Neural Network Ensemble-based Solar Power Generation Short-Term Forecasting

A. Chaouachi, R.M. Kamel, R. Ichikawa, H. Hayashi, and K. Nagasaka

Abstract—This paper presents the applicability of artificial neural networks for 24 hour ahead solar power generation forecasting of a 20 kW photovoltaic system, the developed forecasting is suitable for a reliable Microgrid energy management. In total four neural networks were proposed, namely: multi-layered perceptron, radial basis function, recurrent and a neural network ensemble consisting in ensemble of bagged networks. Forecasting reliability of the proposed neural networks was carried out in terms forecasting error performance basing on statistical and graphical methods. The experimental results showed that all the proposed networks achieved an acceptable forecasting accuracy. In term of comparison the neural network ensemble gives the highest precision forecasting comparing to the conventional networks. In fact, each network of the ensemble over-fits to some extent and leads to a diversity which enhances the noise tolerance and the forecasting generalization performance comparing to the conventional networks.

Keywords—Neural network ensemble, Solar power generation, 24 hour forecasting, Comparative study.

I. INTRODUCTION

THE need for more flexible electric systems, changing regularity and economic scenarios, energy saving and environmental impacts are providing impetus to the development of new power system perception. In this regard, Microgrid (MG) can be considered as one of the most promising concepts; a MG is defined as an integrated power delivery system consisting of interconnected loads, storages facilities and distributed generation mainly composed of Renewable Sources (RS) such as solar and wind energy. As an integrated system, a MG can operate in grid-connected or autonomous mode (island mode). The optimization of the operation cost is an important challenge for MG development and competition, that is greatly depending on the management of the power generation from renewable sources where the generation capacity varies largely with weather conditions, hence the usefulness of the RS power generation forecasting.

A.Chaouachi, Department of Electronic and Information Engineering, Tokyo University of Agriculture and Technology, 2-24-16, Nakamachi, Koganei-shi, Tokyo 184-8588, Japan (e-mail: A.chaouachi@gmail.com, Tel & FAX: +81-42-388-7481)

R.M.Kamel, Department of Electronic and Information Engineering, Tokyo University of Agriculture and Technology, 2-24-16, Nakamachi, Koganei-shi, Tokyo 184-8588, Japan (r_m_kamel@yahoo.com).

R. Ichikawa, Toshiba Corporation 1-1-1 Shibaura, Minato-ku, Tokyo 105-8001, Japan, (ryoichi.ichikawa@toshiba.co.jp).

H. Hayashi, Toshiba Corporation 1-1-1 Shibaura, Minato-ku, Tokyo 105-8001, Japan, (hideki.hayashi@toshiba.co.jp).

K.Nagasaka Department of Electronic and Information Engineering, Tokyo University of Agriculture and Technology, 2-24-16, Nakamachi, Koganei-shi, Tokyo 184-8588, Japan (bahman@cc.tuat.ac.jp).

This study focuses on the power generation forecasting of a photovoltaic system, recently Artificial Neural Network (ANN) has been applied for such purpose regarding to its approximation capability of any continuous nonlinear function with arbitrary accuracy that offer an effective alternative to more traditional statistical techniques proposed so far [1]-[5].

Maher and Mohsen (2007) [6] proposed a neuro-fuzy network associated to Kalman Filter for both medium-term and short-term irradiance forecasting. The neuro-fuzzy estimator presents daily time distribution of meteorological parameters relying on climatic behavior of the previous day. Auto-Regressive Moving Average (ARMA) model of the medium-term forecasting is associated to Kalman filter for short-term forecasting. J.C. Cao and SH.Cao (2006) [7] proposed a Recurrent Back-Propagation Neural Network (RBPNN) combined to wavelet analysis to forecast solar irradiance that was beforehand sampled into time-frequency domain using wavelet transformation, each domain was associated to a RBPNN. K.S.Reddy and Manish Ranjan (2003) [1] compared a Multi-layred Perceptron Neural Network (MLPNN) to other correlation models to estimate monthly and hourly values of global radiation. Yingni Jiang (2008) [8] proposed a MLPNN to predict actual values of monthly mean daily diffuse solar radiation in Zhengzhou (China), solar radiation data was collected from nine stations in different cities. Eight stations were used for the training process while the remaining one was used to test and validate the proposed neural network. Mohandes et al [2] proposed a Radial Basis Function Neural Network (RBFNN) to estimate the monthly mean daily values of solar radiation on horizontal surfaces, the obtained results was compared with classical regression model.

