



Extracting Trends Ensembles in Solar Irradiance for Green Energy Generation Using Neuro-evolution

Mehreen Rehman, Jawad Ali, Gul Muhammad Khan, Sahibzada Ali Mahmud

► To cite this version:

Mehreen Rehman, Jawad Ali, Gul Muhammad Khan, Sahibzada Ali Mahmud. Extracting Trends Ensembles in Solar Irradiance for Green Energy Generation Using Neuro-evolution. 10th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Sep 2014, Rhodes, Greece. pp.456-465, 10.1007/978-3-662-44654-6_45 . hal-01391347

HAL Id: hal-01391347

<https://inria.hal.science/hal-01391347>

Submitted on 3 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Extracting Trends Ensembles in Solar Irradiance for Green Energy Generation using Neuro-Evolution

Mehreen Rehman¹, Jawad Ali², Gul Muhammad Khan³ and Sahibzada Ali
Mahmud⁴

^{1,2,3,4}Department of Electrical Engineering,
University of Engineering and Technology Peshawar, Pakistan 25000.

¹ mehreen@nwfpuet.edu.pk

² jawad.ali@nwfpuet.edu.pk

³ gk502@nwfpuet.edu.pk

⁴ sahibzada.mahmud@nwfpuet.edu.pk

Abstract. *Globally, there are variations in climate, there is fossil fuel depletion, rising fossil fuel prices, increasing concern regarding energy security, and awareness about the environmental impacts of burning fossil fuels. These factors lead to a growing interest around the world in green and renewable energy resources, solar energy being a common one. A Neuro-evolutionary approach is explored to extract the trend ensembles in the solar irradiance patterns for renewable electric power generation, using the data taken from stations in Al-Ahsa, Kingdom of Saudi Arabia. The algorithm, based on Cartesian Genetic Programming Evolved Artificial Neural Network (CGPANN) was developed and trained for hourly and 24-hourly prediction, using the solar irradiance value as the input parameter. It was tested to predict solar irradiance on hourly, daily, and weekly basis. The proposed technique is **95.48%** accurate in solar irradiance prediction.*

Keywords: Solar irradiance Forecasting, Neural Networks, Cartesian Genetic Programming (CGP), Neuro-evolution, Time Series Prediction

1 Introduction

Most places on the earth's surface receive clean and abundant solar energy, free of cost, throughout the year. Owing to the rising fossil fuel costs and the degradation of atmosphere by these fossil fuels, there is a dire need for economical and effective harnessing and utilization of solar energy. Because the world has to pay the cost of losing its energy resources at an alarmingly fast pace, there is an increasing reliance of power generation on renewables, such as solar energy. Worldwide efforts are being taken to increase the capacity of renewable power production. However, the intermittent nature of renewables makes it a challenging task to integrate them into any system. Accurate solar irradiance is therefore

essential in improving the efficiencies of solar and wind energy based applications, obtaining optimized power production [1], overcoming the challenges of integrating them in the power grid and for the efficient management and operation of solar thermal energy plants. Solar radiation data gives us information about the amount of energy that hits the earth's surface during a particular time frame and is required for research on effective solar energy utilization. Because this data is not readily available, and because the intensity and availability of solar radiation is dependent upon several other environmental factors, alternate ways of generating the data must be developed [2].

Artificial Neural Networks is an emerging technology for handling complex, practical problems, such as forecasting. Neural networks are capable of making computationally efficient, faster and more practical predictions than any of the conventional methods.

Our work mainly focuses on introducing a new and efficient solution based on the Neuro-evolutionary technique termed as the Cartesian Genetic Programming Evolved Artificial Neural Network (CGPANN) [3] to extract the trend ensembles in the solar irradiance patterns for renewable power generation. Neuro-Evolution is the phenomenon that involves the artificial evolution of the entities of an Artificial Neural Network (ANN). In CGPANN, Neuro-Evolution exploits the powerful structural properties of Cartesian Genetic Programming [4] and the functional properties of an ANN. This leads to the evolution of all the network parameters; node weights, system inputs, node inputs, node functions, network topology as well as the system outputs [3]. The research solution discussed here is specific only to CGPANN. The proposed technique uses past solar irradiance data to estimate the hourly, 12-, 24-, 48- and 168-hourly solar radiation patterns.

