Revue des Energies Renouvelables Vol. 13 N°1 (2010) 25 – 47

EPNN-based prediction of meteorological data for renewable energy systems

A. Mellit*[,1], M. Drif†[,2] and A. Malek[2]

[1] Department of Electronics, Faculty of Sciences and Technology,
LAMEL, Mohamed Benyahia University, Ouled-Aissa, P.O. Box 98, Jijel, Algeria
[2] Centre of Development of Renewable Energy (CDER),
P.O. Box 62, Bouzareah, Algiers, Algeria

(reçu le 20 Décembre 2009 – accepté le 25 Mars 2010)

Abstract - *In this paper, an application of an Evolving Polynomial Neural Network (EPNN) for prediction of meteorological time series (global solar irradiation, air temperature, relative humidity, and wind speed) is described. Prediction of such data plays a very important role in design of the renewable energy systems. The problem of time series prediction is formulated as a system identification problem, where the input of the system is the past values (y (t - 1), y (t - 2), y (t - 3), …) of a time series and its desired output (y (t), y (t + 1), y (t + 2), …) are the future of a time series. In this study, a dataset of meteorological time series for five years collected in Algiers (Algeria) by the National Office of Meteorology has been used. The obtained results showed a good agreement between both series, measured and predicted. The correlation coefficient (r) is arranged between 0.9821 and 0.9923, the mean relative error over the whole data set is not exceed 15.4 %. The proposed model provides more accurate results than other ANN's architecture, wavenet (wavelet-network) and Adaptive Neuro-Fuzzy Inference Scheme (ANFIS). In order to show the effectiveness of the proposed predictor, the predicted data have been used for sizing, and prediction of the output energy of photovoltaic systems.*

Résumé - *Dans cet article, une application sur un réseau neural polynomial en constante évolution (EPNN) pour la prédiction de séries temporelles météorologiques (rayonnement solaire global, température de l'air, humidité relative et vitesse du vent) est décrite. La prévision de ces données joue un rôle important dans la conception des systèmes d'énergie renouvelable. Le problème de la prédiction de séries temporelles est formulé comme un problème d'identification du système, où l'entrée du système est les valeurs du passé (y (t - 1), y (t - 2), y (t - 3), ...) des séries chronologiques et sa sortie désirée (y (t), y (t + 1), y (t + 2), ...) sont l'avenir d'une série chronologique. Dans cette étude, un ensemble de données de séries chronologiques météorologiques recueillies pendant cinq ans à Alger (Algérie) par l'Office National de Météorologie a été utilisé. Les résultats obtenus montrent un bon accord entre les deux séries, mesurées et prédites. Le coefficient de corrélation (r) est disposé entre 0,9821 et 0,9923, et l'erreur relative moyenne sur l'ensemble des données n'est pas excéder 15,4 %. Le modèle proposé fournit des résultats plus précis que dans d'autres techniques 'Ann's', wavenet (réseau d'ondelettes) et Adaptive Neuro-Fuzzy Inference Scheme (ANFIS). Afin de démontrer l'efficacité de l'indicateur proposé, les données prédites ont été utilisées pour le calibrage, et la prédiction de l'énergie de sortie de systèmes photovoltaïques.*

Keywords: Meteorological data - Prediction - Renewable energy systems - Artificial neural networks - Polynomial neural network - Genetic algorithm.

---

* Associate member at the International Centre for Theoretical Physics (ICTP) , Trieste, Italy www.ictp.it
† Corresponding author Dr. M. Drif (m_drif@hotmail.com)

1. INTRODUCTION

Meteorological data such as solar irradiation, air temperature, relative humidity and wind, are accepted as dependable and widely variable renewable energy sources. It is therefore required to be able to formulate forecasting and estimation models of these meteorological data [1].

Real world problems are described by non-linear and chaotic processes which make them hard to model and predict. However, prediction of meteorological time-series such as global solar irradiation, ambient temperature, wind speed, and relative humidity play very important role in:

Design of the renewable and solar energy systems;
Weather control;
Climate impact assessments of agricultural, and water system management.

For this reason, numerous authors have developed models for modelling and predicting of meteorological time-series based on statistical processes, such as Autoregressive (AR), moving-average (MA), autoregressive moving-average (ARMA) model, autoregressive integrated moving-average (ARIMA) model and Markov chain [2-12].

An obvious problem is that these processors are not able to cope with certain non-stationary signals, and signals whose mathematical model is not linear. On the other hand, Artificial Neural Networks (ANNs) are powerful when applied to problems whose solutions require knowledge which is difficult to specify [1, 13-15], but for which there is an abundance of examples. As time series prediction is conventionally performed entirely by inference of future behaviour from examples of past behaviour, it is a suitable application for a neural network predictor.

The neural network approach to time series prediction is non-parametric in the sense that it does not need to know any information regarding the process that generates the signal. For instance, the order and parameters of an AR or ARMA process are not needed in order to carry out the prediction. This task is carried out by a process of learning from presented examples. The main reasons for using neural networks for prediction rather than classical time series analysis are [13]:

They are computationally at least as fast, if not faster, than most available statistical techniques;
They are self-monitoring (i.e. they learn how to make accurate predictions);
They are as accurate if not more accurate than most of the available statistical techniques;
They provide iterative forecasts;
They are able to cope with nonlinearity and non-stationary of input processes;
They offer both parametric and nonparametric prediction.

Different models based on the neural networks were introduced in literature for predicting of meteorological data [16-18]. An artificial neural network for predicting of hourly mean values of ambient temperature 24 hours in advance is utilized by [19]. Full year hourly values of ambient temperature are used to train a neural network model for the coastal location of Jeddah, Saudi Arabia.

The mean percent deviation between the predicted and measured values is found to be 3.16, 4.17 and 2.83 for three different years. A fuzzy model of solar irradiance on inclined surfaces has been developed by [20]. The fuzzy model includes concepts from earlier models, though unlike these, it considers non-disjunctive sky categories. The proposed model offers performance similar to that of the models with the best results in the comparative analysis of the literature, such as the Perez model [21].

An ensemble of ANNs and learning paradigms for weather forecasting in southern Saskatchewan (Canada) has been developed in [22]. According to the authors the proposed neural networks are not only able to learn better but also to generalize better than conventional multi-layered perceptron network (MLPN), Elman recurrent neural network (ERNN), radial basis function network (RBFN), and Hopfield model (HFM). The modelling results indicate that reasonable prediction accuracy was achieved for most of the models. A simplified hybrid model for generating sequences of total daily solar radiation which combine a neural network and Markov chain, is proposed by [23].

This model is called ANN-MTM (Markov Transition Matrix). The inputs of the proposed model are the geographical coordinates while the outputs are the daily total solar radiation. It can be used for generating sequences of solar radiation at long term and it was applied for Algeria. The unknown validation data set produced very accurate prediction with an RMSE error not exceeding 8 % between the measured and predicted data. A correlation coefficient ranging from 90 % and 92 % has been obtained.

A suitable wavenet (wavelet-network) model for forecasting of daily global solar radiation data is developed by [24]. The obtained result indicates that the proposed wavenet predictor present more perform results compared with the conventional approaches AR, ARMA, Markov and neural networks. A soft computing model based on a radial basis function network (RBFN) for 24-hour weather forecasting of southern Saskatchewan, Canada is presented by [25]. The model is trained and tested using hourly weather data of temperature, wind speed and relative humidity for 2001. The results indicate that the Radial Basis Function Network (RBFN) produces the most accurate forecasts compared to the Multi-layer Perceptron (MLP), Elman Recurrent Neural Network (ERNN) and Hopfield Neural Network (HNN). An artificial neural network for the wind speed prediction of a target station using reference stations data is used by [26]. The maximum mean absolute percentage error was found to be 14.13 % for Antakya meteorological station and the best result was found to be 4.49 % for Mersin meteorological station.

A new model based on neuro-fuzzy for predicting the sequences of monthly clearness index and used it for generating solar radiation is proposed by [27], which has been used for the sizing of a Photovoltaic-system. The authors proposed a hybrid model for estimating sequences of daily clearness index by using an Adaptive Neuro-Fuzzy Inference Scheme (ANFIS) and Markov chain; the proposed model has been used for estimating the daily solar radiation. An application of sizing a Photovoltaic-system is presented based on the data generated by this model. The authors [28] had applied fuzzy set for generating daily solar radiation from air temperature, obtained results demonstrate the potential of modelling solar irradiation using fuzzy set, according to the authors, for improvement the precision, the proposed model could be further developed either by including additional relevant meteorological variable.

