$$

Determination coefficient  $R^2$  is defined by the equation:

$$R^2 = \frac{\left[ \sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x}) \right]^2}{\left\{ \left[ \sum_{i=1}^N (y_i - \bar{y})^2 \right] \left[ \sum_{i=1}^N (x_i - \bar{x})^2 \right] \right\}}. \quad (3)$$

Mean absolute error (MAE) and normalized mean absolute error (nMAE) are defined by the following equation:

$$\begin{aligned} \text{MAE} &= \frac{\sum_{i=1}^N |y_i - x_i|}{N}, \\ \text{nMAE} &= \frac{\sum_{i=1}^N |y_i - x_i|}{N\bar{x}}. \end{aligned} \quad (4)$$

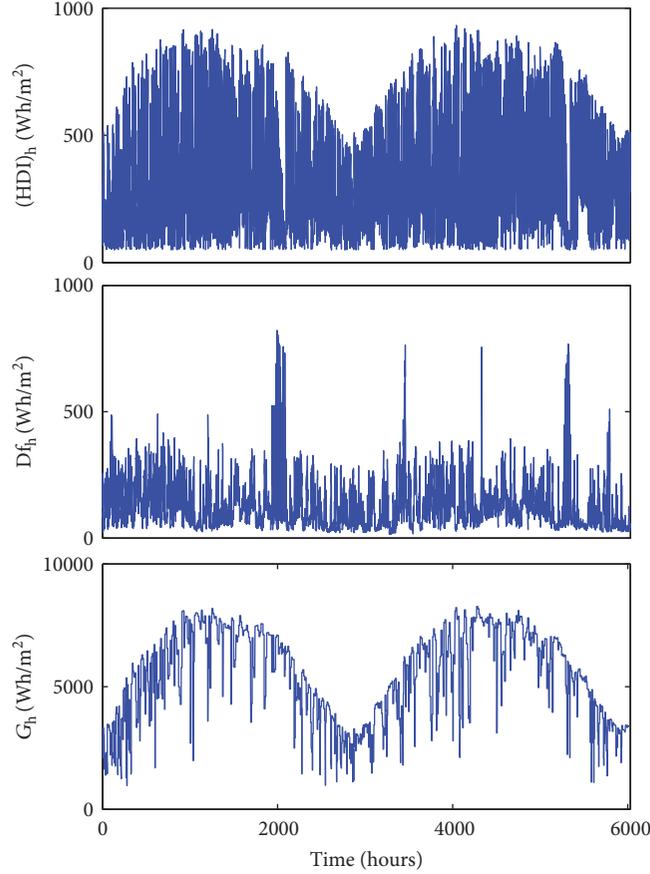


FIGURE 3: Hourly direct, diffuse, and global solar radiation for 2010 and 2011.

$y_i$  denotes estimated values and  $x_i$  the measured ones; their respective average values are defined as  $\bar{x} = \sum_{i=1}^N x_i/N$ ;  $\bar{y} = \sum_{i=1}^N y_i/N$  with  $N$ , the data number.

- (i) Root mean square error (RMSE) shows the difference between the measured values and the predicted ones; it indicates the scattering of data around linear lines. The approximation is better if RMSE is minimal (tends to 0).
- (ii) The determination coefficient ( $R^2$ ) expresses the correlation between the real values and the estimated ones; the best approximation corresponds to the highest  $R^2$  (closer to 1).

## 4. Meteorological Data and Database Development

**4.1. Meteorological Data.** The data used in this paper were measured in a radiometric station supervised by our laboratory and installed in the Faculty of Science and Technology in Fez (Morocco) (latitude:  $33^{\circ}56'N$ ; longitude:  $4^{\circ}99'W$ ) at an altitude of 579 m. The site's climate patterns can be summarized by dry and hot summers and cold winters.

Our station measures global irradiance ( $G_h$ ), direct normal irradiance, diffuse irradiance ( $Df_h$ ), wind speed

( $WS_h$ ), wind direction ( $Wd_h$ ), air temperature ( $T_h$ ) and relative humidity ( $Hr_h$ ). The station has also a pluviometer to measure precipitations and some radiometers to measure spectral components of solar radiation, in particular, the UV and the active photosynthetic radiation APR. All the measure instruments are related to a data acquisition board (CR10X) with a storage module. Data are measured and recorded every 5 seconds and then converted to hourly averages.

Hence, three years of hourly data recorded from the 1st January 2010 to 31st December 2012 are available with 26304 data records for this study. In order to extract outlier values, each parameter was examined. Then data before sunrise and below sunset periods were also deleted to avoid the mask effect or a nonreliable response of pyranometers at high zenith angle on solar data [50]. Therefore, we have 13,085 records for each variable mentioned above. All these variables are going to be used as input parameters to the ANN models. Figure 3 presents the variation of horizontal direct irradiation ( $(HDI)_h$ ), diffuse ( $(Df)_h$ ), and global ( $G_h$ ) solar radiation versus time over two years.

In addition to measured variables, we have calculated the sunshine duration (SD) defined by the World Meteorological Organization as the time interval when direct solar radiation exceeds  $120 \text{ (W m}^{-2}\text{)}$ . It is a radiometric parameter, giving information about the site's nebulosity [51]. Besides, we have considered some astronomical variables: declination

angle ( $\delta$ ) and daylight hours (DH), the first permits to designate a specific day; the second indicates the period of the year considered. We have also calculated sunset hour angle ( $W_{ss}$ ), sun hour angle (HA), extraterrestrial solar radiation on horizontal plan (ER), and the sun fraction ( $S_f$ ), defined as a ratio of solar duration to daylight hours. All these astronomic variables have been computed using analytic expressions detailed in the Appendix.

**4.2. Database Elaboration Criteria.** The ANN models' performance depends on the choice of the best combination of weather variables as input, training algorithm, and ANN architecture design. The most important key task in time series prediction is the selection of the input variables. It is, in fact, a prerequisite stage as there is no systematic approach to adopt for nonlinear ANN models [45]. However, we must take into account some criteria such as the following:

- (i) Parsimony which consists in developing the simplest ANN architecture with a minimum of inputs, hidden layers, and hidden neurons while keeping high performances.
- (ii) Avoid redundant inputs (they contain the same information) and choose the best-correlated variables to solar irradiation. Indeed, too many inputs can reduce the model efficiency [51].

To deal with this issue, we will either compute the correlation between different input variables and the target variable and then decide of the best combination or consider many combinations of variables during the training phase and choose the one that gives the best results.

## 5. Results and Discussion

**5.1. Prediction of Hourly Global Solar Irradiation from Daily Radiation.** As mentioned before, daily global solar radiation, when measures exist, is the most common available measure with relatively long series of data. But to carry out an accurate solar system sizing and evaluation, hourly data are needed. In this section, a FFNN model was trained to generate hourly horizontal global solar radiation. The daily global radiation ( $G_D$ ) is the main input in addition to some calculated variables especially declination, hour angle, sunset hour angle, and extraterrestrial solar radiation on horizontal plan.

