Renewable Energy 126 (2018) 855–864

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Predicting solar energy generation through artificial neural networks using weather forecasts for microgrid control

Fermín Rodríguez [a, b, *], Alice Fleetwood [a, b], Ainhoa Galarza [a, b], Luis Fontán [a, b]

[a] Ceit, Manuel Lardizabal 15, 20018 Donostia/San Sebastián, Spain
[b] Universidad de Navarra, Tecnun, Manuel Lardizabal 13, 20018 Donostia/San Sebastián, Spain

ARTICLE INFO

Article history:
Received 11 December 2017
Received in revised form
9 March 2018
Accepted 26 March 2018
Available online 27 March 2018

Keywords:
Microgrid
Solar energy
Prediction model
Artificial neural network

ABSTRACT

This paper proposes an artificial neural network (ANN) to predict the solar energy generation produced by photovoltaic generators. The intermittent nature of solar power creates two main issues. Firstly, power production and demand have to be balanced to ensure the control of the whole system, and the inherent variability of clean energies makes this difficult. Secondly, energy generation companies need a highly accurate day-ahead or intra-day estimation of the energy to be sold in the electricity pool. For the tool developed in this paper, we address the issue of the complexity of control in systems that are based on solar energies. The tool's ability to predict the parameters that are involved in solar energy production will allow us to estimate the future power production in order to optimise grid control. Our tool uses an ANN which we developed using MATLAB® software. The results were validated by analysing the root mean square error of the prediction for days outside the database used for training the ANN. The difference between the actually produced and predicted energy is about 0.5–9%, meaning that the accuracy of our tool is sufficient enough to be installed in systems which have integrated solar generators.

1. Introduction

The traditional energy network has been undergoing many changes in the last few years due to demand in growth, environmental issues and market deregulation. These changes have raised challenges that traditional energy networks may not be able to overcome [1,2]. Distributed generation combined with the incorporation of renewable technologies may provide a solution that the traditional network cannot give on its own. Some benefits of distributed generation are reliability enhancement, power loss reduction and integration of clean energies. However, integrating distributed generation into the traditional network raises the number of agents that must be controlled, so the complexity of the grid also increases. In addition, negative effects such as not being able to ensure frequency and voltage stability at the grid or an unsecure situation in the traditional network control can emerge. In this context, microgrids can be used as a platform to integrate distributed generation technologies into the traditional network and overcome these negative effects.

A microgrid is a distributed power network which integrates loads, power generators and storage technologies with control components in order to operate connected to, or islanded from, the main network [3]. In islanded mode, the local generators are the only source of power supply for the loads connected to the microgrid [4]. If the microgrid is connected to the network, a power exchange will occur between both systems according to supply and demand levels. Hence, microgrids can operate alternatively between both modes, islanded or connected. For instance, when there are external faults or dynamic changes occur in the main grid, microgrids can change from connected to islanded mode in order to continue operating.

The benefits of microgrids include a reduction of losses and costs and an increase in the contribution of renewable sources in system [4]. In addition, they have been used in recent years as a solution to help to mitigate reliability issues [2]. However, the literature [4] stresses that these benefits will only be achieved through the optimisation and control of the microgrid. Therefore, the coordinated operation and control of elements such as distributed generation and storage devices is central to the concept of microgrids [5].

Incorporating renewable generation in microgrids is an important factor in achieving the goals of reducing the losses of the

Abbreviations: ANN, Artificial Neural Network.

* Corresponding author.
E-mail address: frlalanne@ceit.es (F. Rodríguez).

https://doi.org/10.1016/j.renene.2018.03.070
0960-1481/

overall system and improving the system's stability. The advancement of renewable energies, such as solar photovoltaic (PV) and wind power, is in fact one of the key drivers for their implementation in microgrids [2]. This increase in renewable energy sources is due to recent accelerated technology developments, making renewables affordable and readily available [2,3]. In addition to this, the environmental benefits of renewable energies are making them a popular choice for supplementing traditional generators across the grid [3].

While clean technologies address the need for a sustainable source of energy, their inherent variability and dependence on weather conditions introduces further complications to the power network [6–8]. Because of this, the methods that have been used for balancing generation and consumption before these new technologies appeared must be updated to incorporate them. For example, due to the variability of the power demand on microgrids, the first step taken to improve these control methods consisted of forecasting the demand of the loads in the short term [9–11], as it is necessary to know the amount of energy that will be consumed in order to balance the system with the power generators in an efficient way. The second step, which is addressed in this paper, consisted of forecasting the energy generated by the renewable energies in order to continue updating the traditional methods.

As it is known, solar power generation has a direct relationship with the existing local weather conditions [6]. It must be taken into account that solar power is not only diurnal in nature, but also varies throughout the day with the changes in the level of solar irradiation. This is why a calculation of future power production by this type of source is needed if a balanced system is desired. However, these future power calculations require knowledge of the future weather parameters that affect these technologies.

Predicting weather parameters for energy generation calculations is not a new undertaking, and several methods are provided in literature. Different models and parameters of influence, such as cloudiness, irradiance and so on, have been used [8]. Nevertheless, other authors propose the forecasting of renewable generation as a method for achieving this goal [2,5,12,13]. Forecasting consists of predicting future trends based on historical data. Use of this procedure can result in the desired reduction of uncertainty in the use of clean energies.

This paper presents a new prediction model for solar energy generation that will be implemented in a microgrid. Although different prediction methods are proposed in the literature [12,14], the tool developed here is based on an artificial neural network (ANN). The ANN we have designed is able to predict the solar energy that will be produced by the PV generators with a high standard of accuracy in the short term, 10 min. This information is given to the control of the microgrid, so that it can take the right decisions at each moment, and in so doing improve its reliability.

As we have chosen ANNs to develop the tool, we briefly summarise these networks and the possibilities provided by MATLAB® software to work with them. In addition, an iterative process was used to fix some parameters of the final ANN, so the main given steps and the reasons for those steps are also explained. Finally, our predictions for three different days (sunny, partially cloudy and cloudy) throughout the year are shown and contrasted with the real measured values.

2. Prediction methods

Several methods have been used to forecast the weather parameters needed to obtain the solar energy produced by the PV generators. These can be classified into two different categories: single model approaches and multiple model approaches.

Concerning the single model approaches, these can be separated again into two different groups, physical models, such as numerical weather prediction, and statistical methods [15]. Statistical methods include the persistence method, autoregressive models, artificial neural networks and support vector machines. The strength of these methods depends on two main factors: the prediction horizon and the availability of reliable historical data. The multiple model approaches stem from a combination of single model methods. More et al. [14] look favourably on the multiple models, arguing that the weather parameters cannot be obtained from a single method due to their variability.