A dynamic Neuro-fuzzy and Kalman filter for predicting irradiance and ambient temperature is proposed and used by [29]. The results show that the obtained Normalized Mean Bias Error (NMBR) and Normalized Root Mean Square Error (NMSE) do not exceed 10%, which prove the efficiency of the adopted method when compared to those calculated for various standard models in literature. Recently, an extended review of the application of the artificial intelligence techniques for modelling and forecasting of solar irradiation found in [14]. In this review, several approaches have been compared and analyzed. The authors [30] had applied the MLP for creating the hourly global and diffuse solar radiation, obtained results indicates that a very good agreement, with a satisfactory outcome is verified between global and diffuse solar irradiation hourly data sets obtained by ANNs. A feed-forward neural network for estimating global solar radiation by using air temperature and relative humidity, is used by [31, 32] obtained results show good precision when they used as inputs both parameters (air temperature, relative humidity and sunshine duration), instead when used only air temperature. An ANN predictor for solar radiation from other meteorological data has been implemented on FPGA (Field Programmable Gate Array) [33].

The Genetics-Based Self-Organizing Network (GBSON) method was introduced to overcome the drawbacks of the original Group Method of Data Handling (GMDH) algorithms, since they use local search techniques to obtain an optimal solution [34]. This method has been used for solving real-world problems [35], it has been used for complex system identification, non linear system modelling, time series prediction etc.

In this paper, we investigate the suitability of the combined Group Method of Data Handling (GMDH) and Genetic Algorithm (GA) to predict the meteorological parameters. In the same context, several Artificial Neural Networks (ANN) architectures, wavenet, ANFIS have been investigated and compared with the PNN-GA technique. Therefore, the inputs of the network are the past values, while the outputs are the future values of a time series. An example of the prediction one future day $x(t)$ based on the past three days $\left(x(t-1), x(t-2), x(t-3)\right)$ is presented and discussed.

This paper is organized as follows: the next section provides the problem formulation of the present study. Section 3 deals with the presentation of polynomial neural network, genetic algorithm and the evolving polynomial neural network used in this study. The method implementation and results discussion are given in section 4. An application for sizing and prediction of the output energy from PV systems is presented and discussed in the final section.

2. PROBLEM FORMULATION AND MOTIVATION

Time series prediction involves the determination of an appropriate model, which can encapsulate the dynamics of the system, described by the sample data. The meteorological time series such as solar radiation, air temperature, relative humidity, wind speed, wind direction, pressure etc. relatively are neither completely random nor deterministic, so, the time series prediction is based on their past values, it is necessary to obtain a data record. When obtaining a data record, the objective is to have data that are maximally informative and an adequate number of records for prediction purposes. Hence, future values of a time series $x(t+T)$ can be predicted as a function of past values $x(t-T)$, $x(t-2T)$, $x(t-3T)$, …, $x(t-kT)$ so:

$$\tilde{x}(t+T) = f\left(x(t), x(t-T), x(t-2t), \ldots, x(t-kT)\right) \tag{1}$$

where T is the sampling time (time interval), The problem of time series prediction now becomes a problem of system identification, the unknown system to be identified is the function f with inputs the past values of the time series.

To solve this problem several ANN-predictor have been developed in literature. However, the genetic-based self-organizing network has been applied with success for time series prediction such as, sunspot series [34], explosive forming process [36], exchange rates, etc.

In our knowledge, the genetic-based self-organizing network has not been applied in meteorological time series prediction, so we try to investigate and to apply this kind of network (PNN-GA) in order to evaluate his reliability for predicting of meteorological time series, and compare it with some alternative ANNs-predictor.

3. GENETIC ALGORITHM, POLYNOMIAL NEURAL NETWORK AND GENETIC-BASED SELF-ORGANIZING NETWORK

3.1 Genetic algorithm

Genetic algorithms (GAs) are a kind of stochastic search or optimization algorithms, originally proposed by [37]. They have been invented based on natural genetics and evolution. The outline of a simple GA procedure is following way. Initially, the initial population of individuals which have a binary digit string as the 'chromosome' is generated at random. Each bit of chromosome is called 'gene'. The 'fitness', which is a measure of adaptation to environment, is calculated for each individual. Then, 'selection' operation passing individuals to next generation is performed based on fitness value, and then 'crossover' and 'mutation' are performed on the selected individuals to generate new population by transforming chromosomes into offspring's ones. This procedure is continued until the end condition is satisfied. This algorithm is conforming to the mechanism of evolution, in which the genetic information changes for every generation and the individuals which adapt to environment better survive preferentially. The properties of GAs are that GA is a stochastic parallel search based on the multi search points and GA requires only fitness value based on the objective functions. Particularly GA attracts attentions as a solver of multi-objective optimization problems due to parallel search. Genetic Algorithms begin with a population of chromosomes created randomly. Following, the initial population is evaluated. Three main operators: reproduction, crossover and mutation are used to evolve the initial population towards better solutions. The population is evaluated, and if the termination criteria are not met, the three main operators are applied again. One cycle of these operators and the evaluation procedure is known as a generation in GA terminology.

3.2 Polynomial neural network

The Polynomial Neural Network (PNN) algorithm [38] is based on the GMDH method and utilizes a class of polynomials such as linear, quadratic, and modified quadratic. The individual Preferred Node (PNs) are expressed as a second-order regression equation. In particular, when combining two inputs at each node as the generic structure we arrive at the following relationship:

$$y = A + B x_i + C x_j + D x_i^2 + E x_j^2 + F x_i x_j \tag{2}$$

In the above expression, A, B, C, D, E and F are parameters of the model, while y is the output of this model, x_i and x_j denote two inputs. The outputs obtained from each of these nodes are then combined to obtain a higher-degree polynomial. In this case, a complex polynomial is formed (referred to as an Ivakhnenko polynomial [39]. his function usually takes on the form:

$$y = A + \sum_{i=1}^{n} B_i x_i + \sum_{i=1}^{n}\sum_{j+1}^{n} C_{ij} x_i x_j + \sum_{i=1}^{n}\sum_{j+1}^{n}\sum_{k+1}^{n} D_{ijk} x_i x_j x_k \quad (3)$$

where x_i, x_j and x_k are the nodal input variables, and y is the output of an individual neuron (node). A, B_i, C_{ij}, and D_{ijk} are the coefficients of the Ivakhnenko polynomial [39].

3.3 Genetic-based self-organized network

The Genetic-based self-organized network (GBSON) uses polynomial neural networks to represent the model of the system to be identified. Each layer of the polynomial neural network is regarded as a separate optimization problem [34]. The input to the first layer of the network is the independent variables of the data sample. The output of each layer is the peak nodes obtained by the use of a multi-modal Genetic Algorithm. The peak nodes selected to be the output of a layer are also the inputs for the next layer. The population members of the GA are network nodes represented by an eight-field bit string. The two first fields are used to represent the nodes from the previous layer connected to the present node. The other six fields are used to represent the coefficients of a quadratic function that determines the output of the node y [34]:

$$y = a + b z_1 + c z_2 + D z_1 z_2 + e z_1^2 + f z_2^2 \quad (4)$$

where z_1 and z_2 are the outputs of the connected nodes in the previous layer. The fitness measure of a node is given by calculating its description length.

The description length gives a trade off between the accuracy of the prediction and the complexity of the network [34]. The equation used by Kargupta and Smith for calculating the description length is:

$$l = (1/2)\, n \log\left(D_n^2\right) + (1/2)\, m \log(n) \quad (5)$$

where D_n^2 is the mean-square error, m is the number of coefficients in the model selected and n is the number of observations used to determine the mean square error.

The multi-modal GA used in GBSON incorporates the fitness-sharing scheme, where the shared fitness is given by:

$$f_i' = \frac{f_i}{m_i} \quad (6)$$

f_i is the original fitness of the node and m_i is the niche count defined by:

$$m_i = \sum_{j+1}^{N} sh\left(d_{ij}\right) \quad (7)$$

Where

$$sh\left(d_{ij}\right) = \begin{cases} 1 - \left(\dfrac{d_{ij}}{\sigma_s}\right)^{\alpha} & \text{if} \quad d_{ij} < \sigma_s \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

N is the population size and d_{ij} is the Hamming distance between the members of the population i and j. The niche radius σ_s is determined by the equation:

$$\frac{1}{2^l} \sum_{i=0}^{\sigma_s} \binom{l}{i} = \frac{1}{q} \qquad (9)$$

Where l is the string length and q is the number of nodes in the previous network layer. New populations are obtained after applying the genetic operators of tournament selection, single-point crossover and point mutation. A mating restriction is also applied on the members to be crossed. If a number i is to be crossed, its mate j is selected such that $d_{ij} < \sigma_s$. If no such mate can be found then j is selected randomly. The GBSON procedure continues until the GA converges to a layer with a single node [34].