To implement the network, different combinations of these parameters have been proposed as matrix input (stimuli) to FFNN models with horizontal (HDI) $_D$  as target. For each combination, different configurations of the neural network were tested by changing the number of hidden neurons. Training was repeated up to 10 times for each number of hidden neuron so that we choose the best configuration. Table 1 shows the best results for training through the correlation coefficient  $R$  and RMSE. The training was done over 2 years (8714 records).

According to these results, the best FFNN model corresponds to three nodes in the input layer (HA,  $W_{ss}$ , and  $G_D$ ) and 10 neurons in the hidden layer. To ensure the efficiency

TABLE 1: Best learning's performances with  $G_h$  as target.

Matrix Input	Hidden neurons	RMSE (kWh/m <sup>2</sup> )	$R$
[HA, $W_{ss}$ , $G_D$ ]	10	0.0580	0.9813
[ $\delta$ , HA, $W_{ss}$ , $G_D$ ]	12	0.0583	0.9809
[HA, $W_{ss}$ , ER $_h$ ]	10	0.0783	0.8978

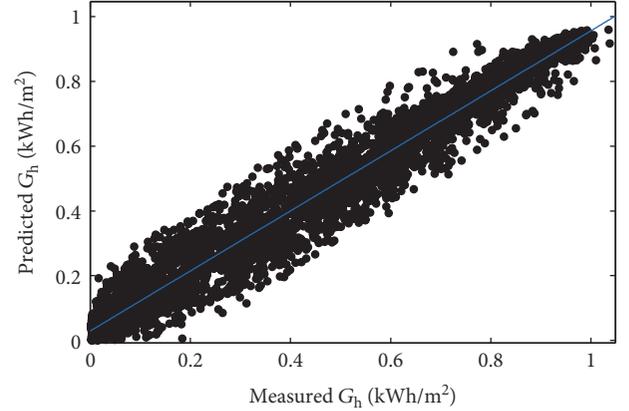


FIGURE 4: Regression plot of measured and predicted  $G_h$ .

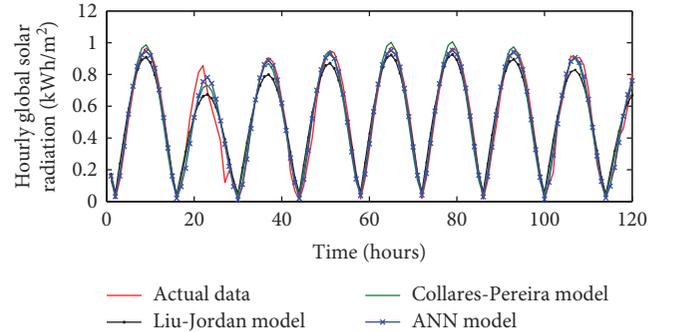


FIGURE 5: Comparison between ANN results and other procedures of estimation.

of the developed network to generate synthetic hourly data, it was tested over unseen data (1 year, 4371 records).

Figure 4 shows a regression plot which represents predicted values of hourly global solar radiation for the year 2012 using the proposed model versus the measured ones. It is clear that there is a good correlation:  $R = 0.98$  and  $RMSE = 0.061$  kWh/m<sup>2</sup>.

To discuss the validity of these results, the proposed model was compared with accurate empirical models. A comparison between our model and Liu-Jordan and Collares-Pereira and Rabel (C-P&R) models is carried out. A brief description of these models is given in the Appendix. In Figure 5, we show a sample (8 solar days) to illustrate the conducted comparison. It appears that the three models can predict hourly solar radiation data accurately in clear sky days. Nevertheless, sometimes the Liu-Jordan model tends to underestimate values. As it can be seen, the proposed model is more accurate than the other two models. However,

TABLE 2: Statistics for different methods of estimation of  $G_h$ .

Model	nMAE (%)	nRMSE (%)	$R^2$
Liu-Jordan	15.6	20.2	0.91
(C-P&R)	14.3	18.8	0.92
FFNN (testing)	12.0	13.1	0.96

for the three models the prediction's accuracy is lower on cloudy days than on sunny days but still acceptable as most of the solar radiation prediction model's accuracies degraded on cloudy days [52–54].

In Table 2, we reported a deepened evaluation of the three models using nMAE, nRMSE, and  $R^2$ . Indeed, one year of hourly data was generated using the three models. Results show that the proposed model exceeds the other models. This implies that the proposed model has more ability to predict future data based on the RMSE value. Moreover, it is more powerful in predicting hourly solar radiation according to the nMAE value.

Consequently, it is clear that the proposed model exceeds the other empirical models. In addition, it has the capacity of learning and handling huge data sets with nonlinear behaviour and stochastic nature. Even so, the empirical models can exceed our FFNN model in case of lack of long series data records necessary for learning processes.

**5.2. Prediction of Hourly Horizontal Direct Solar Irradiation from Daily Radiation.** While most applications adopt the global solar radiation, the concentrating solar plants (CSP) require, generally, an accurate estimation of direct solar irradiation at different time scales [55]. The direct irradiation forecasting represents an important aspect for a full solar energy potential evaluation. Global radiation measurements are usually obtained in most of the radiometric stations, whereas the data availability of its direct component is more limited. Moreover, when the direct component is measured, there is no extensive data series. Models to generate the direct solar resources are needed in order to establish its typical behaviour for energy applications.

In this paper, we considered horizontal direct solar irradiation (HDI), the normal component (NDI), if needed, can be calculated using the following equation:

$$(NDI) = \frac{(HDI)}{\cos(A_z)}, \quad (5)$$

where  $A_z$  denotes the zenith angle.

In particular, in this, paragraph we deal with hourly horizontal direct irradiation  $(HDI)_h$  as target. Thus, to implement a suitable FFNN model, we have first defined the input variables to the model. To this end, a set of heterogeneous parameters has been considered. Especially, daily global radiation ( $G_D$ ) combined to astronomical, climatic, and radiometric variables ( $\delta$ , SD, HA,  $G_D$ ,  $K_t$ ,  $W_{ss}$ , and  $S_f$ ).

$G_D$  is the main variable as it provides information about the climatic and meteorological patterns. The hour angle and declination angle are information utilized to train our FFNN with considering information about sunlight duration

TABLE 3: Best performance of learning processes for different FFNN configurations.