In addition, multiple model approaches add greater complexity to the prediction of the parameters. While there are cases in the literature [12,14] that propose this path, there is a significant portion of the literature [6,16,17] that has obtained satisfactory results from the single model approaches. Due to the lack of consensus in the literature and the results that have been obtained using ANNs in previous studies, the single model approach has been chosen for this study.

2.1. Artificial neural networks

ANNs are a form of artificial intelligence which is based on the computerization of human abilities. It must be taken into account that traditional computer processing happens in a very different way if we compare it with the way the human brain works. The complex, nonlinear and parallel functioning of the human brain allows it to outperform even the best conventional computers in certain areas, particularly in relation to pattern recognition and perception [18]. Artificial intelligence has been able to increase the capabilities of computers by replicating the human biological information processing system. ANNs in particular are able to replicate the biological neural network that exists in the human brain and thus recognize patterns and predict future values.

The elements of an ANN are artificial neurons; they can be represented by neuron k in Fig. 1. These neurons, or processing units [18], act as nonlinear summing devices and are located in layers and connected to each other by adjustable connection weights, also known as synaptic junctions [19].

The neural model in Fig. 1 can be expressed mathematically as,

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \tag{1}$$

$$y_k = \varphi(u_k + b_k) \tag{2}$$

where the input values of a node, x_j, replicate the dendrites of the biological neuron and are multiplied by their respective connection weights, w_{kj}, and then summed. The summation from Eq. (1) is applied in Eq. (2) to the function of node φ, along with the bias b_k, giving as a result the activation value of the node y_k, which represents the axon of the biological neuron. This activation value then acts as an input to the next linked node. Biases b_k are applied externally to the neuron's summing junction; they can be positive or negative, which increases or decreases the net input of the activation. Further information about the relationship between the human brain and ANNs can be found in Refs. [18,19].

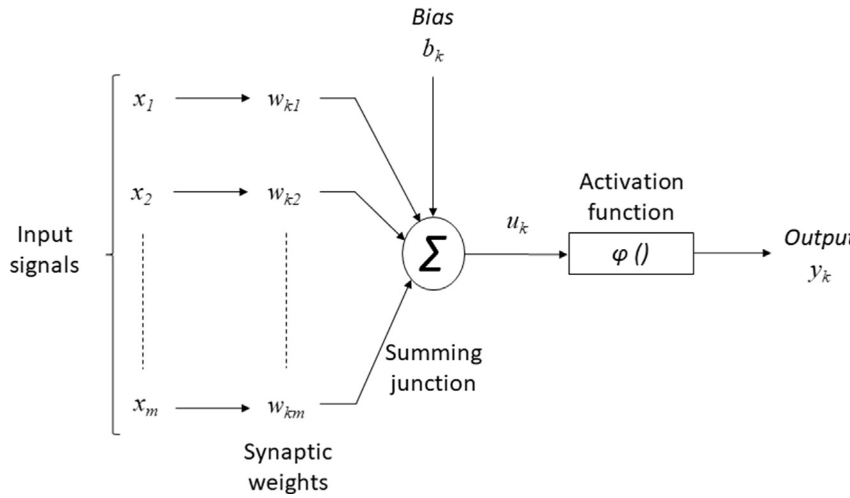

Fig. 1. Artificial neural model.

Producing reasonable outputs from previously unseen inputs is known as generalization [18]. This is achieved through correctly training the network, and it is what allows ANNs to obtain good results on different tasks such as prediction and classification in pattern recognition [20]. ANNs are trained by learning algorithms, a process which consists of modifying the initial synaptic weights in an orderly way until the minimal error is produced by the network [14,18,21]. After reviewing the literature, we chose the Levenberg-Marquardt (LM) learning algorithm due to its robustness [16,22].

In addition, there are three different types of ANNs, as classified by Haykin in Ref. [18]. The types are single-layer feedforward networks, multi-layer networks, and recurrent networks, although each of them can be divided into further categories. After testing several of the recurrent network types, we chose the layer recurrent neural network architecture to develop our tool. This type of ANN is similar to a feedforward network in that the information passes through the network in the direction from input to output. However, the outputs from the hidden layer of the recurrent network are connected back to the input of the hidden layer through time delays.

As the design of the network follows a trial and error process in order to fix its parameters, many changes must be made before the final network can be implemented. Changes that can be made include adjusting the number of hidden neurons or the time delay of the network. In order to compare the results that were obtained through this process, error histogram and regression plots given by the MATLAB *nntraintool* toolbox were used.

Finally, the root mean square error (RMSE) given by Eq. (3) was used in this work to analyse the difference between real and predicted data. This is the most common way to measure the efficiency and correctness of a network during the testing and validation process [23]. However, sometimes the RMSE value is not accurate enough due to the fact that some values can ruin the global result. Thus, some histograms were drawn to see the error dispersion in the ANN prediction.

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(X_i - X_i')^{\,2}} \tag{3}$$

where X_i represents the actual measured values; X_i' represents the predicted points and N is the number of predictions that have been made.

2.2. The importance of the database

Selecting data is arguably the most important step in the process of designing an artificial neural network [22]. It is essential to the performance of the neural network that the database used for learning has the highest accuracy. Proximity to site and validity of data are the most important factors in considering whether data is acceptable for use [19]. For the development of this tool, we obtained a database from Euskalmet (http://www.euskalmet.euskadi.eus), which is a governmental agency for providing meteorological data in the Basque Country. The variables obtained through Euskalmet database extraction tool were irradiation, atmospheric pressure, relative air humidity and air temperature. Historical data provided by Euskalmet is recorded in 10-min intervals.

The literature available on recurrent neural networks suggests that the use of historical data across an extended time period is preferable [16,22,24,25]. For example, Kemmoku et al. [24] used data from a period of six years to predict the daily insolation for the next year. Larson et al. [25] forecasted hourly-averaged day-ahead power output from three years of data. Mellit et al. [16] selected sections of data across a two-year period for 24 hourly predictions. Mohammed et al. [22] used data measured for three years to train, test and validate an hourly prediction ANN. The precedents of Mellit et al. [16], Mohammed et al. [22] and Larson et al. [25] were taken into account, due to the short-term forecasting requirements of our tool. As such, we chose two years' worth of data; the dataset covers the period from the start of January 2015 until the end of December 2016, which is the most recent data that Euskalmet provides.

Before the neural network can be constructed, it is essential to have a clear understanding of the input parameters that will be used. The inputs which are used in an ANN must be variables that characterize the process. Thus, an analysis was undertaken to determine which variables would make up the network's inputs. Through this process, variables with little or no contribution to the network output were eliminated [19].

This analysis was done using the Excel data analysis tool for

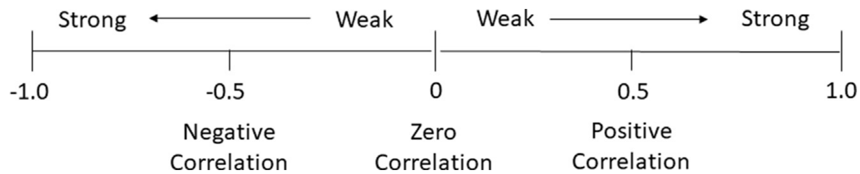

Fig. 2. Pearson correlation coefficient scale.