4. METHODOLOGY AND RESULTS

4.1 Method implementation

The described Polynomial Neural Network-based Genetic Algorithm (PNN-GA) technique is used for predicting the global solar irradiation (G), air temperature (T), wind speed (W) and relative humidity (H). The dataset of these parameters were recorded in Algiers (Algeria) by the National Meteorological Office (ONM) [www.meteo.dz], during five years, from 1996 to 2000 and the time scale is one day (*ie.* Figure 1 shows the evolution of the different meteorological time series). Therefore, the objective of this work is to predict these data based on the past observations.

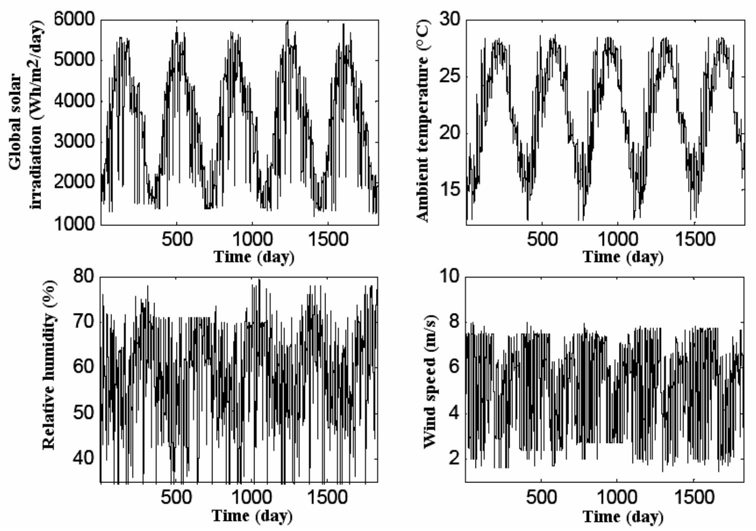


Fig. 1: Evolution of the meteorological time series parameters (global solar irradiation, air temperature, wind speed and relative humidity) used in this simulation

Traditional normalization techniques use linear or logarithmic scaling essentially. These require the designer to supply practical estimates of maximum and minimum values of normalized variables so as to improve the PNN-GA performance. The normalization equation typically used is:

$$y = y_{min} + \frac{x - x_{min}}{x_{max} - x_{min}} \times \left(y_{max} - y_{min} \right) \tag{10}$$

Where $x \in [x_{min}, x_{max}]$ and $y \in [y_{min}, y_{max}]$. Here, x is the original data value (H, G, W or T) and y is the corresponding normalized variable. In this study $y_{min} = 0.1$ and $y_{max} = 0.9$.

Subsequently, the input pattern was assigned as $x(t-1)$, $x(t-2)$, $x(t-3)$, and the desired output was $x(t) = f\left(x(t-1), x(t-2), x(t-3)\right)$.

From the 1825 available data points, 1460 points (four years) were used for training the network while 365 (one year) data points were used for testing and validating the models for each parameter (i.e. H, G, T and W).

A soft computing program is developed by using Matlab (Ver.7.5) environment. The number of generation (Gen) is 500, the crossover probability (P_{cross}) is 0.9 and the mutation probability (P_{mut}) of 0.01. The different step of the used PNN-GA can be summarized as follows:

Step 1. Determine system input variables, form training, and the testing data (normalized data).
Step 3. Determine initial information for constructing the PNN structure.
Step 4. Apply genetic algorithm for determining the PN structure using.
- Representation
- Fitness evaluation
- Selection
- Crossover and mutation
- Elitism.

Step 5. Estimate the coefficients of the polynomial corresponding to the selected node (PN).
Step 6. Select nodes (PNs) with the best predictive capability, and construct their corresponding layer. All nodes of the corresponding layer of PNN architecture are constructed by genetic optimization.
Step 7. Check the termination criterion.
Step 8. Determine new input variables for the next layer. The PNN algorithm is carried out by repeating steps 4-8 consecutively.

4.2 Results and discussion

In this section, we present the simulation results for generating a single-step (one day) prediction based on past observations (three days). However, the PNN-GA predictor can be used for any scale time (*ie.* min, hour, day, etc). For various simulations, the polynomials neural networks constructed by PNN-GA for each parameter (H, G, T and W) are shown in figure 2.

The neurons in the input layer and the output were fixed previously for each parameters, however, the number of neurons in hidden layers where estimated during the learning process. The results of simulation are shown in figure 3a-, b-, c- and d-, from the above curves we can observe that a good agreement between both series (actual and measured) for the different parameters used in this simulation.

In order to test and validate these results we compare the mean, the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE), the Variance (Var), and the correlation coefficient (r) between actual and predicted data by the PNN-GA predictor. The results are illustrated in **Table 1**, according to this table it should be noted that the MAE over the whole data set is not exceed 15.4 % and the RMSE is arranged between 0.00032 and 4.0032, the coefficient of correlation (r) is arranged between 0.9821 to 0.9923. Therefore, based on these goodness tests, we can conclude that the PNN-GA predictor is very suitable for predicting the meteorological time series data in Algiers.

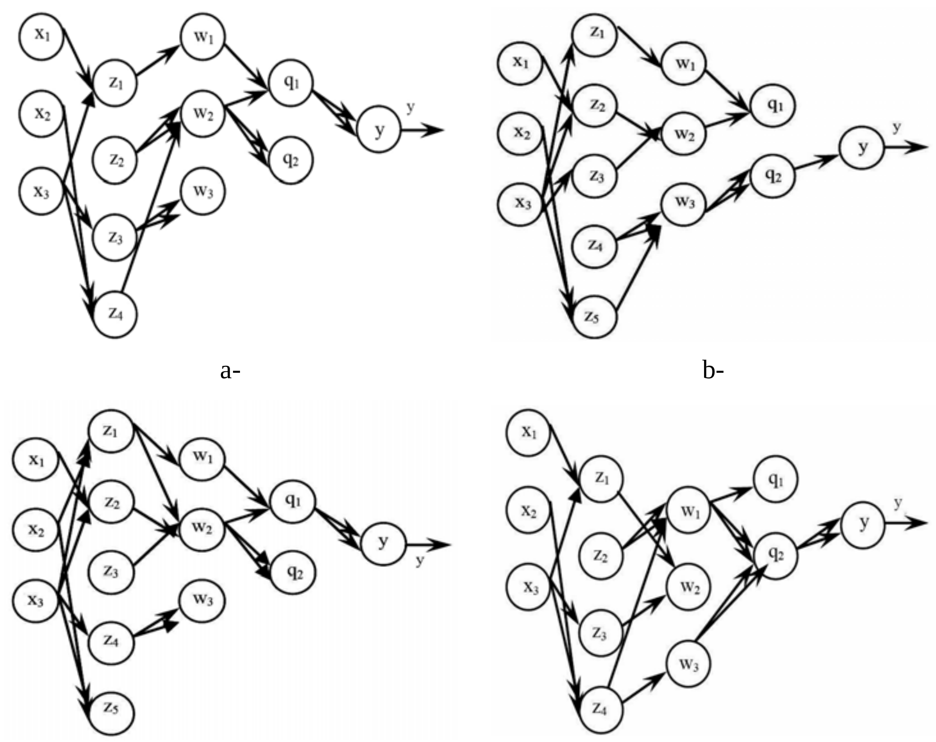


Fig. 2: Network architecture estimated by PNN-GA, a- wind speed,
b- global solar irradiation, c- relative humidity and d- Air temperature

In the neural networks prediction, when the dataset used for training the network is large the obtained results are very accurate. In the present study, four years have been used for training the PNN-GA and one year for testing the PNN-GA, however, less than four years can be used for training the PNN-GA but we can not get good results.