Model	Input matrix	Hidden neurons	$R$	RMSE (kWh/m <sup>2</sup> )
1	SD, HA, $G_D$ , $K_t$	15	0.980	0.0707
2	SD, $G_D$ , $K_t$	10	0.846	0.1224
3	$S_p$ , HA, $G_D$ , $K_t$	15	0.953	0.0707
4	$S_f$ , HA, $K_t$	12	0.892	0.1140
5	SD, HA, $K_t$	10	0.920	0.0836
6	SD, HA, $G_D$	10	0.891	0.1090
7	$\delta$ , SD, HA, $G_D$ , $K_t$	20	0.984	0.0390
8	$\delta$ , SD, HA, $G_D$ , $K_t$ , $W_{ss}$	10	0.950	0.0707

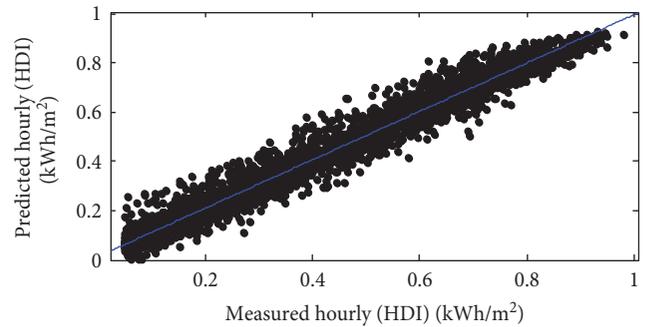


FIGURE 6: Predicted versus measured hourly HDI.

and the day considered. Mainly, the HA has an effect on the optical path length through the atmosphere; then, it can replace the relative air mass. The clearness index  $K_t$  represents the most relevant factor in the  $(HDI)_h$  prediction. In fact, it represents an indirect measure of the atmosphere filtering action.

Table 2 shows the best performance for different architectures of the FFNN model for training process. It is worth recalling that the learning process is done over data for 2010 and 2011 and training repeated up to 10 times for each number of hidden neurons with the Levenberg-Marquardt algorithm to do fitting of the hourly horizontal direct solar radiation.

From Table 3, we notice that models 1 and 7 (7 has the same variables as 1 in addition to  $\delta$ ) give the best results. However, model 7 surpasses slightly model 1 especially in terms of RMSE. Thus, model 7 with five variables was adopted to carry on the hourly (HDI) generation as  $\delta$  is a costless parameter. The generalisation of the model was evaluated by testing the model over data corresponding to 2012. Results are satisfactory as it can be seen in the regression plot of Figure 6 which represents predicted  $(HDI)_h$  versus the measured ones. Indeed,  $R = 0.982$  and  $RMSE = 0.041 \text{ kWh.m}^{-2}$  during test process. This high correlation value implies that the proposed model makes accurate predictions.

In addition to that, both predicted and measured hourly (HDI) values were plotted versus time for the whole year. We also plotted the error defined as the difference between predicted and measured values. For a good illustration,

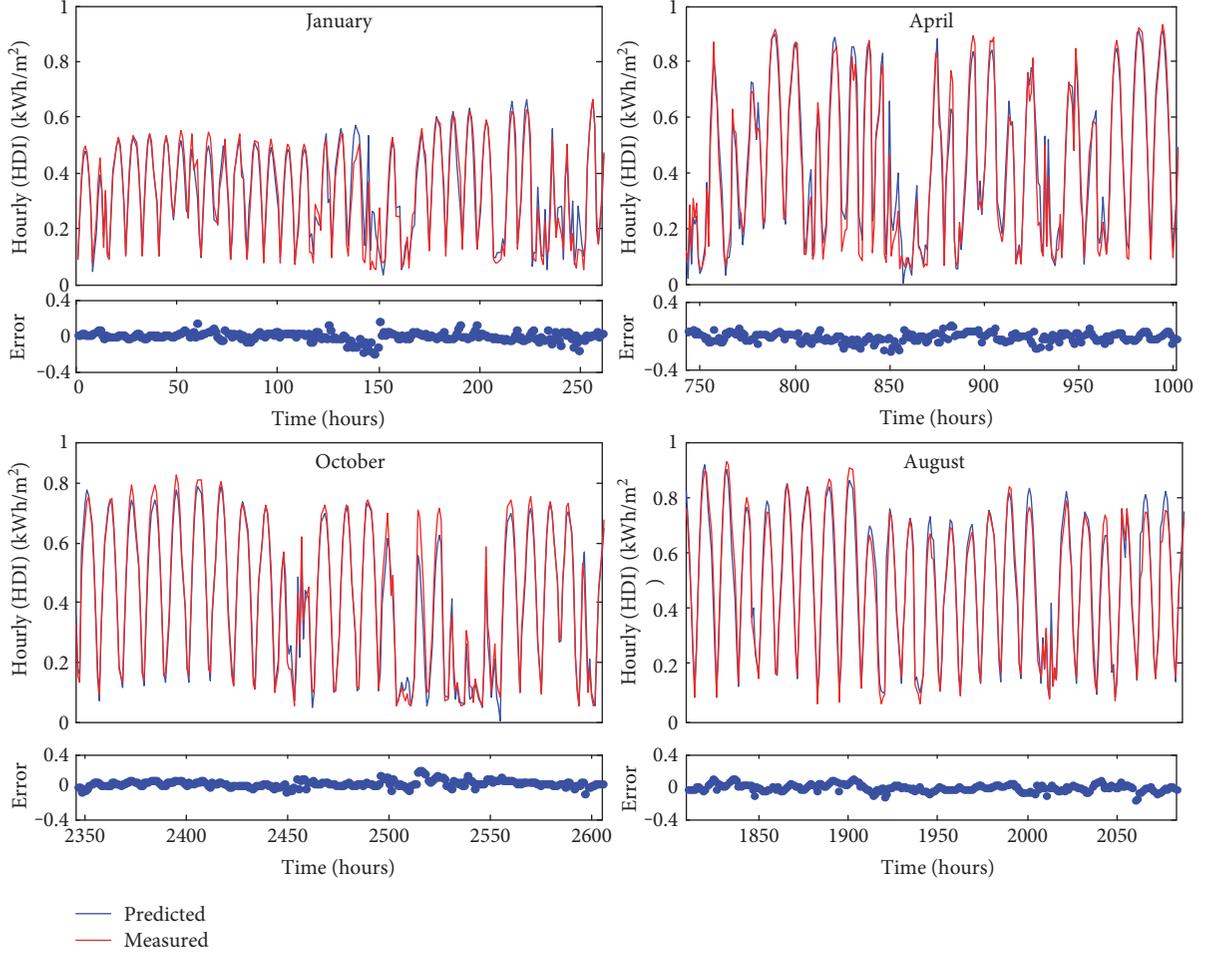


FIGURE 7: Measured and predicted hourly (HDI) for four months representing different seasons.

Figure 7 presents one month from each season. Generally, there is a good agreement. However, the quality of the fitting differs from a season to another and, as it can be seen, the error scatter is getting closer and closer to zero from winter to summer.

Concerning the diffuse component, it is worth mentioning that once the direct radiation is predicted, the diffuse component can be determined by subtracting direct radiation from global radiation on a horizontal surface.

**5.3. Prediction of Daily Horizontal Direct Solar Irradiation from Daily Radiation.** In some meteorological stations, there are relatively long series of only daily values of global solar radiation, temperature, and relative humidity. The aim in this section is to show how we can use these data to predict daily HDI  $(HDI)_D$  using FFNNs.