Table 1
Pearson R correlation coefficients on sample days of pressure, humidity, and temperature versus irradiation.

Sample Day	Pressure	Humidity	Temperature
01/01/2015	0.15	−0.65	0.48
15/04/2015	0.22	−0.24	0.66
13/09/2015	−0.32	−0.51	0.37
06/02/2016	−0.08	−0.43	0.29
26/05/2016	−069	−0.87	0.86
21/09/2016	−0.53	−0.81	0.65

Pearson correlation. Correlation analyses are used in the statistical evaluation of data to check if there is a true relationship between two different variables. A Pearson correlation is represented through its degree of strength, varying from a strong to a weak correlation or a zero correlation [26]. The symbol *r* is used for the correlation coefficient, which ranges from -1 to $+1$, as shown graphically in Fig. 2.

Once the variables have been chosen, it is necessary to collect all the data and prepare them before they are used in the training of the network. Before training the network, selected data inputs and target outputs were normalized. Maximum and minimum values for each input variable were used to scale them from 0 to 1. This process is necessary to obtain faster learning and better results, as was demonstrated in Ref. [19]. Since the input data to the artificial neural network is normalized, so too is the output. Thus, the output needs to be de-normalized before it can be compared with real measurements.

3. Solar ANN design and implementation

3.1. Data analysis

Since the target of the solar ANN is irradiation, one of the inputs to the network must be irradiation. For this tool, we decided that for each predicted value the irradiation measurements of the previous 24 h would be used. In addition, the relationship between irradiation and atmospheric pressure, air temperature and relative air humidity was analysed to determine whether these variables should be also included as inputs in the solar ANN. The results of the correlation analysis of these variables for a few days are shown in Table 1.

Fig. 3 shows the correlation between pressure, humidity, and temperature versus solar irradiation. As can be seen in these figures and in Table 1, the correlation coefficients vary significantly between sample days depending on what season the measurement was done. Concerning the correlation coefficient values for pressure vs. solar irradiation, it can be said that there is no correlation between these parameters owing to the fact that the values are distributed in both directions independently of the season, so this variable is not going to be used as an input to the ANN. The correlation values for relative air humidity vs. solar irradiation and temperature vs. solar irradiation show that there is a negative and positive relationship between these variables, respectively. However, depending on the season and the sample day these tests are run on, the correlation coefficients show stronger or weaker relationships. This indicates that the relationships between these variables are not stable, and for that reason they were not included as inputs in the ANN.

Moreover, due to the temporal variation of solar irradiation, we analysed the variation throughout the year. This analysis was done to determine if the database should be divided into seasons, as shown by Alados et al. in Ref. [27]. The variation of solar irradiation between seasons was analysed using sample days from the database. The results are shown in Table 2.

As can be seen in Table 2 and Fig. 4a, the maximum irradiation value during the day is significantly higher in summer than in winter. Following this analysis, the data for the solar ANN was divided into seasons so that the season for each day could be given as an input (1–4) to the neural network. The database was divided as follows: winter (1), spring (2), summer (3) and autumn (4).

Furthermore, from this analysis it was evident that sunrise and sunset times vary throughout the year, so the number of sun hours is again higher in summer than in winter. As can be seen in Fig. 4b, there is a direct relationship between the number of sun hours and the accumulated irradiation throughout a day. The Pearson correlation test has given a value of 0.84, so it can be said that a positive strong correlation exists between both variables. To accommodate this, the data has been split among different periods: Early Morning (1) intervals between 00:00 and 5:50, Morning (2) intervals between 06:00 and 11:50, Afternoon (3) intervals between 12:00 and 17:50, and Night (4) intervals between 18:00 and 23:50.

3.2. ANN development

From the data analysis, a vector of 146 values was given to the network as training input, *X*, and just one single training target, *T*. The first of these input values was the season of the day being predicted, the second was related to the time of day, and the remaining 144 vales represented the irradiation values in 10 min intervals for the previous 24 h. The single output from the ANN (*T*) is the predicted irradiation value.

To determine the optimal design of the solar ANN, data from January to April 2015 was used to test different network architectures. The performance and error of predictions were evaluated for a network with different numbers of hidden neurons and delays to choose the best architecture. To determine the delay, the number of hidden neurons was held constant and the value of the delay was changed incrementally in each test. And the procedure for fixing the number of hidden neurons was exactly the same, but in this case the delay was held.

It has been verified that an ANN with the same architecture and trained on different times will make similar predictions, but not

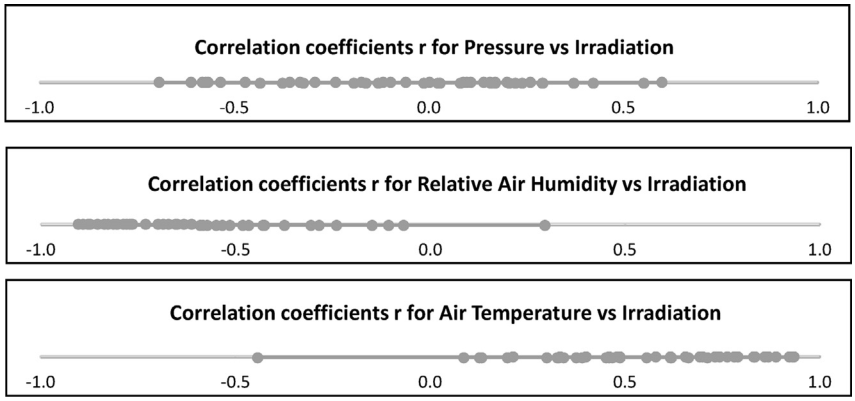

Fig. 3. Correlation coefficients between weather parameters and irradiation.

Table 2
Changes in sunrise and sunset times and maximum irradiation across seasons.

	Sunrise	Sunset	Maximum Irradiation (W/m^2)	Sun Hours
Winter				
01/01/2015	7:40	16:40	430	9
02/12/2016	6:50	16:40	416	9:50
Summer				
15/07/2016	4:50	20:00	993	15:10
09/08/2016	4:40	19:40	972	15:00
Spring				
02/04/2015	5:50	18:40	891	12:50
04/05/2015	5:00	19:20	791	14:20
Autumn				
21/09/2016	5:50	18:20	766.3	12:30
02/10/2016	6:10	18:00	715.1	11:50

give exactly the same values. To ensure the best architecture, each test in Table 3 was repeated five times. While the "RMSE Training Data" has been obtained by the average of the RMSE values from the five tests from 2 January 2015 to 11 January 2015, the "RMSE Validation Data" was obtained from the average of the RMSE values from 2 January 2017 to 11 January 2017.