2 Literature Review

2.1 Computational Intelligence Techniques for Solar Irradiance Forecasting

Estimation methods used in the literature for forecasting solar radiation include stochastic, analytic [5], and artificial neural networks [6]. The analysis by Mubiru and Banda [7] shows that owing to the nonlinear, non-stationary nature of solar radiation; ANNs are thought to be superior to empirical methods in estimating the global solar radiation.

Autoregressive (AR) and autoregressive moving average (ARMA) are models frequently used for the estimation of solar irradiance [8, 9]. Angstrom [10] was the pioneer to apply Regression models based on sunshine duration. Ahhi et al. [11] have used various input parameters of an ANN to predict solar radiation.

Kalogirou Model applies the standard back propagation learning algorithm. The model uses the hourly records for one year to predict hourly solar radiation. Other ANN models are; The Mohandes Model [6]; The Kemmoku Model [1]; The Reddy Model [13]; The Sozen Models [14] that has explored Scaled conjugate gradient (SCG), Pola-Ribiere conjugate gradient (CGP), and Levenberg-Marquardt

(LM) learning algorithms and logistic sigmoid transfer function; The Cao Model [15] that establishes a recurrent BP network for forecasting of solar irradiance; The Soares Model [16] that develops a neural network exploring feature determination and pattern selection techniques. Linares-Rodriguez et. al. [17] explores an artificial neural network ensemble model for estimating global solar radiation from Meteosat satellite images. This model uses clear-sky estimates and satellite images as input variables. Tymvios et al., Angela et. Al [20] have introduced a single-parameter model for solar irradiance prediction. Lorenz et al., [21] use numerical weather predictions (NWP) as input parameters to NNs to estimate global irradiance. Hocaoglu et al., [22] study and compare feed-forward NNs with seasonal AR models. Cao and Lin, [23] combined NNs with wavelets to predict next day hourly values of global irradiance, incorporating various meteorological parameters as inputs to the models. Jain and Goel, [24] developed a multilayer feed forward neural network model for solar radiation forecasting in India. For the input to the network, the geographical, solar and meteorological parameters were taken from different areas of India. Azadeh et. al. [25] proposed an ANN model based on multilayer perceptron approach, incorporating all climatological and meteorological factors as input variables for prediction of solar global radiation. The results of the integrated ANN-MLP model have been compared to those of angstroms model [10]. There is a notable decrease in the MAPE value from 14.78% by the Angstrom model to 7.5% by the integrated ANN model.

2.2 Cartesian Genetic Programming evolved Artificial Neural Network (CGPANN)

CGPANN is a Neuro-Evolutionary technique based on Cartesian Genetic Programming where the nodes are replaced by artificial neurons, the basic elements of CGPANN. A CGPANN node or neuron is composed of inputs, weighted connections, non-linear activation functions and outputs. The genotype is generated by encoding all these attributes of the neural networks [3]. The parent genotype undergoes mutation, producing the offspring. The network and its parameters continue to evolve until the best possible solution is obtained [4]. In this paper, CGPANN is used as the principle estimator in the prediction system. For the production of the next generation population in CGPANN, (1+) evolutionary strategy is recommended. The 10 % mutation rate used implies that only 10% of the genes take part in mutation to give the offspring. During the process of evolution some neurons, whose outputs remain unconnected (Junk Nodes) may not participate at all in generating the system output whereas others (Active Nodes) may play an active role in producing the system output. The weights are generated randomly (ranging from -1 to 1). The output can either be the system inputs or the output of any of the nodes. The neurons in the network are not fully connected. The systems inputs are not connected with all of the neurons in the input layer. The fundamental entities of a typical node in a CGPANN can be observed in Figure 1. The selection of the inputs is from the input array I_p , such that

$$I_p = [i_{p1}, i_{p2}, i_{p3}, \dots, i_{pn}] \quad (1)$$

The weighing matrix [array of weights] W_g ,

$$W_g = w_{g1}, w_{g2}, w_{g3}, \dots, w_{gn} \quad (2)$$

for a particular genotype g , consists of random weights ranging from +1 and -1 . The output of a junction in ANN that takes x_i as junction input, where the input is multiplied by a randomly allotted weight, w_{gi} can be represented by