For this analysis, we computed daily values for all variables from the recorded hourly ones to implement FFNN models. We kept the same scripts for variables and just replaced subscript “h” by “D.”

1095 data records are available; 730 are used for the learning process, and the rest is used to test the model (365 records). As mentioned before, a correlation analysis was carried out between the input variables and  $(HDI)_D$ . Especially,

TABLE 4: Correlation of different parameters to  $(HDI)_D$ .

Variables	Correlation to $(HDI)_D$
Mean temperature ( $T_D$ )	0.520
Sunshine duration (SD)	0.927
Declination angle ( $\delta$ )	0.65
Relative humidity ( $Hr_D$ )	-0.669
Wind speed ( $WS_D$ )	0.687
Global solar radiation ( $G_D$ )	0.940
Daylight hours (DH)	0.558
Extraterrestrial radiation ( $ER_D$ )	0.556
Sunshine fraction ( $S_F$ )	0.865

a set of nine heterogeneous parameters have been considered: horizontal extraterrestrial solar radiation ( $ER_D$ ), global solar radiation ( $G_D$ ), mean temperature ( $T_D$ ), sunshine duration (SD), sunshine fraction ( $S_F$ ), wind speed ( $WS_D$ ), declination angle ( $\delta$ ), daylight hours (DH), and relative humidity ( $Hr_D$ ).

Results, reported in Table 4, show that the most effective input parameters are global solar radiation and sunshine duration.

TABLE 5: Best performance of FFNN architectures.

Input matrix	Hidden neurons	nRMSE (%)	R
SD, $G_D$	12	15.40	0.97
$S_F$ , $G_D$	15	15.46	0.96
SD, $G_D$ , $\delta$	10	7.73	0.992
SD, $G_D$ , $ER_D$	10	8.60	0.989
$S_F$ , $G_D$ , $ER_D$	10	12.01	0.98

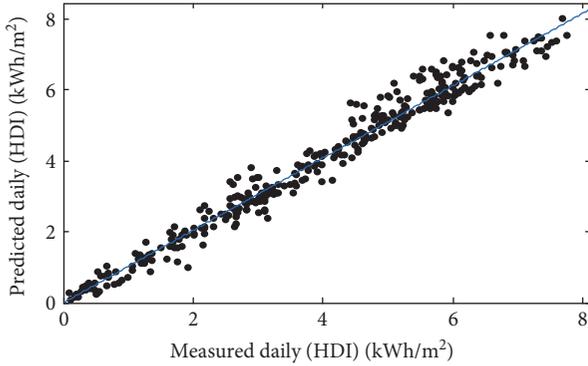


FIGURE 8: Regression plot of measured and predicted daily horizontal direct irradiation.

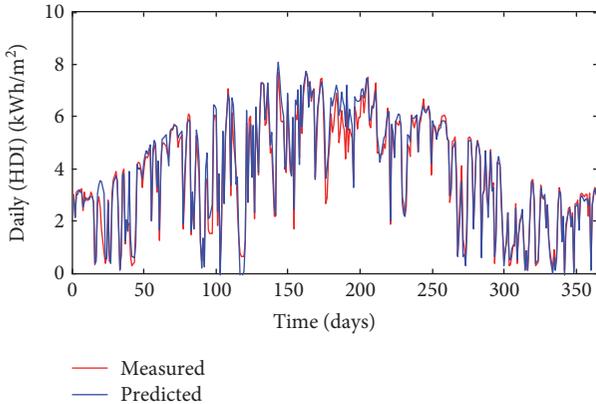


FIGURE 9: Measured and predicted daily horizontal direct irradiation for test process.

Once the most effective input variables were defined, different combinations of these parameters have been proposed as matrix input to FFNN models with horizontal  $(HDI)_D$  as target. Table 5 shows the best results of training for all the analyzed combinations in terms of nRMSE and  $R$ .

It can be noted that the network with three input variables (sunshine duration (SD), global solar radiation ( $G_D$ ), and declination angle ( $\delta$ )) reaches the best results for the  $(HDI)_D$  prediction.

This configuration has been used to test the capability of the model of generalisation. Indeed, the model was

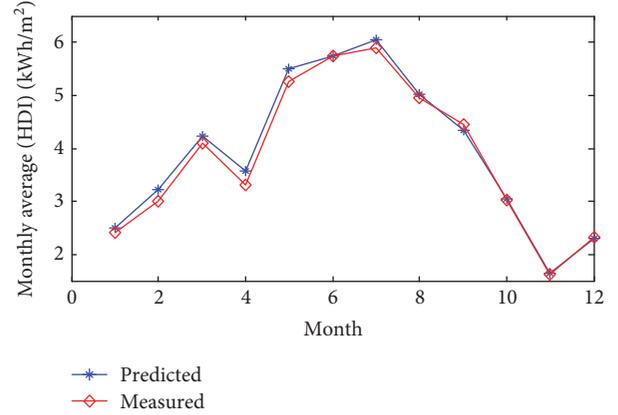


FIGURE 10: Comparison between the monthly averages of measured and predicted horizontal direct irradiation (HDI).

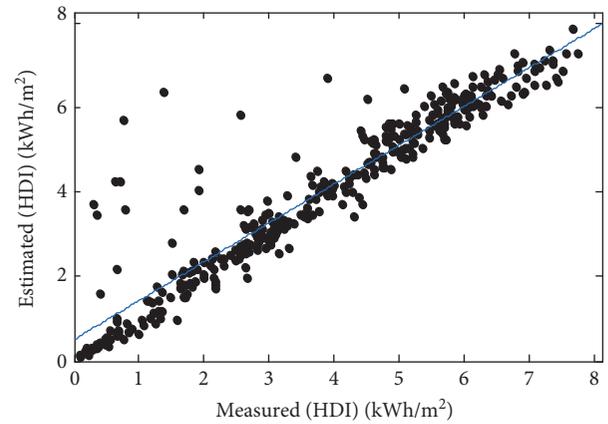


FIGURE 11: Regression plot of measured and estimated horizontal direct solar radiation.

tested over unseen data corresponding to 2012 (not used for learning process).

Figure 8 shows a regression that represents the predicted daily direct solar radiations versus the corresponding measured values in a test phase. It is clear from this plot that there is a good agreement between the predicted values and the measured ones. Such result is confirmed by  $R = 0.99$  and a  $nRMSE = 8.05\%$ .

In Figure 9, we plotted both measured and predicted daily direct solar radiation versus time corresponding to test period. Globally, there is no significant difference between the two curves.

For more illustration and to get an idea of the monthly direct solar irradiation in Fez, we represent, in Figure 10, the monthly average of predicted and measured direct solar radiation. There is a net improvement compared to Figure 9 as daily fluctuations are attenuated in monthly averages. Moreover, we notice that even for months with low solar radiation the ANN model gives relatively accurate prediction. Most of the monthly averages of  $(HDI)_D$  fluctuate in the range of 2.5–6  $kWh/m^2$ .