Once the tests had been run and the errors had been compared, it was decided to create an ANN with a 1:2 delay and 10 hidden neurons. A summary of the designed network is given in Table 4 and the network diagram is in Fig. 5, where *W* is related to connection weights between neurons and *b* are the bias which increases or decreases the net input of the activation.

4. Results and discussion

Once the solar ANN was trained, its performance was evaluated. Fig. 6 shows the regression of the test data, which has a value of 0.98981 at the 9th iteration and demonstrates a strong relationship between network outputs and target values. Although a few outliers can be seen in the figure, most training arrays produced results along the line of best fit, given in the figure as a continuous line.

Moreover, Fig. 7a shows the result that the trained network provided in predicting a day within the training database. While the real measurements of solar irradiation are given by the discontinuous line, the values predicted by the network are given by the continuous one. This performance was checked to evaluate the success of the trained ANN. The RMSE for this sample day was 27.67 W/m^2, which can indicate that the training was completed

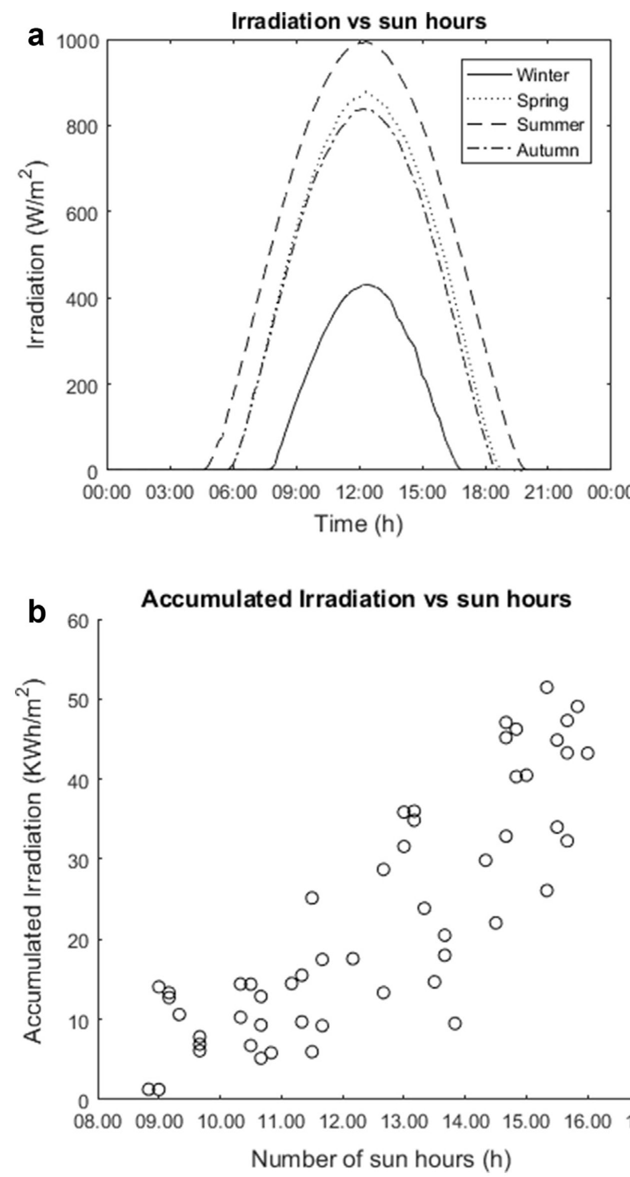

Fig. 4. a) Relationship between irradiation and sun hours in different seasons, b) Correlation between accumulated irradiation and sun hours throughout sample days.

Table 3
Solar ANN architecture design tests and error evaluation for test and non-test data.

Fix	Test	Neurons	Delays	Iterations	RMSE Training Data (W/m^2)	RMSE Validation Data (W/m^2)
Delay	1	10	1:2	6	17.60	19.84
	2	10	1:3	6	19.34	20.21
	3	10	1:4	6	19.07	20.50
	4	10	1:5	6	19.49	20.64
	5	10	1:6	6	19.02	20.00
No Neurons	6	5	1:2	6	20.13	20.72
	7	15	1:2	6	20.05	21.04
	8	20	1:2	6	19.39	20.64

Table 4
Summary of solar ANN design and architecture.

Network type	Layer recurrent network
Inputs	146 – season, time, irradiation
Outputs	1 – irradiation
Number of Layers	3 – input, hidden, output
Number of Hidden Neurons	10
Number of Delays	1:2
Functions	Log-sigmoid, linear
Learning Algorithm	Levenberg-Marquardt

successfully. It must be taken into account that the lower the RMSE, the better the prediction is. The lowest RMSE that was obtained with the tool on a predicted day was 3.38 W/m^2. Nevertheless, the RMSE do not usually represent the error made by the ANN with too much accuracy, due to the fact that punctual large errors can distort the overall result. This is the reason why in Fig. 7b an error histogram is shown with a Normal distribution of that error.

As can be seen in Fig. 7a, the difference between real and predicted values were not too large, and the trend of the prediction line follows the actual line at every moment. In addition, if Fig. 7b is analysed, the major density of errors is located between −20 and 20 W/m^2, which is inside the bell of the normal distribution. Furthermore, it can be observed in Fig. 7b how the values that distort the RMSE value are outside the normal distribution bell. These errors are related with the ups and downs of the solar irradiance that make a higher error than if these do not appear.

After doing this analysis and checking that the results obtained by the ANN for other sample days are similar to the day shown at Fig. 7a and b, the training process of the ANN was concluded.

After the training process was concluded, we checked to see whether the ANN was able to predict the irradiation of days outside the database that had been used for training. Figs. 8a, 9a and 10a show the results of predicting three different days outside the database, and Figs. 8b, 9b and 10b show the error in those predictions through an error histogram and the normal distribution of this error. These three days were chosen in different seasons in order to determine whether the ANN was able to predict more sun hours in the summer than in the winter; more sun hours is related with more irradiation. In addition, these days were chosen in order to see the ANN's degree of accuracy with different percentages of cloudiness.