$$O_{n'} = \sum_{i=1}^N x_i w_{gi} \quad (3)$$

If the inputs to a node are N , output $O_{nj'}$ is

$$O_{nj} = \xi^j(O_{n'}) = \xi^j\left(\sum_{i=1}^N x_i w_{gi}\right) \quad (4)$$

here, ξ is the node function. Also, j is defined by

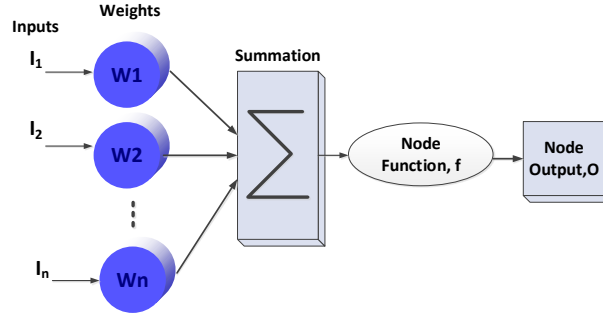


Fig. 1: The Nth node of CGPANN

$$\{j|j \in N \wedge N_T \geq j \geq 1\} \quad (5)$$

If I_g is the input set to G_k , where each input i has unique value:

$$i \in R \wedge 1 \geq i \geq 0 \quad (6)$$

then, network G_k consists of randomly chosen inputs I_g , node outputs, O_{nj} , for a single output O_p such that

$$O_p = \frac{1}{n} \sum_{i=1}^N O_i \quad (7)$$

where

$$O_i = \xi \left(\sum (O_{nj} W_{gj} + O_{nj-1} W_{gj-1} + \dots + O_{n1} W_{g1} + I W_{gk}) \right) \quad (8)$$

Here, W_{gj} , W_{gj-1} , ... W_{g1} are Random subsets of W such that W_{gk} is a subset W_{gj} and

$$W_{gj} = \{w_{gk} | w_{gk} \in R \wedge +1 \geq w_k \geq -1\} \quad (9)$$

now

$$G_k = \{I, O_{nj}, O_{nj-1}, \dots, O_{n1}, O_p\} \quad (10)$$

if we have two successive genotypes G_k and G_l . G_l is the product of G_k by mutating $\mu\%$ characteristics of the genotypes. Let the total entries in G_k be N_{cfw} then

$$N'_{cfw} = \mu \times N_{cfw} \quad (11)$$

N_{cfw} represents genotype entries that are to be mutated to get G_l . The set ζ has ω as a unique entry such that

$$\{\omega | \omega \in \zeta \forall (W_g, I, O_{nj})\} \quad (12)$$

Each value in ω is defined Pseudo-randomly. N_{cfw} values are provided by the available N_{cfw} entities using the equivalent expression

$$\left\{ \omega_i | \omega_i \in \{1, 2, 3, \dots, N_{cfw}\} \wedge \omega_i \subset \{1, 2, 3, \dots, N_{cfw}\} \right\}, i = 1, 2, 3, \dots, N'_{cfw} \quad (13)$$

Where ω_i can its value from $\{N_{cfw}, \dots, 3, 2, 1\}$ so each entry in ζ is

$$\gamma(i) = \left\{ \omega_i | \omega_i \in \{1, 2, 3, \dots, N_{cfw}\} \wedge \{1 \leq \omega_i \leq N_{cfw}\} \right\} \quad (14)$$

The network takes inputs from array I_p . The network inputs are fed to the primary layer, the output of which is fed into secondary processing nodes, known as intermediate layer nodes. Activation function ξ , that is employed for this particular CGPANN is Log-sigmoid function, given by Eq. 16.

$$\xi(x) = \frac{1}{1 + e^{-x}} \quad (15)$$

Here x is input to the activation function $\xi(x)$. x is defined by the Eq. 17.

$$O(c, g, j) = \sum_{i=1}^N w_g(c, i) \cdot I(g, c, i) \quad (16)$$

Whereby N defines the quantity of inputs to the given node j for $n = 1, 2, 3 \dots N$ numbered inputs. K are the network inputs in set $k = 1, 2, 3 \dots N \dots K$. g represents a particular genotype in a population and c is the node taking part in

the operation. $w(i, c)$ is the randomly assigned weight to each input and $I(c, g, i)$ is defined as a randomly selected input to the node j .

$$I(g, c, i) = PRG([I(g, c, 1), I(g, c, 2), \dots, I(g, c, N) \dots \dots, I(g, c, k)] : [O(g, c, 1), O(g, c, 2), \dots O(g, c, j1)]) \quad (17)$$

A typical CGPANN genotype and phenotype is shown in Fig. 2. The figure shows a single row, four nodes, with three inputs per nodes, five inputs and outputs network.