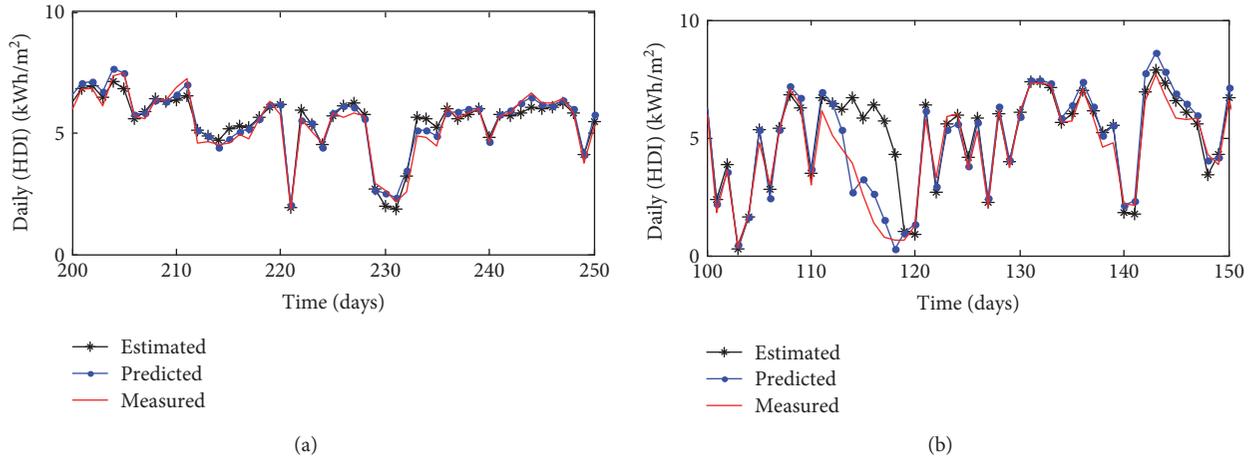


FIGURE 12: Comparison between estimated, predicted, and measured daily direct solar radiation.

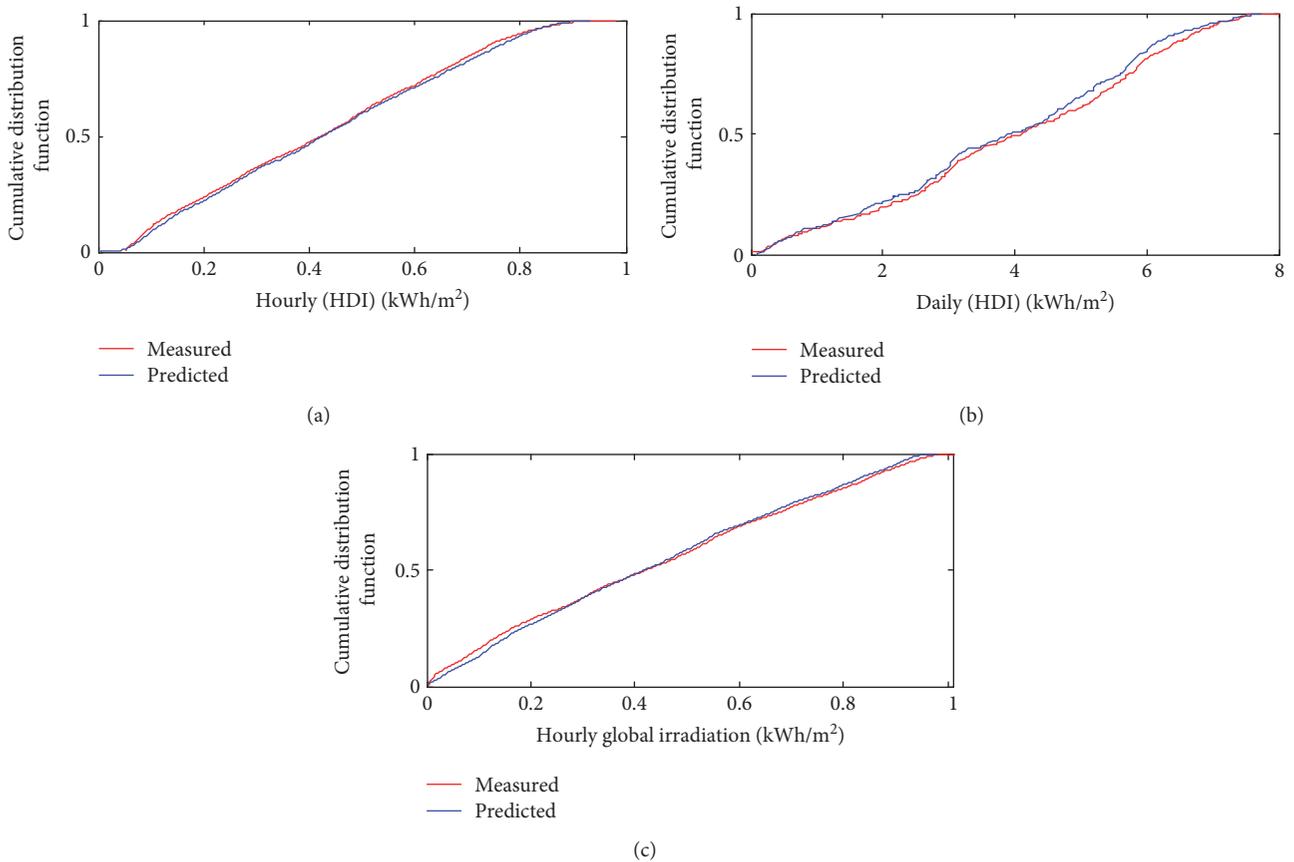


FIGURE 13: Cumulative distribution function between predicted and measured values.

To evaluate our model, we compare its prediction with estimation of  $(HDI)_D$  with another approach. The direct solar radiation analysis, for a specific location, is often calculated starting from the global irradiance data registered. It is estimated by means of the decomposition model [56].

To estimate  $(HDI)_D$  using  $G_D$ , we begin by estimating the diffuse  $Df_D$  using empirical statistical relations between the

diffuse fraction ( $K_d$ ) defined as the ratio of diffuse radiation to global radiation and the clearness index  $K_t$ .

Generally, such relations are locally dependent, even if in many studies it is claimed to develop relations for different locations. Thus, to deal with this issue, we used our data base to plot  $K_d$  versus  $K_t$ . The fit of the resulting scatter points can be expressed by the following equation:

$$K_d = 3.74K_t^3 - 5.75K_t^2 + 1.37K_t + 0.72. \quad (6)$$

The fitting is good since the correlation coefficient is 95% and RMSE equals to 0.07.