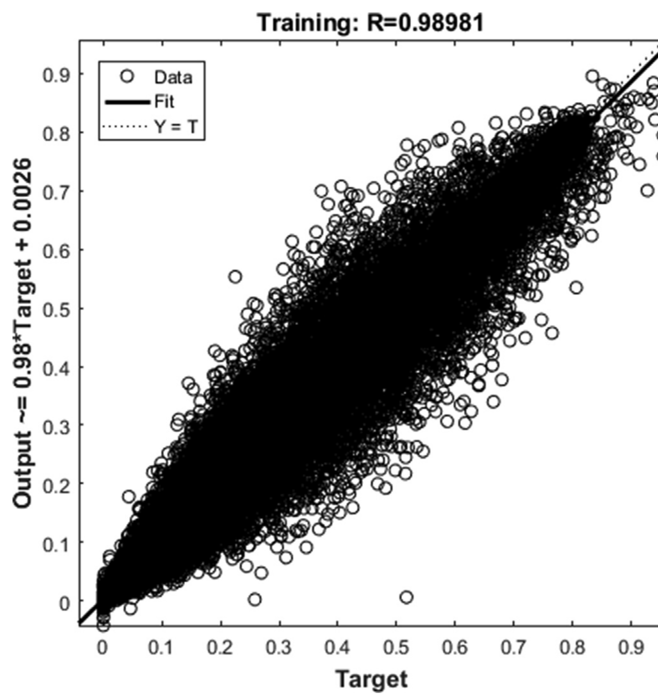

Fig. 6. Obtained regression for Solar ANN.

Fig. 8a shows the prediction that the ANN made on 2 January 2017. For Figs. 8a, 9a and 10a, the discontinuous line indicates the real measurements, while the continuous line represents the predicted values. The RMSE that the ANN calculated on this day is 29.89 W/m^2. Although it is true that the RMSE is higher than in the

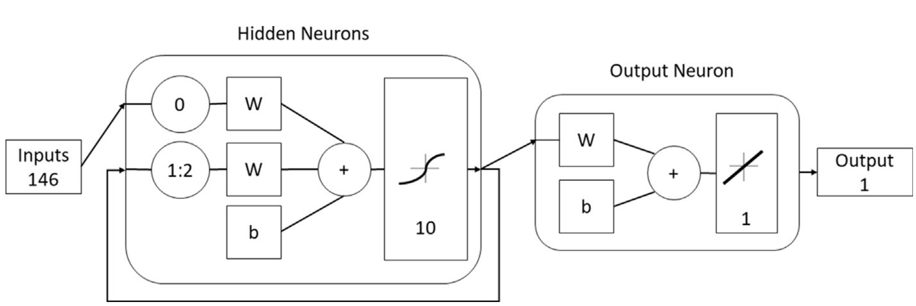

Fig. 5. Solar ANN design showing number of inputs, outputs, and hidden neurons.

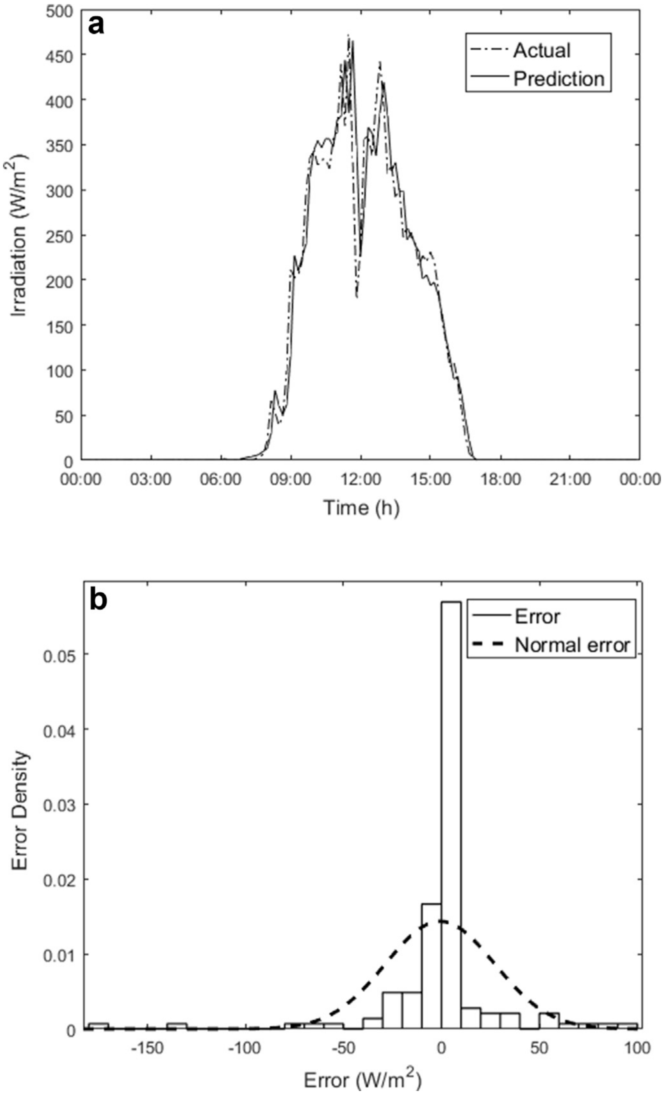

Fig. 7. a) Prediction of a day within the database, b) histogram and normal distribution of error committed.

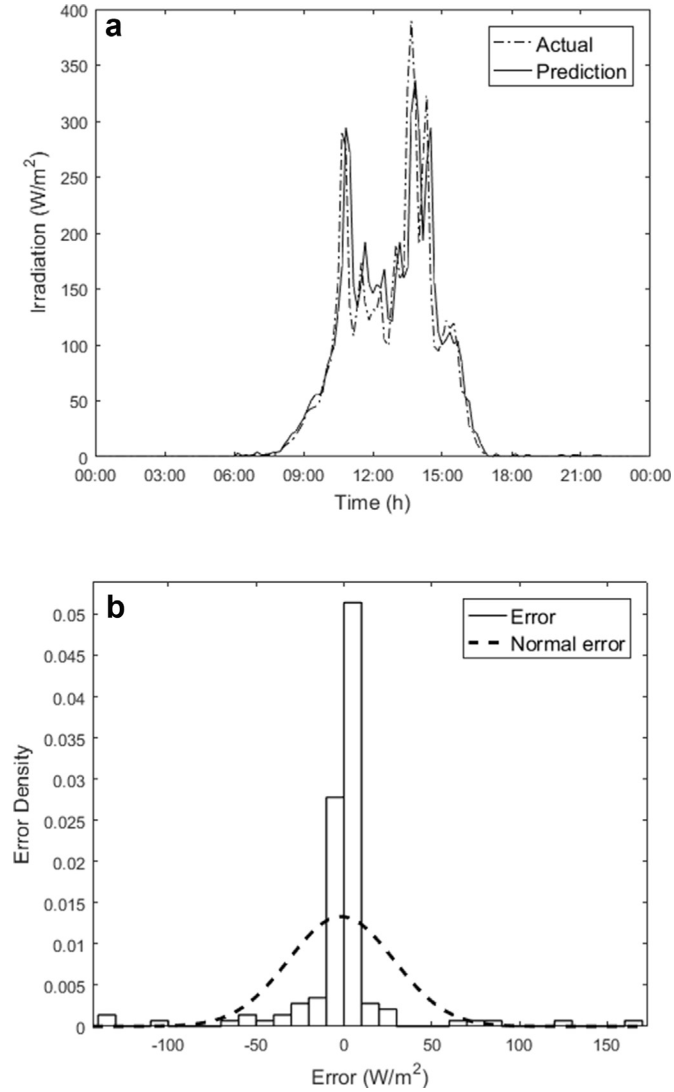

Fig. 8. a) Prediction of solar irradiation on 2 January 2017, b) histogram and normal distribution of error.

previous case where the RMSE is 27.67 W/m^2, we can see in Fig. 8b how the great majority of the errors are inside the bell and near zero error.