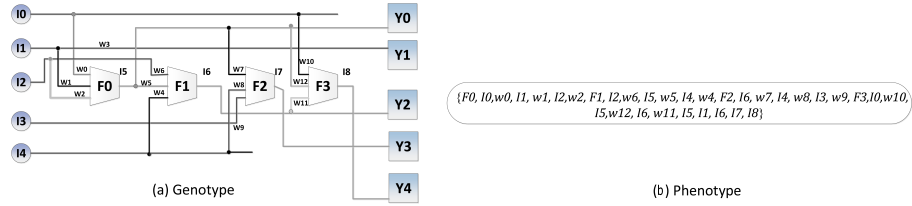


Fig. 2: A typical CGPANN genotype and phenotype

3 Application of CGPANN to solar irradiance forecasting

3.1 Experimental Setup

The solar irradiance data of 1997, collected from Saudi Arabian stations in Al-Ahsa [26], were used to train the network for hourly and 24-hourly prediction and subsequently testing was done on the data acquired from 1997 as well as 1998, for hourly, 12-, 24-, 48- hourly and 7 days prediction capability of the model. The inputs to the network are the solar irradiance values, where irradiance, given in W/m^2 , is the intensity of electromagnetic radiation incident per unit area. Al-Ahsa lies at an altitude of 178 m, $25.3^\circ N$ latitude and $49.48^\circ N$ longitude. The average temperature is 304 K, relative humidity 23% and the daily average solar global radiation is $21.6 MJ/m^2$. Although the network has been tested and trained using the data from 1997, the datasets for both belong to different months which mean that the data used for training is not being used for testing. The Mean Absolute Percentage Error (MAPE) value is used as the performance criterion to compare the offspring to the fittest network (the parent) and decide which of these are fit enough to be promoted to the next generation. The parent network further produces nine more networks by mutation. The process continues until the desired fitness is achieved or the maximum number of generations is completed. For training the network, all experiments are run for one million generations. Because a mutation rate of 10% is thought to produce more satisfactory results in comparison to other rates, it has been

used in our experiment (Kadilar & Alada, 2009; Huang et al., 2006; Chen et al., 2008).

MAPE is given by:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left(\left| \frac{L_F - L_A}{L_A} \right| \right) \times 100\% \quad (18)$$

$$Fitness = 100\% - MAPE \quad (19)$$

Where L_F is the predicted value, L_A is the actual value and n is the number of days.

3.2 Results and Analysis

The proposed CGPANN model has been evaluated using 24 hours input network for the hourly, daily and weekly prediction of solar irradiance. The network takes 24 entries for the prediction of the next hour, next 12 hours, 24 hours, 2 days and 7 days. While the number of nodes on the network is varied between 50 and 500 with a step size of 50, training and testing experiments are repeatedly performed. Subsequently, a phenotype was being translated by the genotype that showed the best fitness level and then testing was carried on hourly, daily, and weekly basis. All the network nodes however may not be utilized by the final phenotype. Usually 5 to 10 percent of nodes participate in the production of the phenotype.

During testing, an independent dataset was given to the network for its validation. Table 1 contains the MAPE values calculated using 24 hours input network for the prediction of the next one hour, 12hrs, and 1 day. The network used here has been trained for hourly prediction. The values change with the changing number of nodes. MAPE values increase as the network is being used for long-term forecasting. The best result has been for the hourly prediction for 500 nodes i.e. a MAPE of only 4.52%.

Table 2 shows the testing results of predictions on 12-hourly, daily, and weekly basis, using the network model that has been trained for 24-hourly prediction. A lowest MAPE value of 5.36 has been achieved for the 12-hourly prediction. Hence, with the increase in time span, the forecasting of the model is affected owing to environmental, technical and climatic issues. The difference in values of Tables 1 and 2 explain the fact that no matter what the number of nodes, the CGPANN model exhibits better results when forecasting over a shorter period. However, the model that is trained to predict a single instance is also capable of predicting up to 24 hours without much error (see table 1). And the model that is trained to predict 24 hours can predict up to 7 days (168 hours) in advance with more than 90% accuracy (see table 2), indicating the robustness of the model. Another fact revealed is that the duration for which the network is trained is directly proportional to the efficiency of the model. This is indicated by the lower MAPE values in Table 2 corresponding to 24-hourly prediction training compared to those in Table 1 that correspond to hourly prediction training.