Hence, the diffuse  $Df_D$  can be calculated from (6) and  $(HDI)_D$  can be estimated using the following equation:

$$(HDI)_{est} = G_D - Df_D. \quad (7)$$

In order to evaluate results from this procedure, we first plotted estimated daily direct solar radiation  $(HDI)_{D_{est}}$  versus the real values of  $(HDI)_D$  (Figure 11). It is clear that the quality of fit is not as good as in Figure 8; this degradation is seen indeed in  $R$  that becomes equal to 0.93 and  $nRMSE = 17.5\%$  instead of 8.6% for FFNN prediction.

Moreover, by considering data for 2012, we plotted, simultaneously, the estimated direct solar radiation using (6), the one predicted with the proposed FFNN and their relative real data. To better illustrate the difference between the two approaches, we represent a zoom in zones of good (Figure 12(a)) and bad agreements (Figure 12(b)).

From these figures, it appears clearly that prediction with the proposed model is more accurate and exceeds estimation with statistical relations (6).

Finally, we have plotted the cumulative distribution function for both measured and predicted series for each solar component studied (Figures 13(a), 13(b), and 13(c)); as can be seen, obtained results are almost confused with measured data.

Therefore, in point of view statistical test (goodness test), we can conclude that the predicted solar irradiance components with different FFNN's topologies are satisfactory.

## 6. Conclusion

To overcome the lack of accurate long series of solar data, needed for solar systems optimal sizing, especially for concentrating solar plant, an ensemble of FFNN models were trained to predict different components of solar radiation was presented in this research paper. These components are hourly global, daily, and hourly direct solar radiation. Three different architectures of FFNN models were used. In case of having daily global radiation for long series thanks to our model, hourly global radiation can be predicted. The generated data are not only satisfactory ( $R = 0.98$ ,  $RMSE = 0.061 \text{ kWh/m}^2$ ) but surpass those estimated from empirical methods. For concentrating solar plant design, the developed FFNN models can provide long series of either  $(HDI)_h$  or  $(HDI)_D$  with a good accuracy ( $R = 0.98$ ;  $nRMSE = 9.5\%$  and  $R = 0.99$ ;  $nRMSE = 8.05\%$ , resp.). The horizontal diffuse component is determined by subtracting direct radiation from global radiation on a horizontal surface.

Additionally,  $(HDI)_D$  series were estimated using an equation developed for Fez. Results' accuracy shows the superiority of FFNN model.

The cumulative distribution functions between generated and measured data were computed for each studied component, and results confirm the good performances of FFNN models.

Finally, our models can provide synthetic series for different solar radiation components to be used in optimal sizing and planning of solar energy systems, especially for concentrating solar plant.

We look forward to apply this approach in further studies using data from other locations to develop a model that represents all Moroccan's locations.

## Appendix

### A. Calculation of Horizontal Extraterrestrial Solar Radiation $(ER)_h$

To calculate horizontal extraterrestrial solar radiation  $(ER)_h$  for each hour, two geographical parameters must be calculated: solar declination " $\delta$ " and zenith angle  $A_z$ .

- (i) *Solar Declination* ( $\delta$ ).  $\delta$  is the angle between the Sun's rays and the Earth's equatorial plane and depends on the day number  $d_n$  via the day angle  $B$  defined by Iqbal [41] as follows:

$$B = \frac{2\pi(d_n - 1)}{365}$$

$$\delta(\text{rad}) = 0.006918 - 0.399912 \cos B + 0.070257 \sin B$$

$$- 0.006758 \cos 2B + 0.000907 \sin 2B$$

$$+ 0.002697 \cos 3B + 0.00148 \sin 3B. \quad (A.1)$$

- (ii) *Zenith Angle* ( $A_z$ ).  $A_z$  is the angle between the sun and a horizontal surface [41]; it influences the quantity and the quality of the solar radiation; when it is low, its cosines are maximal and the optical path is minimal; and the incident radiation is less absorbed and the solar radiation maximum. The zenith angle is calculated at the middle of the considered hour by the following equation:

$$\cos A_z = \sin(\delta)\sin(La) + \cos(\delta)\cos(La)\cos(HA), \quad (A.2)$$

with " $La$ " the latitude and  $(HA)$  the hour angle calculated from the true solar time  $t_s$  by the following equation:

$$(HA)(\text{degrees}) = 15(12 - t_s). \quad (A.3)$$

The horizontal extraterrestrial solar radiation  $(ER)_h$  in  $\text{Wh m}^{-2}$  is computed by integration of the extraterrestrial solar radiation  $I_e$  in  $\text{W.m}^{-2}$  on the time period.  $I_e$  is calculated by Iqbal [41] as follows:

$$I_e = I_0 E_0 \cos A_z, \quad (A.4)$$

where  $I_0$  is the solar constant ( $I_0 = 1367 \text{ W m}^{-2}$ ),  $E_0$  the eccentricity factor taking into account the variation of the earth-sun distance (dimensionless) [38]:

$$E_0 = 1.00110 + 0.034221 \cos G + 0.001280 \sin G + 0.000719 \cos 2G + 0.000077 \sin 2G. \quad (\text{A.5})$$

## B. Empirical Models for Calculating Mean Hourly Solar Radiation

According to the literature, there are some empirical models developed for calculating hourly solar radiation from daily solar radiation. Among the most accurate models, Liu and Jordan proposed in [57] the following to calculate hourly solar radiation:

$$\frac{G_h}{G_D} = \frac{(\pi/24) (\cos(\text{HA}) - \cos W_{ss})}{\sin W_{ss} - 2\pi((W_{ss})/360)\cos W_{ss}}. \quad (\text{B.1})$$

On the other hand, Collares-Pereira and Rabel (C-P&R) verified the previous model in [58] and propose the following for calculating mean hourly solar radiation:

$$\frac{G_h}{G_D} = (a + b \cos(\text{HA})) \frac{\pi/24 (\cos(\text{HA}) - \cos W_{ss})}{\sin W_{ss} - 2\pi((W_{ss})/360)\cos W_{ss}}, \quad (\text{B.2})$$

where the coefficients  $a$  and  $b$  are defined as follows:

$$\begin{aligned} a &= 0.409 + 0.5016 \sin(W_{ss} - 60), \\ b &= 0.6609 - 0.4767 \sin(W_{ss} - 60). \end{aligned} \quad (\text{B.3})$$

## Nomenclature

ANN:	Artificial neural network
$A_z$ :	Zenith angle
$\delta$ :	Declination angle
$Df_D$ :	Daily diffuse irradiance
$Df_h$ :	Hourly diffuse irradiance
DH:	Day light hours
$E_0$ :	Eccentricity factor
$ER_D$ :	Daily extraterrestrial solar radiation
$ER_h$ :	Hourly extraterrestrial solar radiation
FFNN:	Feed forward neural network
$G_D$ :	Daily global radiation
$G_h$ :	Hourly global irradiation
HA:	Sun hour angle
$(HDI)_D$ :	Daily horizontal direct irradiation
$(HDI)_{est}$ :	Estimated (HDI)
$(HDI)_h$ :	Horizontal direct irradiation
$Hr_D$ :	Daily relative humidity
$Hr_h$ :	Hourly relative humidity
$I_0$ :	Solar constant
$K_t$ :	Clearness index
$K_d$ :	Diffuse fraction
La:	Latitude
MAE:	Mean absolute error
NDI:	Normal direct irradiation
nMAE:	Normalized mean absolute error
nRMSE:	Normalized root mean square error
RMSE:	Root mean square error
$S_F$ :	Sunshine fraction
$T_D$ :	Daily air temperature

$T_h$ :	Hourly air temperature
$t_s$ :	True solar time
$Wd_D$ :	Daily wind direction
$Wd_h$ :	Hourly wind direction
$Wsd$ :	Daily wind speed
$Wsh$ :	Hourly wind speed
$W_{ss}$ :	Sun set hour angle.

## Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] M. A. Atwater and J. T. Ball, "A numerical solar radiation model based on standard meteorological observations," *Solar Energy*, vol. 21, pp. 163–170, 1978.
- [2] D. Elizondo, G. Hoogenboom, and R. W. McClendon, "Development of a neural network model to predict daily solar radiation," *Agricultural and Forest Meteorology*, vol. 71, no. 1-2, pp. 115–132, 1994.
- [3] D. B. Williams and F. S. Zazueta, "Solar radiation estimation via neural network," in *Proceedings of the 6th International Conference on Computers in Agriculture*, pp. 140–146, Cancun, Mexico, 1994ASAE.
- [4] D. Williams and F. Zazueta, "Solar radiation estimation via neural network," in *Proceedings of the 6th International Conference on Radiation, Computers in Agriculture*, pp. 1143–1149, Cancun, Mexico, 1996.
- [5] M. Mohandes, S. Rehman, and T. O. Halawani, "Estimation of global solar radiation using artificial neural networks," *Renewable Energy*, vol. 14, no. 1-4, pp. 179–184, 1998.
- [6] S. M. Al-Alawi and H. A. Al-Hinai, "An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation," *Renewable Energy*, vol. 14, no. 1-4, pp. 199–204, 1998.
- [7] A. Guessoum, S. Boubkeur, and A. Maafi, "A global irradiation model using radial basis function neural network," in *Proceedings of the World Renew Energy Congress (WREC '98)*, pp. 332–336, Oxford, UK, 1998.
- [8] L. Hontoria, J. Riesco, P. Zufiria, and J. Aguilera, "Improved generation of hourly solar radiation artificial series using neural networks," in *Proceedings of the Engineering Applications of Neural Networks (EANN '99)*, Warsaw, Poland, 1999.
- [9] P. J. Zufiria, A. Vázquez-López, J. Riesco-Prieto, J. Aguilera, and L. Hontoria, "A neural network approach for generating solar irradiation artificial series," in *Proceedings of the International Work-Conference on Artificial Neural Networks (IWANN'99)*, Alicante, Spain, June 1999.
- [10] M. Mohandes, A. Balghonaim, M. Kassas, S. Rehman, and T. O. Halawani, "Use of radial basis functions for estimating monthly mean daily solar radiation," *Solar Energy*, vol. 68, no. 2, pp. 161–168, 2000.
- [11] L. Hontoria, J. Riesco, P. Zufiria, and J. Aguilera, "Application of neural networks in the solar radiation field. Obtainment of solar radiation maps," in *Proceedings of the 16th European Photovoltaic for Chemical Engineers*, vol. 3, pp. 385–408, Elsevier, Amsterdam, The Netherlands, 2000.
- [12] A. Sfetsos and A. H. Coonick, "Univariate and multivariate forecasting of hourly solar radiation with artificial