On the one hand, Fig. 9a shows the prediction that the ANN made on 6 April 2017. As can be seen in Fig. 9a, the number of sun hours and solar irradiation have increased relative to the prediction made for 2 January 2017. Therefore, we can deduce that the ANN works properly. The RMSE that the ANN calculated on this day is 73.00 W/m^2. Although it is true that the RMSE is higher than in the previous cases, it has been taken into account the fact that the more sun hours there are in a day, the more mistakes will be made. In addition, we can deduce that the more periods with clouds a day has, the higher the RMSE will be.

On the other hand, if we analyse Fig. 9b we can observe how the great majority of the errors are located next to the zero error. As in the previous case, there are some cases which are outside the bell curve. These are the values which are responsible for the distortion that affects the RMSE value. Thus, the prediction given by the trained ANN can be accepted.

Fig. 10a shows the prediction that the ANN made on 4 July 2017. In this figure, we can observe how both the predicted and measured values overlap. The RMSE that the ANN calculated on this day is 14.67 W/m^2. Fig. 10b shows that the large majority of the errors are inside the bell curve and near zero error and few values are out of the bell. We can conclude, after analysing these three different days, that the more sunny the day we want to predict, the better accuracy we get with our ANN. Table 5 shows the relationship between the RMSE and the degree of cloudiness. Although it is true that on some sample days the RMSE reaches values of 77.76 W/m^2, when they are analysed in greater depth, as is done in Figs. 8b, 9b and 10b, it can be seen that the highest density of errors are close to zero.

We also analysed the power generation calculated from the forecasted weather parameters. The power output from a PV solar generator is given by Eq. (4).

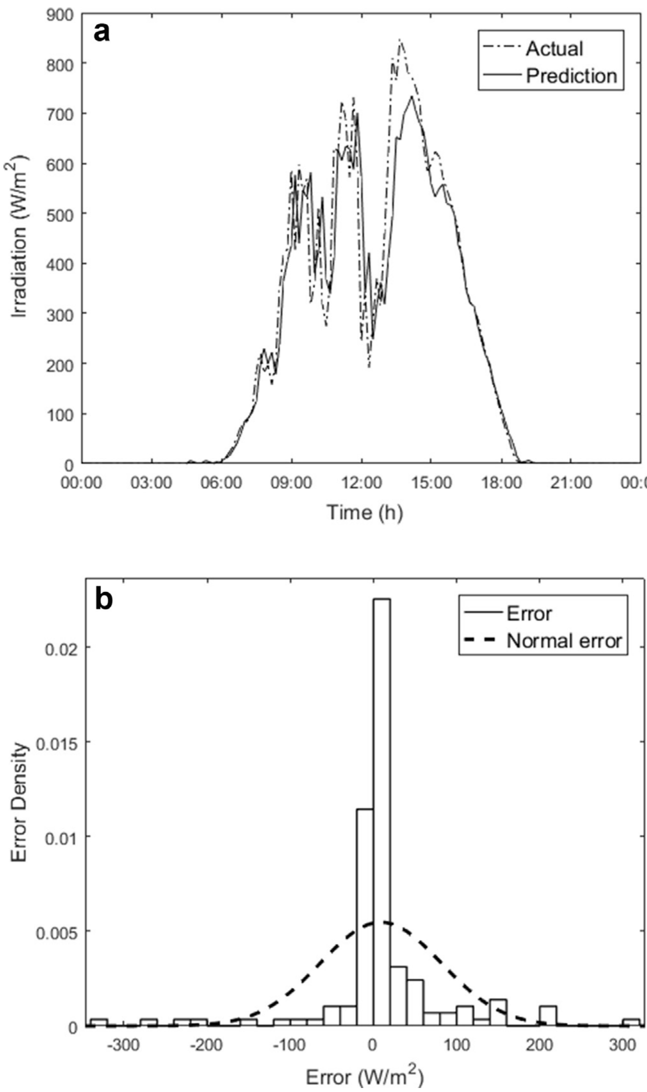

Fig. 9. a) Prediction of solar irradiation on 6 April 2017, b) histogram and normal distribution of error.

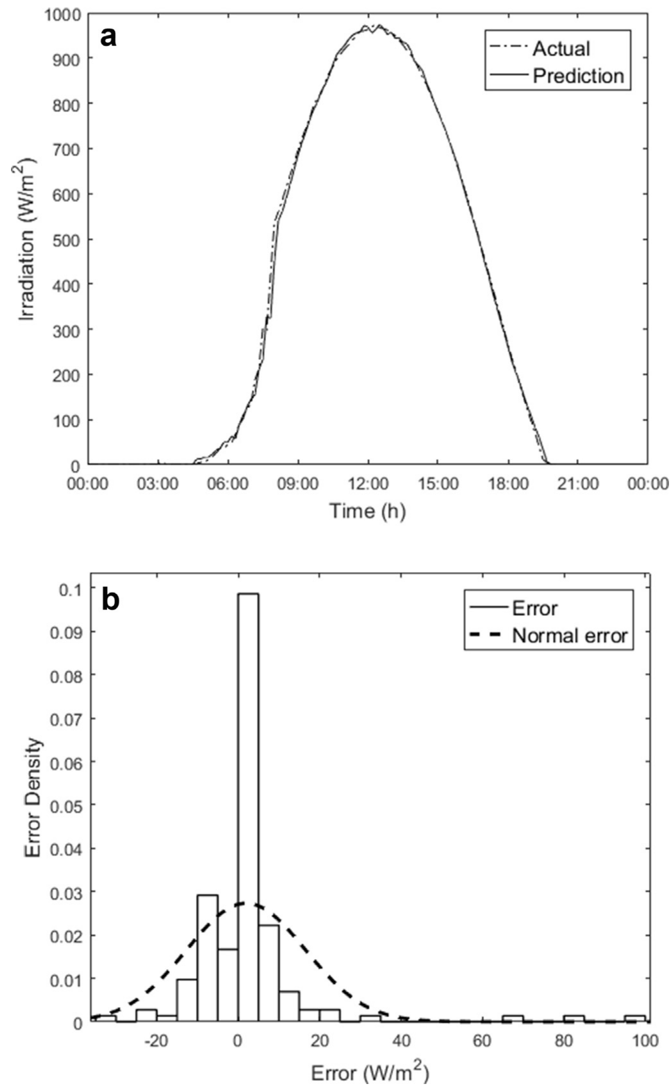

Fig. 10. a) Prediction of solar irradiation on 4 July 2017, b) histogram and normal distribution of error.

$$P_S = \eta SI(1 - 0.005(t_0 - 25)) \tag{4}$$

where P_S is the generated electrical power (W), η is the conversion efficiency coefficient, S is the area of the module (m^2), I is the solar irradiation (W/m^2) and t_0 is the measured temperature (°C). To calculate the electrical power generated, one PV solar panel was chosen whose parameters are $\eta = 17.59\%$ and $S = 1.6767\ m^2$.