Table 1: Comparison between observed and predicted meteorological data

Data	**Actual mean**	**Estimated mean**	**RMSE**	**MAE** (%)	**Variance**	r
G (Wh/m^2/day)	3.0361e+3	3.0523e+3	0.00416	5.32	0.0813	0.9821
T (°C)	23.452	23.491	0.00032	6.12	0.0145	0.9923
H (%)	63.641	63.625	0.00173	7.62	0.0212	0.9812

W (m/s)	10.537	10.491	0.00213	15.4	0.0745	0.9895

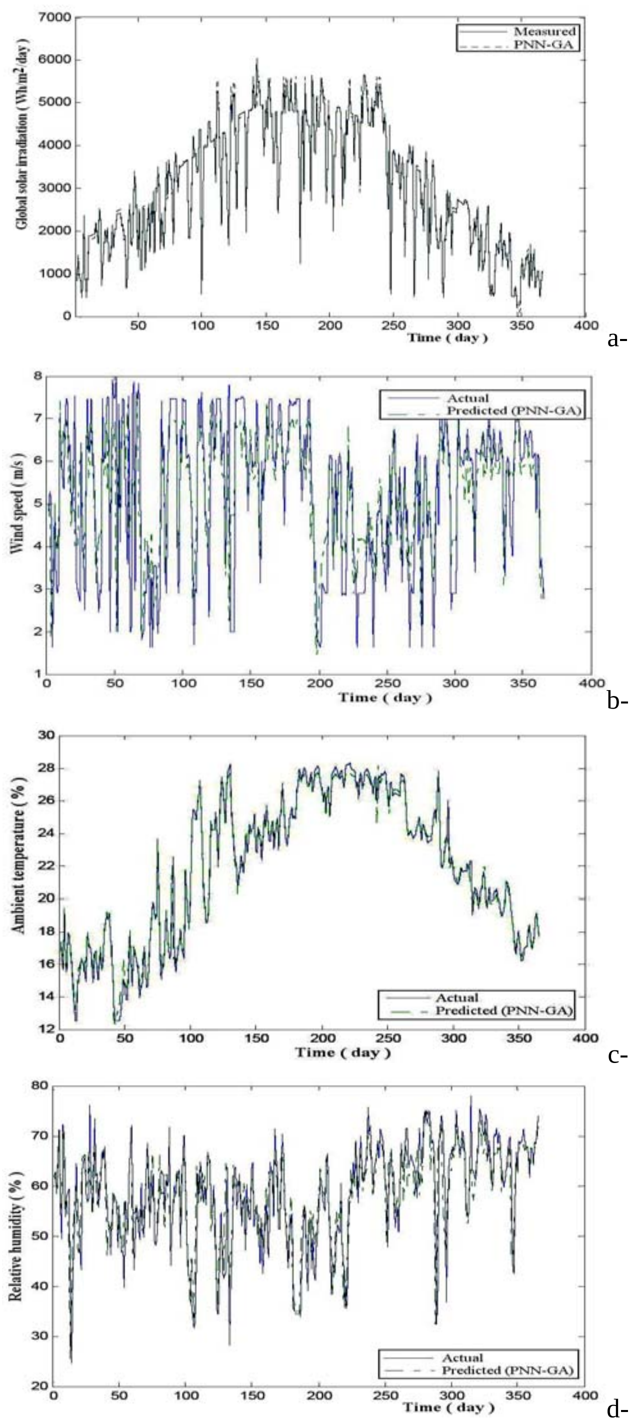


Fig. 3: Actual and predicted meteorological data

In order to show the potential of the PNN-GA predictor to the other ANN architectures (Feed forward neural network 'FFNN', Radial basis function network 'RBFN', Recurrent neural network 'RNN'), wavelet-network and adaptive neuron-fuzzy inference scheme 'ANFIS', (see figure 4a-, b-, c-, d- and e-) we compare the cumulative distribution function F_X between actual and predicted meteorological time series data.

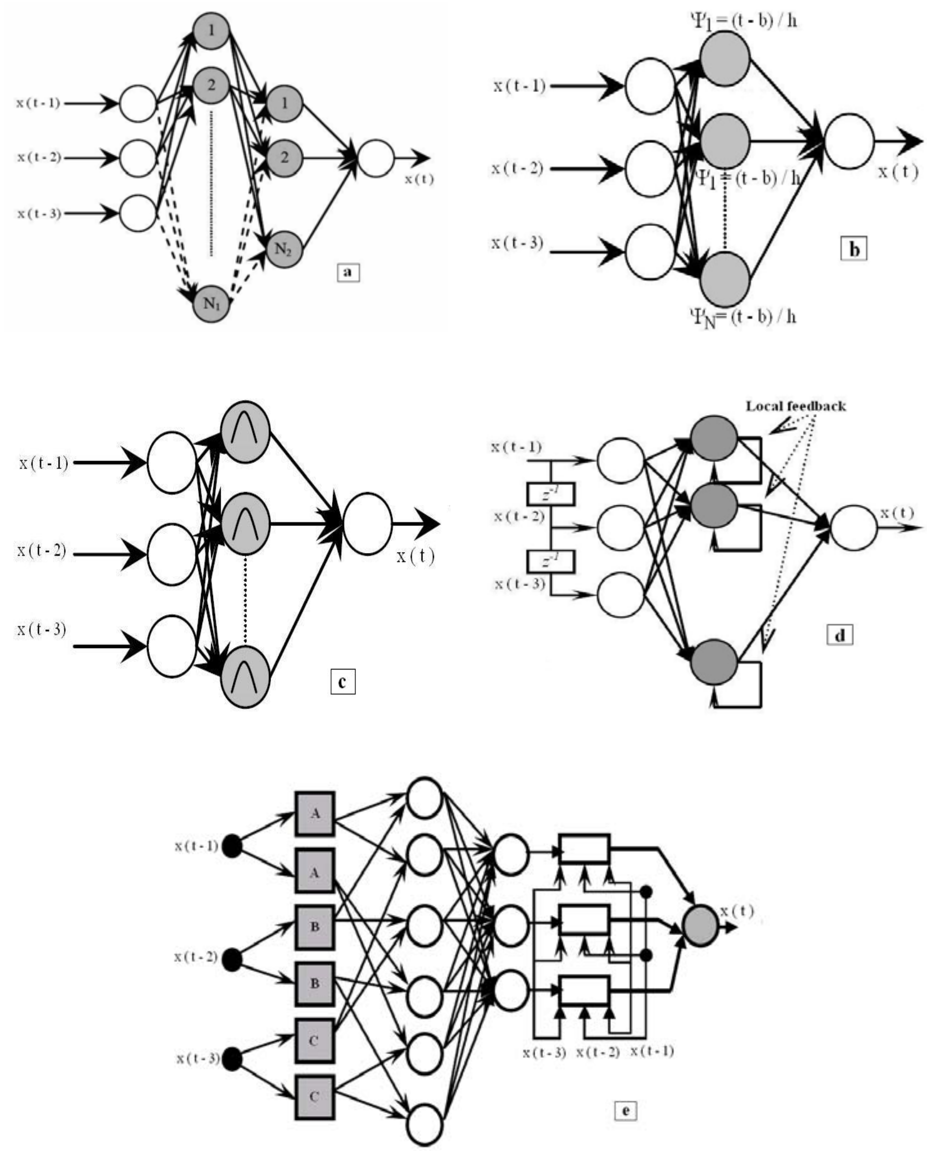


Fig. 4: Different ANNs architecture: a- FFNN, b- Wavenet, c- RBFN, d- RNN with local feedback and e- ANFIS

For the previous ANN architectures, ANFIS, and wavenet we have followed the similar procedure as the PNN-GA predictor. However, for the wavenet model, we have used the Morlet basic wavelet function (Ψ). The ANFIS is a technique for automatically tuning (Backpropagation algorithm) Sugeno-type fuzzy inference system based on some collection of input-output data, the ANFIS predictor uses first-order Sugeno-type systems, single output derived by weighted defuzzification, and membership function type that is a generalized bell curve. Obtained results are shown in figure 5a-, b-, c- and d-.

As can be seen from the figure 5a-, b-, c- and d- the measured meteorological time series are very close the predicted ones by ANN's, Wavenet and PNN-GA, especially for air temperature, relative humidity, global solar irradiation, while for the wind speed time series the results are approximately accurate.