Figure 3 represents a graph that displays the deviation between the actual and the estimated data in terms of the normalized values of irradiance. It is obvious that the estimated data is following the frequent fluctuations in the actual data throughout the time frame i.e. 300 hours. The graph indicates only a minor difference between the two sets of values, confirming the accuracy and robustness of the CGPANN model.

Table 1: Testing Results for Different Number of Nodes using 24 Hours Input Network for Prediction of the Next One Hour, 12 hours, and 1 Day, for Hourly Prediction Training

Prediction	Year	50	100	150	200	250	300	350	400	450	500
Hourly	1998	6.65	4.83	4.61	4.79	4.65	4.77	5.05	5.15	4.94	4.52
12-hourly	1997	9.88	7.48	7.03	7.06	6.98	6.64	7.44	7.57	7.22	6.78
	1998	11.01	8.37	7.97	7.96	7.87	7.55	8.36	8.49	8.14	7.68
24-hourly	1997	10.48	8.42	7.99	8.20	8.09	7.55	8.73	8.94	8.31	7.83
	1998	11.72	9.50	9.10	9.27	9.13	8.62	9.83	10.03	9.41	8.88

Table 2: Testing Results for various Number of Nodes for 24 Hours Input Network and Prediction of the Next 12 hours, 1 day, 2 Days, and 1 Week, for 24-hourly Prediction Training

Prediction	Year	50	100	150	200	50	300	350	400	450	500
12-hourly	1997	6.78	5.41	5.37	5.36	5.39	5.42	5.41	5.38	5.42	5.37
	1998	8.15	6.67	6.64	6.63	6.65	6.67	6.67	6.65	6.68	6.64
24-hourly	1998	8.08	6.68	6.65	6.64	6.66	6.70	6.68	6.66	6.69	6.65
48-hourly	1997	7.80	5.98	5.92	5.90	5.95	6.03	5.99	5.93	6.00	5.92
	1998	9.41	7.45	7.41	7.41	7.43	7.50	7.45	7.42	7.46	7.42
7 days	1997	8.30	7.43	7.28	7.17	7.39	7.52	7.56	7.26	7.59	7.31
	1998	12.41	8.52	8.37	8.29	8.48	8.66	8.62	8.36	8.66	8.40

4 Conclusion

The use of renewable energy necessitates exact estimation of renewable potential throughout the globe. Our research is an effort to introduce a technique for solar irradiance forecasting based on the Neuro-evolutionary technique called Cartesian Genetic Programming evolved Artificial Neural Network (CGPANN). The use of this technique could be very useful for improving efficiencies of solar based applications, the integration of renewable energy resources or the management of solar energy systems. The experiments display promising results, yielding an accuracy of 95.48 % and MAPE of 4.52% for hourly prediction. As the only

input to the learning model is the solar radiation patterns, in the future, other parameters such as environmental conditions can be incorporated as inputs to produce even more accurate results.

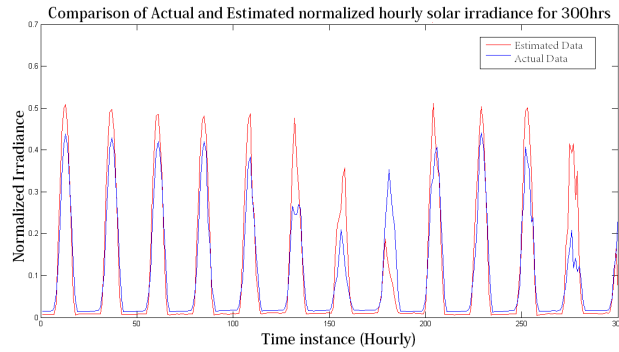


Fig. 3: The actual and the estimated normalized solar irradiance for 300 hours

References .