Fig. 11a and b represent the difference between the predicted and simulated energy generated throughout the chosen panel in two different situations; while Fig. 11a is an overcast day, Fig. 11b is a cloudless day. Keeping in mind the fact that this tool will be incorporated in microgrids not only to estimate future energy generation in the short term but also to improve control, it is also important to ensure the accuracy of the generated energy predictions throughout the day. Fig. 11a shows the predicted and actual energy generated for 2 January 2017, and the difference between the predicted and actual energy generated is 12.4 Wh and the percentage error is 3.55%. Fig. 11b, is related to the 4 July 2017and the difference between the predicted and actual energy generated is 14.5 Wh and the percentage error is 0.61%. Hence, the more sun hours and the less the cloudiness, the better the energy prediction

Table 5
Variation of the RMSE depending on degree of cloudiness.

Type of Day	Date	RMSE (W/m^2)
Sunny	06/01/2017	5.33
	08/01/2017	9.46
	07/04/2017	13.18
	09/04/2017	5.16
Partially cloudy	04/01/2017	15.39
	03/04/2017	34.82
	03/05/2017	41.25
	10/06/2017	36.95
Cloudy	09/02/2017	58.28
	02/03/2017	57.48
	06/04/2017	73.00
	02/06/2017	77.76

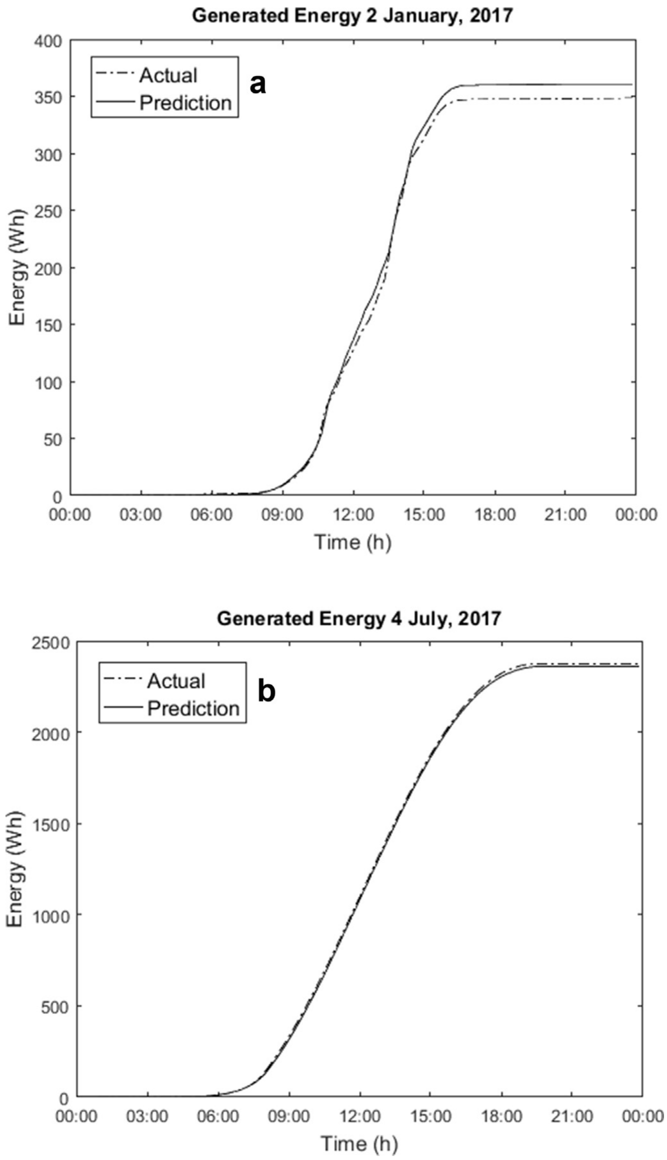

Fig. 11. a) Energy prediction for 2 January 2017 and, b) energy prediction for 4 July 2017.

and the lower the error.

Table 6 demonstrates that the predictions made by our tool are good enough. Even though the RMSE are high for some samples depending on the weather, the difference between the predicted

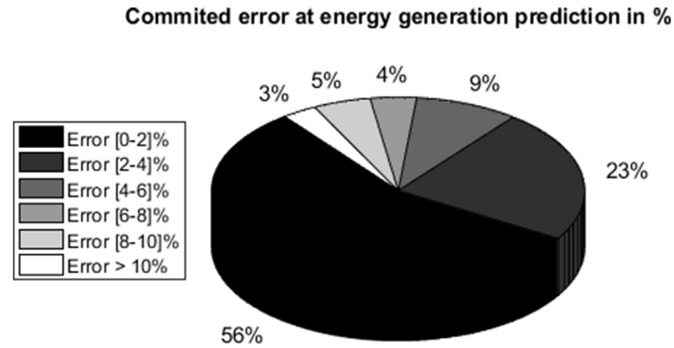

Fig. 12. Distribution of the error at analysed days.

and generated energy is more constant no matter what the weather is.

Finally, to ensure the accuracy of our tool, the days between January and August 2017 were analysed. Fig. 12 shows the degree of error between the real and the predicted energy as a percentage. The results demonstrate that in 56% of the samples, the tool has been able to predict the generated energy with an error rate that is under 2%, and the mean value of the error for the analysed period of time is 2.81%. However, the higher the ups and downs in solar irradiation, the less accurate the prediction. This can be observed by comparing the results in Figs. 8a and 9a.

5. Conclusions

This paper presents a tool which is able to predict solar energy in the short term, 10 min. It is based on an ANN which has been trained with a database from Euskalmet. The results show that the accuracy given by the final ANN for the solar irradiation parameter is sufficiently high that it can be used to help microgrids with their instantaneous control, due to the fact that the uncertainty of solar irradiation is reduced. For instance, the mean value of the error in % for the first eight months of 2017 was 2.81. Furthermore, Figs. 7a, 8a and 9a demonstrated that the tool's precision is sufficiently high owing to the fact that the trend of the prediction line follows the actual line at every moment.

Since microgrids are connected to the internet, they are able to pick up data from the forecasts put out by meteorological agencies and then use this data and the developed tool to predict the future energy provided by the PV generators. Therefore, microgrid controllers can weigh the results obtained from the ANN. For example, if we have a day that is sunny or only has a few clouds, the predictions obtained by the tool will be reliable and the control of the microgrid will not switch on additional subsystems to ensure

Table 6
Comparison of RMSE and error (%).