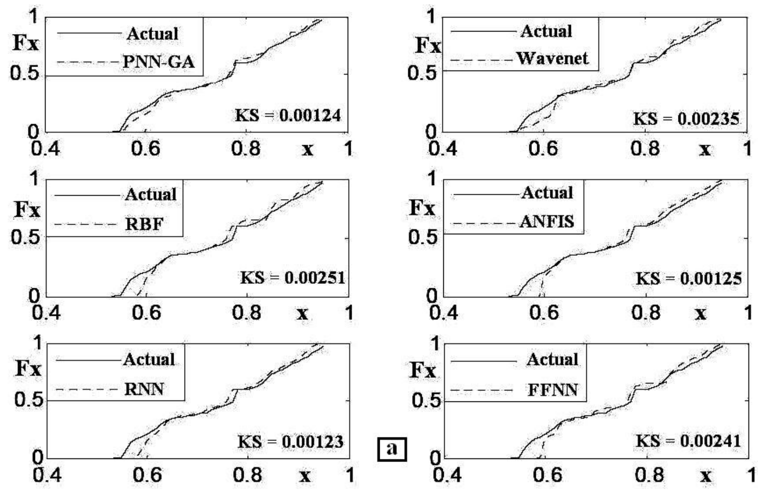


Fig. 5a-: The cumulative distribution functions between actual and predicted global solar irradiation by using PNN-GA, Wavelet, RBFN, ANFIS, RNN and FFNN

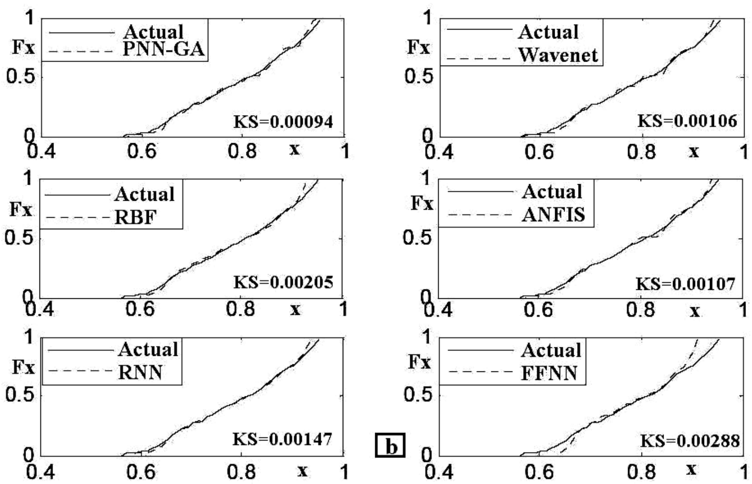


Fig. 5b-: The cumulative distribution functions between actual and predicted air temperature by using PNN-GA, Wavelet, RBFN, ANFIS, RNN and FFNN

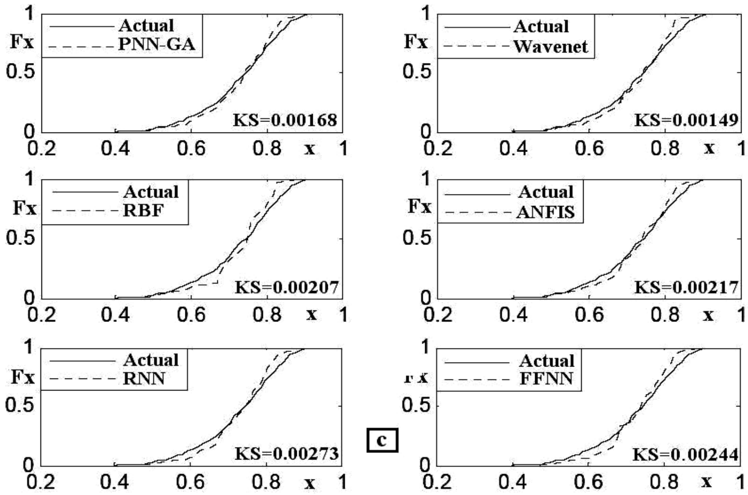


Fig. 5c-: The cumulative distribution functions between actual and predicted relative humidity by using PNN-GA, Wavelet, RBFN, ANFIS, RNN and FFNN

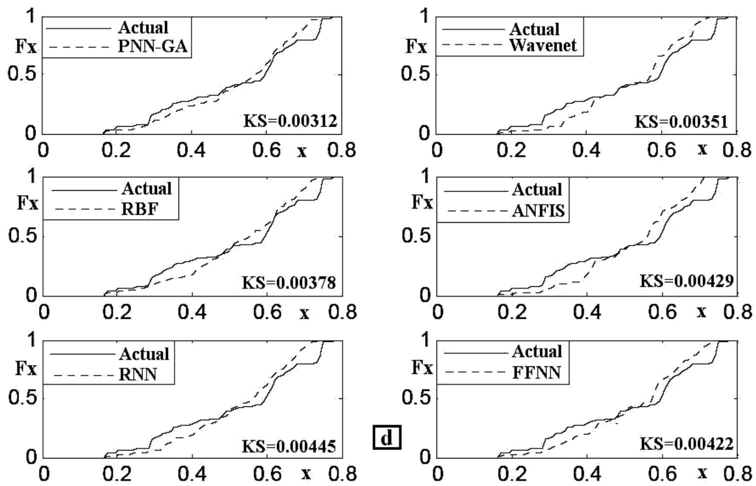


Fig. 5d-: The cumulative distribution functions between actual and predicted wind speed by using PNN-GA, Wavelet, RBFN, ANFIS, RNN and FFNN

In order to test and compare the different models, we have used the correlation coefficient r. The simulation results are illustrated in **Table 2**. Form this table we can conclude the following remarks:

For all meteorological time series used in this study, the correlation coefficient r obtained by the PNN-GA is more than 0.98 , compared to the other ANNs, wavenet and ANFIS which is less than 0.98. The number of generation is fixed at 500 generations, while the architecture is changed according to the nature of the time series data; therefore, we can found the same architecture for global solar irradiation and relative humidity <3x5x3x2x1>, three neurons in the input layer, three hidden layers within 5, 3 and 2 neurons respectively, one neuron in the output layer. This architecture <3x4x3x2x1> is for both air temperature and wind speed.

Table 2: Comparison between measured and predicted data by using PNN-GA, Wavenet, RBFN, ANFIS, RNN and FFNN

Models	Global solar irradiation ($Wh/m^2/day$)		
	Architecture	Nb of iterations	r
PNN-GA	<3x5x3x2x1>	500 gen.	0.9854
Wavenet	<3x8x1>	935	0.9822
RBF	<3x17x1>	1425	0.9793
ANFIS	<3x3x2x3x1>	890	0.9813
RNN	<3x11x1>	1165	0.9811
FFNN	<3x17x12x1>	4512	0.9787
	Air temperature (°C)		
	Architecture	Nb of iterations	r
PNN-GA	<3x4x3x2x1>	500 gen.	0.9923
Wavenet	<3x13x1>	1245	0.9902
RBF	<3x15x1>	1322	0.9871
ANFIS	<3x3x2x3x1>	932	0.9897
RNN	<3x9x1>	995	0.9911
FFNN	<3x13x9x1>	3117	0.9751
	Relative humidity (%)		
	Architecture	Nb of iterations	r
PNN-GA	<3x5x3x2x1>	500 gen.	0.9832
Wavenet	<3x11x1>	1211	0.9811
RBF	<3x17x1>	1622	0.9771
ANFIS	<3x3x4x3x1>	1132	0.9794
RNN	<3x13x1>	1045	0.9792
FFNN	<3x12x11x1>	4274	0.9732
	Wind speed (m/s)		
	Architecture	Nb of iterations	r
PNN-GA	<3x5x3x2x1>	500 gen.	0.9807
Wavenet	<3x5x3x2x1>	1211	0.9785
RBF	<3x12x1>	1244	0.9726
ANFIS	<3x3x4x3x1>	1132	0.9722
RNN	<3x11x1>	1020	0.9745
FFNN	<3x17x14x1>	5551	0.9718

Concerning the wavenet model, from this table we can observe that the (r) is arranged between 0.9785 and 0.9902. The architecture of this model is varied according to the time series data, the maximum amount of neurons in the hidden layer is 13. Generally, one layer is sufficient for predicting data, this model is relatively fast compared to the other ANN architectures.

As can be seen, from this table the ANFIS and the RNN neural network can provides approximately similar results according to the correlation coefficient (r) test. However, there is a difference between both architectures. In the RNN architecture, one hidden layer is sufficient for prediction the meteorological time series. However, the RNN need much computing time compared with the ANFIS.

The RBFN and the FFNN also give acceptable results but both networks need more computing time compared to the previously architectures, it should be noted that the FFNN need more than one hidden layer for obtaining a good accurately results, the maximum amount of neurons in the hidden layers is varied according to time series data. The computing time of the RBF is less than needed by the FFNN. In the point of view correlation coefficient (r), we can observe approximately the same results by both models, r is approximately 0.975.

In the point of view computing time and precision, we can classify the different models as follow: PNN-GA, wavenet, RNN, ANFIS, RBF and FFNN. Additionally, we can consider the ANFIS and RNN as the similar level.

Finally, according to the above remarks we conclude that the PNN-GA predictor can provides more accurate results than other investigated ANN architectures, wavenet and ANFIS. However, we can not negligee the results obtained by the other ANN architectures, especially RNN, ANFIS and wavenet.