- intelligence techniques,” *Solar Energy*, vol. 68, no. 2, pp. 169–178, 2000.
- [13] G. Mihalakakou, M. Santamouris, and D. N. Asimakopoulos, “The total solar radiation time series simulation in Athens, using neural networks,” *Theoretical and Applied Climatology*, vol. 66, no. 3–4, pp. 185–197, 2000.
- [14] S. A. Kalogirou, “Artificial neural networks in renewable energy systems applications: a review,” *Renewable and Sustainable Energy Reviews*, vol. 5, no. 4, pp. 373–401, 2001.
- [15] L. Hontoria, J. Aguilera, J. Riesco, and P. Zufiria, “Recurrent neural supervised models for generating solar radiation synthetic series,” *Journal of Intelligent and Robotic Systems*, vol. 31, no. 1–3, pp. 201–221, 2001.
- [16] L. Hontoria, J. Aguilera, and P. Zufiria, “Generation of hourly irradiation synthetic series using the neural network multilayer perceptron,” *Solar Energy*, vol. 72, no. 5, pp. 441–446, 2002.
- [17] A. S. S. Dorvio, J. A. Jervase, and A. Al-Lawati, “Solar radiation estimation using artificial neural networks,” *Applied Energy*, vol. 71, pp. 307–319, 2002.
- [18] F. Tymvios, C. P. Jacovides, and S. C. Michaelides, “The total solar energy on a horizontal level with the use of artificial neural networks,” in *Proceedings of the 6th Hellenic Conference of Meteorology, Climatology and Atmospheric Physics*, pp. 26–28, Ioannina, Greece, September 2002.
- [19] S. Kalogirou, S. C. Michaelides, and F. S. Tymvios, “Prediction of maximum solar radiation using artificial neural networks,” in *Proceedings of the 7th World Renewable Energy Congress VII (WREC '02)*, Cologne, Germany, June 2002.
- [20] K. S. Reddy and M. Ranjan, “Solar resource estimation using artificial neural networks and comparison with other correlation models,” *Energy Conversion and Management*, vol. 44, no. 15, pp. 2519–2530, 2003.
- [21] A. Sözen, E. Arcaklioglu, M. Özalp, and E. G. Kanit, “Use of artificial neural networks for mapping of solar potential in Turkey,” *Applied Energy*, vol. 77, pp. 273–286, 2004.
- [22] A. Sözen, E. Arcaklioglu, and M. Özalp, “Estimation of solar potential in Turkey by artificial neural networks using meteorological and geographical data,” *Energy Conversion and Management*, vol. 45, no. 18–19, pp. 3033–3052, 2004.
- [23] A. Mellit, M. Benganem, A. Hadj Arab, and A. Guessoum, “Modeling of global solar radiation data from sunshine duration and temperature using the radial basis function networks,” in *Proceedings of the IASTED International Conference on Modelling, Identification and Control (MIC '04)*, Grindelwald, Switzerland, February 2004.
- [24] L. Hontoria, J. Aguilera, and P. Zufiria, “An application of the multilayer perceptron: solar radiation maps in Spain,” *Solar Energy*, vol. 79, no. 5, pp. 523–530, 2005.
- [25] A. Sözen, E. Arcaklioglu, M. Özalp, and N. Çağlar, “Forecasting based on neural network approach of solar potential in Turkey,” *Renewable Energy*, vol. 30, pp. 1075–1090, 2005.
- [26] F. S. Tymvios, C. P. Jacovides, S. C. Michaelides, and C. Scouteli, “Comparative study of Ångström’s and artificial neural networks’ methodologies in estimating global solar radiation,” *Solar Energy*, vol. 78, no. 6, pp. 752–762, 2005.
- [27] A. Mellit, M. Benganem, and M. Bendekhis, “Artificial neural network model for prediction solar radiation data: application for sizing stand-alone photovoltaic power system,” in *IEEE Power Engineering Society General Meeting, 2005*, vol. 1, pp. 40–44, San Francisco, CA, USA, June 2005.
- [28] G. López, F. J. Batlles, and J. Tovar-Pescador, “Selection of input parameters to model direct solar irradiance by using artificial neural networks,” *Energy*, vol. 30, no. 9, pp. 1675–1684, 2005.
- [29] L. F. Zarzalejo, L. Ramirez, and J. Polo, “Artificial intelligence techniques applied to hourly global irradiance estimation from satellite-derived cloud index,” *Energy*, vol. 30, no. 9, pp. 1685–1697, 2005.
- [30] S. Alam, S. C. Kaushik, and S. N. Garg, “Computation of beam solar radiation at normal incidence using artificial neural network,” *Renewable Energy*, vol. 31, no. 10, pp. 1483–1491, 2006.
- [31] H. K. Elminir, Y. A. Azzam, and F. I. Younes, “Prediction of hourly and daily diffuse fraction using neural network, as compared to linear regression models,” *Energy*, vol. 32, no. 8, pp. 1513–1523, 2007.
- [32] G. López and C. A. Gueymard, “Clear-sky solar luminous efficacy determination using artificial neural networks,” *Solar Energy*, vol. 81, no. 7, pp. 929–939, 2007.
- [33] T. Khatib, A. Mohamed, K. Sopian, and M. Mahmoud, “Solar energy prediction for Malaysia using artificial neural networks,” *International Journal of Photoenergy*, vol. 2012, Article ID 419504, 16 pages, 2012.
- [34] M. Tiris, Ç. Tiris, and I. E. Türe, “Correlations of monthly-average daily global, diffuse and beam radiations with hours of bright sunshine in Gebze, Turkey,” *Energy Conversion Management*, vol. 37, pp. 1417–1421, 1996.
- [35] G. López, M. A. Rubio, and F. J. Batlles, “Estimation of hourly direct normal from measured global solar irradiance in Spain,” *Renewable Energy*, vol. 21, pp. 175–186, 2000.
- [36] H. Loutfi and A. Khtira, “Stochastic analysis and generation of synthetic sequences of daily global solar irradiation: Rabat site (Morocco),” *Renewable Energy*, vol. 2, pp. 129–138, 1992.
- [37] P. Poggi, G. Notton, M. Muselli, and A. Louche, “Stochastic study of hourly total solar radiation in Corsica using a Markov model,” *International Journal of Climatology*, vol. 20, pp. 1843–1860, 2000.
- [38] R. Aguiar and M. Collares-Pereira, “TAG: a time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation,” *Solar Energy*, vol. 49, no. 3, pp. 167–174, 1992.
- [39] J. A. Davies and D. C. McKay, “Estimating solar irradiance and components,” *Solar Energy*, vol. 29, pp. 55–64, 1982.
- [40] C. Gueymard, “Critical analysis and performance assessment of clear sky solar irradiance models using theoretical and measured data,” *Solar Energy*, vol. 51, pp. 121–138, 1993.
- [41] M. Iqbal, *An Introduction to Solar Radiation*, Academic Press, Canada, 1983.
- [42] A. Mellit, “Artificial intelligence technique for modelling and forecasting of solar radiation data: a review,” *International Journal of Artificial Intelligence and Soft Computing*, vol. 1, pp. 52–76, 2008.
- [43] S. A. Kalogirou and A. Sencan, “Artificial Intelligence Techniques in Solar Energy Applications,” in *Theory and Applications*, R. I. Manyala, Ed., p. 444, Book, 2010.
- [44] S. Rehman and M. Mohandes, “Artificial neural network estimation of global solar radiation using air temperature and relative humidity,” *Energy Policy*, vol. 36, pp. 571–576, 2008.
- [45] A. Mellit, H. Eleuch, C. Elaoun, and A. Massi Pavan, “An adaptive model for predicting of global, direct and diffuse hourly solar irradiance,” *Energy Conversion and Management*, vol. 51, pp. 771–782, 2010.

- [46] A. Teke, H. B. Yildirim, and Ö. Çelik, "Evaluation and performance comparison of different models for the estimation of solar radiation," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1097–1107, 2012.
- [47] A. Jain, J. Mao, and K. Mohiuddin, "Artificial neural networks: a tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, 1996.
- [48] The MathWorks, "MATLAB," 2013, <http://www.mathworks.com/products/neuralnet>.
- [49] T. Krishnaiah, S. Srinivasa Rao, K. Madhumurthy, and K. S. Reddy, "Neural network approach for modelling global solar radiation," *Journal of Applied Sciences Research*, vol. 3, no. 10, pp. 1105–1111, 2007.
- [50] A. de Miguel, J. Bilbao, R. Aguiar, H. Kambezidis, and E. Negro, "Diffuse solar irradiation model evaluation in the North Mediterranean belt area," *Solar Energy*, vol. 70, pp. 143–153, 2001.
- [51] A. K. Yadav and S. S. Chandel, "Solar radiation prediction using artificial neural network techniques: a review," *Renewable and Sustainable Energy Reviews*, vol. 33, pp. 772–781, 2014.
- [52] K. Dahmani, G. Notton, C. Voyant et al., "Multilayer perceptron approach for estimating 5-min and hourly horizontal global irradiation from exogenous meteorological data in locations without solar measurements," *Renewable Energy*, vol. 90, pp. 267–282, 2016.
- [53] A. Ahmad and T. Anderson, "Global solar radiation prediction using artificial neural network models for New Zealand," in *Proceedings of the 52nd Annual Conference, Australian Solar Energy Society*, pp. 141–150, Melbourne, 2014.
- [54] T. Khatib, A. Mohamed, and K. Sopian, "A review of solar energy modeling techniques," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 5, pp. 2864–2869, 2012.
- [55] A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: a review," *Progress in Energy and Combustion Science*, vol. 34, no. 5, pp. 574–632, 2008.
- [56] C. Renno and F. Petit, "Choice model for a modular configuration of a point-focus CPV/T system," *Energy and Buildings*, vol. 92, pp. 55–66, 2015.
- [57] B. Y. H. Liu and R. C. Jordan, "The interrelationship and characteristic distribution of direct, diffuse and total solar radiation," *Solar Energy*, vol. 4, no. 3, pp. 1–19, 1960.
- [58] M. Collares-Pereira and A. Rabl, "The average distribution of solar radiation—correlations between diffuse and hemispherical and between daily and hourly insolation values," *Solar Energy*, vol. 22, no. 2, pp. 155–164, 1979.