Type of Day	Date	RMSE (W/m^2)	Predicted Energy (Wh)	Generated Energy (Wh)	Error (%)
Sunny	06/01/2017	5.33	816.15	815.91	0.03
	08/01/2017	9.46	769.83	767.01	0.37
	07/04/2017	13.18	2021.21	2035.95	0.72
	09/04/2017	5.16	2045.75	2057.44	0.57
Partially cloudy	04/01/2017	15.39	693.52	686.89	0.96
	03/04/2017	34.82	2015.47	2025.31	0.49
	03/05/2017	41.25	2054.61	2098.97	2.11
	10/06/2017	36.95	2231.09	2257.06	1.15
Cloudy	09/02/2017	58.28	452.55	461.07	1.84
	02/03/2017	57.48	1268.25	1289.48	1.64
	06/04/2017	73.00	1491.63	1557.44	4.22
	02/06/2017	77.76	1704.87	1666.89	2.27

control. However, if the day is cloudy day and there are many ups and downs in terms of solar irradiation, the microgrid will have to switch on additional back-up subsystems, such as fuel converters, in order to be able to feed the power demand. Hence, the benefit that we obtain from the developed tool is an improvement on the efficiency and reliability of the microgrid control, as the back-up systems will be switched on only when they are really needed.

References

[1] S. A. Kalogirou, C. Neocleous, S. Paschiardis, C. Schizas, "Wind speed prediction using artificial neural networks". In: Proc. European Symp. On Intelligent Techniques ESIT'99, Crete (Greece).

[2] R. Huang, T. Huang, R. Gadh, Solar generation prediciton using the ARMA model in a laboratory-level micro-grid, in: IEEE SmartGridComm 2012 Symposium - Support for Storage, Renewable Sources, and MicroGrid, Tainan, 2012.

[3] C. Nayar, Innovative remote micro-grid systems, Int. J. Environ. Sustain. 1 (3) (2012) 53–65.

[4] A.M. Zein Alabedin, E.F. El-Saadany, M.M.A. Salama, Generation scheduling in microgrids under uncertainties in power generation, in: IEEE Electrical Power and Energy Conference, London, 2012.

[5] C. Chen, S. Duan, T. Cai, B. Liu, G. Hu, Smart energy management system for optimal microgrid economic operation, IET Renew. Power Gener. 5 (3) (2010) 258–267.

[6] A.K. Mishra, L. Ramesh, Application of neural networks in wind power (generation) prediction, in: International Conference on Sustainable Power Generation and Supply, 2009.

[7] N. Sharma, P. Sharma, D. Irwin, P. Shenoy, Predicting solar generation from weather forecasts using machine learning, in: IEEE International Conference of Smart Grid Communications, Brussels, 2011.

[8] V. Lara-Fanego, J.A. Ruiz-Arias, D. Pozo-Vazquez, F.J. Santos-Alamillos, J. Tovar Pescador, Evaluation of the WRF model solar irradiance forecasts in Andalusia (Southern Spain), Sol. Energy 86 (2012) 2200–2217.

[9] L. Hernández, C. Baladrón, J.M. Aguiar, L. Calavia, B. Carro, A. Sánchez-Esguevillas, P. García, J. Lloret, Experimental analysis of the input variables' relevance to forecast next Day's aggregated electric demand using neural networks, Energies 6 (6) (2013) 2927–2948.

[10] L. Hernández, C. Baladrón, J.M. Aguiar, L. Calavia, B. Carro, A. Sánchez-Esguevillas, J. Sanjuán, A. González, J. Lloret, Improved short-term load forecasting based on two-stage predictions with artificial neural networks in a microgrid environment, Energies 6 (9) (2013) 4489–4507.

[11] M. Rana, I. Koprinska, Forecasting electricity load with advanced wavelet neural networks, Neurocomputing 182 (2016) 118–132.

[12] P. Zhao, J. Wang, J. Xia, Y. Dai, Y. Sheng, J. Yue, Performance evaluation and accuracy enhancement of a day-ahead wind power forecasting system in China, Renew. Energy 43 (2012) 234–241.

[13] M. Paulescu, E. Paulescu, P. Gravila, V. Badescu, Weather Modeling and Forecasting of PV Systems Operation, Springer, 2013, pp. 17–42.

[14] A. More, M.C. Deo, Forecasting wind with neural networks, Mar. Struct. 16 (2003) 35–49.

[15] M. Hable, C. Meisenbach, G. Winkler, Economically optimised power dispatch in local systems using evolutionary algorithms and dynamic programming, in: Power System Management and Control, 2002.

[16] A. Mellit, A. Massi, A 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at Trieste, Italy, Sol. Energy 84 (2010) 807–821.

[17] S.M. Al-Alawi, H.A. Al-Hinai, An ANN-based approach for predicting global radiation in locations with No direct measurement instrumentation, Renew. Energy 14 (1–4) (1998) 199–204.

[18] S. Haykin, Neural Networks and Learning Machines, third ed., Upper Saddle River: Pearson Education, Inc, 2009.

[19] S.A. Kalogirou, Artificial neural networks in renewable energy systems applications: a review, Renew. Sustain. Energy Rev. 5 (2001) 373–401.

[20] J. E. Dayhoff and J. M. DeLeo, "Artificial neural networks", in Conference on Prognostic Factors and Staging in Cancer Management: Contributions of Artificial Neural Networks and Other Statistical Models.

[21] H. Demuth, M. Beale, Neural Network Toolbox User's Guide, The MathWorks, Inc, 2002.

[22] L.B. Mohammed, M.A. Hamdan, E.A. Abdelhafez, W. Shaheen, Hourly solar radiation prediction based on nonlinear autoregressive exogenous (Narx) neural network, Jordan J. Mech. Ind. Eng. 7 (1) (2013) 11–18.

[23] N. Manikandan, S. Subha, Software design challenges in time series prediction systems using parallel implementation of artificial neural networks, Sci. World J. 2016 (2016) 10.

[24] Y. Kemmoku, S. Orita, S. Nakagawa, T. Sakakibara, Daily insolation forecasting using a multi-stage neural network, Sol. Energy 66 (3) (1999) 193–199.

[25] D.P. Larson, L. Nonnenmacher, C.F. Coimbra, Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest, Renew. Energy 91 (2016) 11–20.

[26] F.L. Coolidge, Chapter 6: an introduction to correlation and regression, in: Statistics: a Gentle Introduction, third ed., SAGE Publications, Thousand Oaks, 2013, pp. 211–267.

[27] I. Alados, I. Foyo-Morenoa, F. Olmoa, L. Alados-Arboledas, Relationship between net radiation and solar radiation for semi-arid shrub-land, Agric. For. Meteorol. 116 (2003) 221–227.