5. APPLICATION FOR SIZING AND PREDICTION OF THE OUTPUT ENERGY FROM PHOTOVOLTAIC SYSTEMS

In this section, we will use the predicted meteorological time series by the PNN-GA predictor for some applications in renewable energy systems: sizing of stand-alone photovoltaic (PV) systems, sizing of hybrid wind photovoltaic system, estimating of the output energy from a stand-alone photovoltaic system. In order to test the effectiveness of the developed PNN-GA predictor we make a comparison with experimental meteorological time series.

5.1 Application for sizing of stand-alone PV systems

The most important parameter used for sizing of photovoltaic systems is the global solar irradiation. The prediction of this parameter play very important role for estimating the number of photovoltaic modules and batteries , therefore, in this subsection we present an example for sizing of stand-alone PV system by using the measured (G) and predicted global solar irradiation ($\tilde{G}$) by the developed PNN-GA.

A basic configuration of a stand-alone photovoltaic system with constant load was considered (Fig. 6). According to [40] the sizing pair C_A and C_S can be given by the following formulae:

$$C_A = \frac{\eta_{PV}.A.G}{L} \qquad C_S = \frac{C}{L} \tag{11}$$

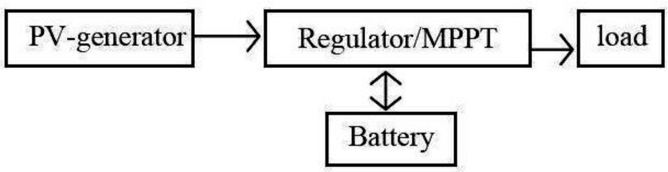


Fig. 6: Simplified schematic of stand-alone PV system

Given a location and a load, two general ideas are intuitive: Firstly, it is possible to find many different combinations of C_A and C_S leading to the same Loss of Load

Probability (LLP) value. Secondly, the larger the PV-system size is, the greater the cost and the lower the LLP. The numerical method [40] is chosen for this purpose. Figure 7 shows the sizing curve based on the measured and predicted global solar radiation for a LLP of 1 % and load of 1 kW/day. As can be seen, good agreement is obtained between both curves; the correlation coefficient is 97 % which is very satisfactory. The optimal pair (C_{Aop} , C_{Sop}) can be estimated based on the economical cost, which allow to calculate the number of PV modules and batteries required for a load of 1 kWh/day [41].

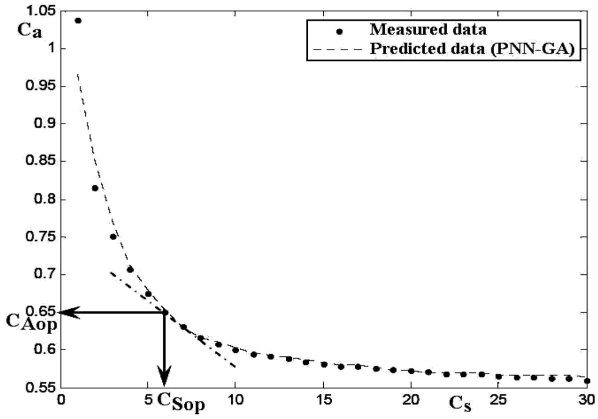


Fig. 7: Comparison between the sizing curves for LLP = 1 % and L = 1 kW/day, by using measured and predicted (PNN-GA) solar irradiation

5.2 Application for sizing of hybrid photovoltaic wind power generation systems

Hybrid photovoltaic systems most commonly take the form of photovoltaic systems combined with wind turbines or diesel generators. Due to their almost complementary power production characteristics, they are usually used in hybrid system configurations. A simplified schematic of a hybrid photovoltaic–wind power generation (HPVWPG) system is shown in figure 8, this system consists of PV generator, wind turbine, battery, inverter, controller and data acquisition system.

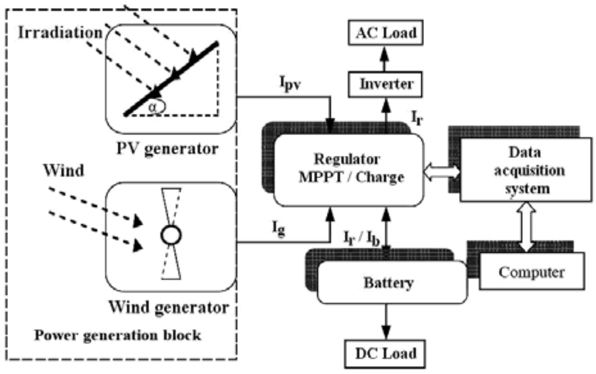


Fig. 8: Simplified schematic of hybrid photovoltaic-wind power generating system

The optimization method can help to guarantee the lowest investment with adequate and full use of the solar system, wind system and battery bank, so that the hybrid system can work at optimum, however, in this kind of systems three meteorological parameters are necessary: global solar radiation, wind speed and air temperature.

The genetic algorithm is used for optimal sizing of the HPVWPG, which allow to estimate the number of PV modules, WGs, and batteries capacity for a specific load. The economic cost function of the system is essentially based on the PV-array cost, battery cost, wind generator cost and the maintenance cost. The total cost of the system can be given by the following formulas [42]:

$$J\left(N_{pv}, N_{wg}, N_{bat}\right) = C_{pv} \times N_{pv} + C_{bat} \times N_{bat} + C_{wg} \times N_{wg} + C_M \qquad (12)$$

$$C_{pv} = \frac{H \cdot \eta_{pv} \cdot N_{pv} \cdot M_{pv}}{L}, \; C_{bat} = \frac{\eta_{bat} \cdot N_{bat} \cdot P_{bat}}{L}, \; C_w = \frac{\eta_{wgt} \cdot N_{wg} \cdot M_{wg}}{L} \qquad (13)$$

Table 3 shows the analyzed cost used in this simulation (PV-array, battery and wind characteristic).

Table 3: Analyzed cost used in the simulation

Photovoltaic array characteristic

V_{max} (V)	I_{max} (A)	V_{oc} (V)	I_{sc} (A)	P_{max}(W)	Capital cost	Maintenance cost
19	5	23	6.5	110	655.00 euro	6.00 euro

Battery characteristic

Voltage (V)	Capacity(Ah)	Efficiency (%)	Capital cost	Maintenance cost
12	240	80	270 euro	3

Wind characteristic

Power rating (W)	H low (m)	H high (m)	Capital cost	Maintenance cost
500	10	17	600 euro	14

The estimated total cost by using the genetic algorithm based on measured data and those predicted by PNN-GA are illustrated in **Table 4**. We can observe that, 4 modules, 1 wind generator and 4 batteries is obtained based on the measured (G, W and T), and we found approximately the same results when we use the predicted data ($\tilde{G}$, $\tilde{W}$ and $\tilde{T}$) which confirm the validity of the developed predictor.

Table 4: Total cost estimated by using measured data and predicted (PNN-GA) solar irradiation, wind speed and air temperature

Method	# - PV mod.	# of WGs	# - Batteries	Tilt (θ)	Total cost (eu)
	Measured data (H, G and T)				
Genetic Algorithm	3	1	4	36°43	2413
	Predicted data ($\tilde{G}$, $\tilde{W}$ and $\tilde{T}$) by PNN-GA				
Genetic Algorithm	3	1	4	36°43	2413

5.3 Application for predicting the output energy form the photovoltaic generator

The technology for power production from renewable energy resources is now widely available, reliable and matured. The use of renewable energy resources, such as photovoltaic (PV) systems is rapidly expanding and has an increasing role in electricity generation, providing pollution-free and secure power. Prediction the output of a solar PV system with reasonable accuracy should be useful in managing and planning an electrical power system to meet the power requirement, it can be used for studying the performance of PV systems under different climate conditions. In this situation, two parameters can play very important role in order to predict the output energy from the photovoltaic generator, which the global solar radiation and the air temperature.