- [1]Koeppel, G., Korpas, M., (2006). Using storage devices for compensating uncertainties caused by non-dispatchable generators. 2006 International Conference on Probabilistic Methods Applied to Power Systems, pp. 18.
- [2]Dorvlo, A. S., Jervase, J. A., & Al-Lawati, A. (2002). Solar radiation estimation using artificial neural networks. *Applied Energy*, 71(4), 307-319.
- [3]Khan, M. M., Khan, G. M., & Miller, J. F. (2010). Efficient representation of recurrent neural networks for markovian/non-markovian non-linear control problems. In *Intelligent Systems Design and Applications (ISDA)*, 2010 10th International Conference on (pp. 615-620). IEEE.
- [4]Miller, J. F., (2011). *Cartesian genetic programming*. Springer.
- [5]Dagestad, K., (2005). Estimating global radiation at ground level from satellite images, Ph.D. thesis, Department of Meteorology, University of Bergen, Bergen, Norway.
- [6] Mohandes M, Balghonaim A, Kassa M, Rehman S, Halawani TO. (2000) Use of radial basis functions for estimation monthly mean daily solar radiation. *Solar Energy* 2000;68(2):1618.
- [7]Mubiru J., and Banda, E. J. K. B., (2008), Estimation of monthly average daily global solar irradiation using artificial neural networks, *Solar Energy*, vol. 82, no. 2, pp. 181187.
- [8]Reikard, G., (2009) Predicting solar radiation at high resolutions: a comparison of time series forecasts, *Solar Energy*, vol. 83, no. 3, pp. 342-349.
- [9]Torres, J. L., Garcia, A., De Blas, M., and De Francisco A., (2005) Forecast of hourly average wind speed with ARMA models in Navarre (Spain), *Solar Energy*, vol. 79, pp. 65-77.

- [10] Angstrom A. (1956). On the computation of global solar radiation from records of sunshine. *Arkiv Geophysik*, 3(23):551
- [11] Ahhi, M. Shamisi, and M. Jama, (2010). Prediction of monthly average daily global radiation in Al Ain city- UAE using artificial neural networks, in *Proc. 4th RES 10*, pp. 109-113.
- [12] Reddy, K. E., & Ranjan, M. (2003). Solar resource estimation using artificial neural networks and comparison with other correlation models. *Energy Conversion and Management*, 44(15), 2519-2530.
- [13] Sozen, A, E. Arcaklioglu, M. Ozalp and E. G. Kanit (2005), Solar-energy potential in Turkey, *Applied Energy*, 80(4), 367381
- [14] Cao S., and Cao, J., (2005). Forecast of solar irradiance using recurrent neural networks combined with wave analysis, *Applied Thermal Engineering*, 25 (2-3), 161-172.
- [15] Soares, J., Oliveira, A. P., Boznar, M. Z., Mlakar, P., Escobedo JF and Machado, AJ., (2004) Model-ling hourly diffuse solar radiation in the city of Sao Paulo using artificial neural network technique, *Applied Energy*, 79 (2), 201-214.
- [16] Linares-Rodriguez, A., Ruiz-Arias, J. A., Pozo-Vazquez, D., & Tovar-Pescador, J. (2013). An artificial neural network ensemble model for estimating global solar radiation from Meteosat satellite images. *Energy*, 61, 636-645.
- [17] Bacher, P., Madsen, H., & Nielsen, H. A. (2009). Online short-term solar power forecasting. *Solar Energy*, 83(10), 1772-1783.
- [18] Angela, K., Ssenyonga Taddeo, and Mubiru James (2011). Predicting Global Solar Radiation Using an Artificial Neural Network Single-Parameter Model. *Advances in Artificial Neural Systems*.
- [19] Lorenz, E., Heinemann, D., Wickramaratne, H., Beyer, H., Bofinger, S., (2007). Forecast of ensemble power production by grid-connected PV systems. In: *Proceedings of the 20th European PV Conference*, Milano.
- [20] Hocaoglu, F.O., Gerek, O.N., Kurban, M., (2008). Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural networks. *Solar Energy* 82 (8), 714726.
- [21] Cao, J., Lin, X., (2008). Study of hourly and daily solar irradiation forecast using diagonal recurrent wavelet neural networks. *Energy Conversion and Management* 49 (6), 13961406.
- [22] Jain, R., & Goel, B. (2013). Prediction of Global Solar Radiation Using Artificial Neural Network. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, Vol.2 Issue 10.
- [23] Azadeh, A., Maghsoudi, A., & Sohrabkhani, S. (2009). An integrated artificial neural networks approach for predicting global radiation. *Energy Conversion and Management*, 50(6), 1497-1505.
- [24] Yang, K., Koike, T., & Ye, B. (2006). Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets. *Agricultural and Forest Meteorology*, 137(1), 43-55.