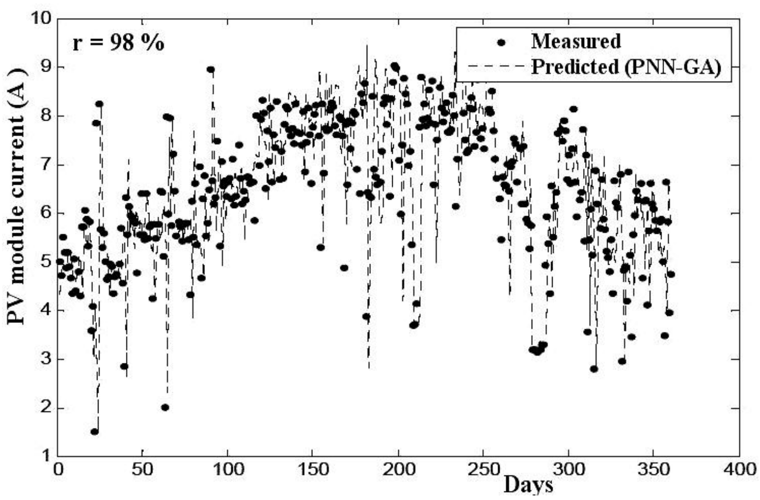
In this subsection, we attempt to use the measured and predicted global solar irradiation ($\tilde{G}$) and air temperature ($\tilde{T}$) for predicting the output energy generation from a photovoltaic module. The *Isofoton* PV module employed consists of 36 square single crystal silicon cells. The total peak power of the PV module is 75 W_p, the maximum PV module voltage V_{max} is 17.3 and the maximum PV module current is I_{max} 20 A. The main specifications of the PV module are shown in **Table 5**.

Table 5: Main specifications of the PV module employed

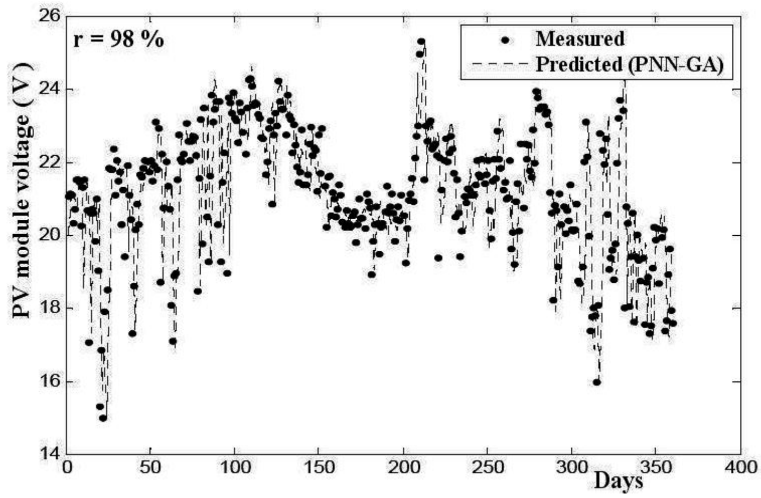
H (W/m^2)	T (°C)	V_{oc} (V)	I_{sc} (A)	V_{max} (V)	I_{max} (A)	P_{max} (W)
1000	25	21.6	4.67	17.3	4.34	75

Figure 9a-, b- and c- show a comparison between measured and predicted PV module current, power and voltage by using the measured and predicted global solar irradiation and air temperature. From these results it is clearly shown that almost agreement are obtained between both data. The correlation coefficient is arranged between 97 % and 98 %.

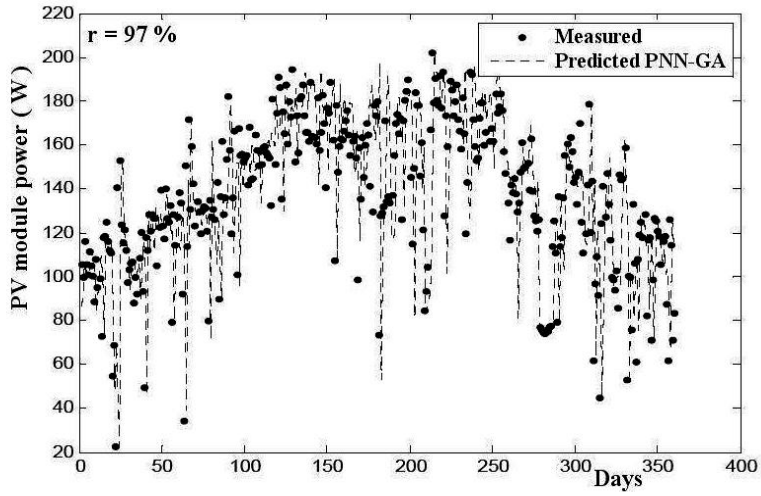
From the above applications, we can conclude that developed PNN-GA is very suitable for meteorological time series prediction.



a-



b-



c-

Fig. 9: Comparison between measured and predicted: a- PV module current, b- voltage and c- power, based on the measured and predicted (PNN-GA) irradiation and air temperature

6. CONCLUSION AND PERSPECTIVES

In this paper, a methodology for predicting the meteorological time series (global solar irradiation, air temperature, wind speed and relative humidity) used for renewable energy systems is developed. This is based on the use of hybrid polynomial neural network combined with a genetic algorithm, called Genetics-Based Self-Organizing Network. The developed PNN-GA predictor has been used for predicting the future

($x(t)$) data based on the past three observed values ($x(t-1), x(t-2), x(t-3)$) with a very accurate results compared to the experimental data.

Additionally, it has been shown that the obtained results by the PNN-GA predictor are more accurate compared with those obtained by other ANNs architecture, wavenet and ANFIS. The potential and the effectiveness of the PNN-GA has been demonstrated in this paper for predicting the meteorological time series, The correlation coefficient (r) is between 0.9821 and 0.9923, the mean relative error over the whole data set is not exceed 15.4 %. In addition, the PNN-GA predictor can be used for different time scale (min, hour, and day).

An application of the predicted data for sizing, and predicting the output energy from photovoltaic systems (stand-alone PV, and wind-PV) demonstrate the effectiveness of the developed predictor (PNN-GA).

For future work, an attempt will be made to predict the meteorological time series for more than one day, in addition, the prediction of one meteorological parameter based on the other parameters, i.e. predicting of the air temperature based on relative humidity, precipitation, dew point, clouds, wind speed and wind direction. Also, we try to use the FPGA (Field Programmable Gate Array) in order to implement the developed intelligent predictor PNN-GA.

NOMENCLATURE

η_{bat} : Batteries efficiency (%)
η_{pv} : PV-array efficiency (%)
η_{wgt} : Wind-generator efficiency (%)
N_{bat} : Number of battery
N_{pv} : Number of PV-array
N_{wg} : Number of wind generator
T : Temperature (°C)
W : Wind speed (m/s)
WG : Wind generator
C_{bat}, C_{pv}, C_{W} : Sizing parameters
C_u : Useful capacity - battery (Wh)
G : Irradiation (Wh/m^2/day
L : Load (W)
PV : Photovoltaic
RMSE : Root mean square error
P_{cross} : Crossover probability
P_{mut} : Mutation probability

ACKNOWLEDGMENTS

This work was supported by the Ministry of Higher Education & Scientific Research (Algiers) under project number: J0201720080012.

REFERENCES

[1] A. Mellit and S.A. Kalogirou, '*Artificial Intelligence Techniques for Photovoltaic Applications: A Review*', Progress in Energy and Combustion Science, Vol. 34, N°5, pp. 574 – 632, 2008.*

[2] D. Sharon, '*On the Further Development of Gringorten's Stochastic Model for Climatological Predictions*', Journal of Applied Meteorology, Vol. 64, N°4, pp. 625 – 630, 1967.

[3] R.J. Aguiar, M. Collares-Pereira and J.P. Conde, '*Simple Procedure for Generating Sequences of Daily Radiation Values using Library of Markov Transition Matrices*', Solar Energy, Vol. 40, N°3, pp. 269 – 279, 1988.

[4] H. Shuichi, M. Mamoru and I. Toshikazu, '*Statistical Time Series Models of Solar Radiation and Outdoor Temperature - Identification of Seasonal Models by Kalman Filter*', Energy and Buildings, Vol. 15, N°3-4, pp. 373 – 383, 1990 -1991.

[5] M.F. Macchiato, L. La Rotonda, V. Lapenna and M. Ragosta, '*Time Modelling and Spatial Clustering of Daily Ambient Temperature: An Application in Southern Italy*', EnvironMetrics, Vol. 6, N°1, pp. 31 – 53, 1995.

[6] H. Bouhaddou, M.M. Hassani, A. Zeroual and A.J. Wilkinson, '*Stochastic Simulation of Weather Data using Higher Order Statistics*', Renewable Energy, Vol. 12, N°1, pp. 21 – 37, 1997.

[7] A.M. Omer, '*Compilation and Evaluation of Solar and Wind Energy Resources in Sudan*', Vol. 12, N°1, pp. 39 - 69, 1997.

[8] D. Wilks and R.L. Wilby, '*The Weather Generation Game: A Review of Stochastic Weather Models*', Progress in Physical Geography, Vol. 23, N°3, pp. 329 – 357, 1999.

[9] I. Matyasovszky, '*A Nonlinear Approach to Modelling Climatological Time Series*', Theoretical and Applied Climatology, Vol. 69, pp. 139 – 147, 2001.

[10] J. Svec and M. Stevenson, '*Modelling and Forecasting Temperature Based Weather Derivatives*', Global Finance Journal, Vol. 18, N°2, pp. 185 – 204, 2007.

[11] C.G. Kilsby, P.D. Jones, A. Burton, A.C. Ford, H.J. Fowler, C. Harpham, P. James, A. Smith and R.L. Wilby, '*A Daily Weather Generator for Use in Climate Change Studies*', Environmental Modelling and Software, Vol. 22, pp. 1705 – 1719, 2007.

[12] I. Moradi, '*Quality Control of Global Solar Radiation using Sunshine Duration Hours*', Energy, Vol. 34, N°1, pp. 1 – 6, 2009.

[13] P.M. Danilo and A.C. Jonathon, '<u>Recurrent Neural Networks for Prediction</u>', John Wiley and Sons Ltd. 2001.

[14] A. Mellit, '*Artificial Intelligence Techniques for Modelling and Forecasting of Solar Radiation Data: A Review*', International Journal of Artificial Intelligence and Soft Computing, Vol. 1, N°1, pp. 52 – 76, 2008.

[15] R.P. Lang and S.F. Lombargo, '*Atmospheric Turbulence, Meteorological Modeling and Aerodynamics*', [A. Mellit, 'Artificial intelligence technique for modeling and forecasting of meteorological data: a survey'], Nova Publisher, Zditor. Hauppauge NY, pp. 101 – 105, 2009.

[16] J.B. Elsner and A.A. Tsonis, '*Nonlinear Prediction, Chaos and Noise*', Bulletin of the American Meteorological Society, Vol. 73, N°1, pp. 49 – 60, 1992.

[17] Y. Yuval, '*Neural Network Training for Prediction of Climatological Time Series, Regularized by Minimization of the Generalized Cross-Validation Function*', Monthly Weather Review, Vol. 128, N°5, pp. 1456 – 1473, 2000.

[18] R. Kretzschmar, P. Eckert, D. Cattani and F. Eggiman, '*Neural Network Classifiers for Local Wind Prediction*', Journal of Applied Meteorology, Vol. 43, N°5, pp. 727 – 738, 2004.

[19] I. Tasadduq, S. Rehman and K. Bubshait, '*Application of Neural Networks for the Prediction of Hourly Mean Surface Temperatures in Saudi Arabia*', Renewable Energy, Vol. 25, N°4, pp. 545 – 554, 2002.

[20] V. Gomez and A. Casanovas, '*Fuzzy Modelling of Solar Irradiance on Inclined Surfaces*', Solar Energy, Vol. 75, N°4, pp. 307 – 315, 2003.

[21] R. Perez, R. Seals, P. Ineichen, R. Stewart and D. Menicucci, '*A New Simplified Version of the Perez Diffuse Irradiance Model for Tilted Surfaces*', Solar Energy, Vol. 39, N°3, pp. 221 – 231, 1987.

[22] I. Maqsood and A. Abraham, Weather Analysis using Ensemble of Connectionist Learning Paradigms', Applied Soft Computing, Vol. 7, pp. 995 – 1004, 2007.

[23] A. Mellit, M. Benghanem, A. Hadj Arab and A. Guessoum, '*A Simplified Model for Generating Sequences of Global Radiation Data for Isolated Sites: Using Artificial Neural Network and a Library of Markov Transition Matrices*', Solar Energy, Vol. 79, N°5, pp. 468 – 482, 2005.

[24] A. Mellit, M. Benghanem and S.A. Kalogirou, '*An Adaptive Wavelet-Network Model for Forecasting Daily Total Solar Radiation*', Applied Energy, Vol. 83, N°7, pp. 705 – 722, 2006.

[25] I. Maqsood, M.R. Khan and A. Abraham, '*An Ensemble of Neural Networks for Weather Forecasting*', Neural Computing and Application, Vol. 13, N°2, pp. 112 – 122, 2004.

[26] M. Bilgili, B. Sahin and A. Yasar, '*Application of Artificial Neural Networks for the Wind Speed Prediction of Target Station Using Reference Stations Data*', Renewable Energy, Vol. 32, N°14, pp. 2350 – 2360, 2007.

[27] A. Mellit, S.A. Kalogirou, S. Shaari, H. Salhi and A. Hadj Arab, '*Methodology for Predicting Sequences of Mean Monthly Clearness Index and Daily Solar Radiation Data in Remote Areas: Application for Sizing a Stand-Alone PV System*', Renewable Energy, Vol. 33, N°7, pp. 1570 – 1590, 2008.

[28] E. Tulcan-paulescu and M. Paulescu, '*Fuzzy Modelling of Solar Irradiation Using Air Temperature Data*', Theoretical and Applied Climatology, Vol. 91, N°1-4, pp. 181 – 192, 2007.

[29] M. Chaabene and M. Ben Ammar, '*Neuro-Fuzzy Dynamic Model with Kalman Filter to Forecast Irradiance and Temperature for Solar Energy Systems*'. Renewable Energy, Vol. 33, N°7, pp. 1435 – 1443, 2008.

[30] K. Moustris, A.G. Paliatsos, A. Bloutsos, K. Nikolaidis, I. Koronaki and K. Kavadias, '*Use of Neural Networks for the Creation of Hourly Global and Diffuse Solar Irradiance Data at Representative Locations in Greece*', Renewable Energy, Vol. 33, N°5, pp. 928 – 932, 2008.

[31] S. Rehman and M. Mohandes, '*Artificial Neural Network Estimation of Global Solar Radiation using Air Temperature and Relative Humidity*', Energy Policy, Vol. 32, N°2, pp. 571 – 576, 2008.

[32] M. Benghanem, A. Mellit and S.N. Alamri, '*ANN-Based Modeling and Prediction of Daily Solar Radiation*', Energy Conversion and Management, Vol. 50, N°7, pp. 1644 – 1655, 2009.

[33] A. Mellit, S. Shaari, H. Mekki and N. Khorissi, '*FPGA-Based Artificial Neural Network for Prediction of Solar Radiation Data from Sunshine Duration and Air Temperature*', IEEE International Conference on Computational Technologies in Electrical and Electronics Engineering, Sibircon-2008, pp. 118-123, Novosibirsk, Russia, 21-25 July, 2008.

[34] A. Foka, '*Time Series Prediction Using Evolving Polynomial Neural Networks'*, MSc dissertation, Dept. of Elec. Eng. & Elec., Univ. of Manchester, Institute of Science and Technology, 1999.

[35] E.F. Vasechkina and V.D. Yarin, '*Evolving Polynomial Neural Network by Means of Genetic Algorithm: Some Application*', Journal of Complexity International Vol. 9, pp. 729 – 744, 2001.

[36] N. Nariman-Zadeh, A. Darvizeh, A. Jamali and A. Moeini, '*Evolutionary Design of Generalized Polynomial Neural Networks for Modelling and Prediction of Explosive Forming Process*', Journal of Materials Processing and Technology, Vol. 164 – 165, pp. 1561 – 1571, 2005.

[37] J.H. Holland, '*Adaptation in Natural and Artificial Systems*', University of Michigan Press; 1975

[38] O. Sung-Kwun, P. Witold, K. Wan-Su and K. Hyun-Ki,'*GA-Based Polynomial Neural Networks Architecture and Its Application to Multi-Variable Software Process*', Q. Yang and G. Webb (Eds.): PRICAI 2006, LNAI 4099, pp. 834 – 838, 2009.

[39] A.G. Ivakhnenko, '*Heuristic Self-Organising Systems in Cybernetics*', Naukova Dumka, Kiev, 1971.

[40] E. Lorenzo, '*Solar Electricity. Engineering of Photovoltaic Systems*', Progensa, Seville. 1994.

[41] A. Mellit, S.A. Kalogirou, L. Hontoria and S. Shaari, '*Artificial Intelligence Technique for Sizing of Photovoltaic Systems: A Review*', Renewable and Sustainable Energy Reviews,Vol. 13, N°2, pp. 406 – 419, 2009.

[42] A. Mellit, '*Application of Genetic Algorithm and Neural Network for Sizing of Hybrid Photovoltaic Wind Power Generation (HPVWPG) Systems*', International Journal of Renewable Energy Technology (IJRET), Vol. 1, N°2, pp. 139 - 154, 2009.