As mentioned above, several solar radiation parameters has been estimated or forecasted in the literature such as: global solar radiation, irradiance, irradiation and clearness index. Some authors reported that their results can be extended to predict solar energy conversion for a photovoltaic system. Nevertheless, the prediction error would logically increase regarding to the incertitude of the solar system model and the deviation between the provided forecasting and the local solar radiation level associated to the photovoltaic dispositive. The objective of this study is to validate several ANN models that supplies an immediate and reliable 24 hour ahead of Solar Power Generation (SPG) forecasting for a 20 kW photovoltaic system located in Tokyo University of Agriculture and Technology (TUAT). In total four ANN's has been developed, consisting on a MLPNN, RBFNN, Recurrent Neural Network

(RNN) and Neural Network Ensemble (NNE). Each neural network will be evaluated for different ranges of climatic conditions basing on error forecasting performance by the mean of statistical and graphical methods.

II. PROPOSED ARTIFICIAL NEURAL NETWORKS
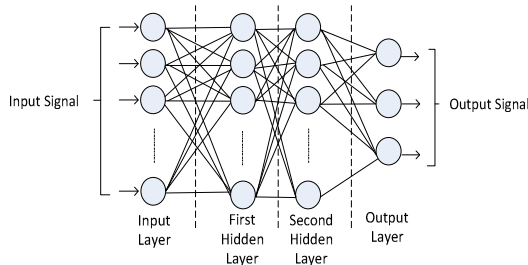
A. Multi Layred Perceptron Neural Network

Fig. 1. Architecture graph of a MLPNN with two hidden layers

Multi-layred Perceptrons has been applied successfully to solve some difficult and diverse problems basing on a preliminary supervised training with error back propagation algorithm using an error correction learning rule. Basically, error back learning consists in two pass through the different layers of the network, a forward pass and backward pass. In the forward pass an activity pattern (input vector) is applied to the sensory nodes of the network, its effect propagates through the network layer by layer to produce an output as actual response. During the backward pass synaptic weights are adjusted in accordance to an error correction-rule. The error signal (subtracted from a desired value) is then propagated backward through the network against the direction of the synaptic connections [9]. In general MLPNN's can have several hidden layers (Fig.1), however according to K.M.Hornik [10] a neural network with single hidden layer is able to approximate a function of any complexity. If we consider a MLPNN with one hidden layer, tanh as an activation function and a linear output unit, the equation describing the network structure can be expressed as:

$$O_k = v_{ok} \sum_{j=1}^{q} v_{jk} \tanh\left(w_{oj} + \sum_{i=1}^{p} w_{ij} x_i\right) \tag{1}$$

Where o_k is the output of the kth output unit, v_{jk} and w_{oj} are the network weights, p is the number of network inputs, and q is the number of hidden units. During the training process, weights are adjusted in such a way that the difference between the obtained outputs o_k and the desired outputs d_k is minimized, which is usually done by minimizing the following error function:

$$E = \sum_{k=1}^{r} \sum_{e=1}^{n} \left(d_{e,k} - o_{e,k}\right)^2 \tag{2}$$

Where r is the number of network outputs and n is the number of training examples. The minimization of the error function is usually done by gradient descent methods, which have been extensively studied in the field of optimization theory [11].

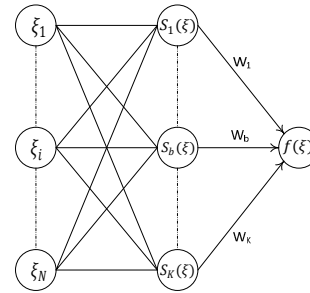### B. Radial Basis Function Neural Network

Fig. 2. Architecture graph of a RBFNN

RBFNN's have been successfully employed in many real world tasks in which they have proved to be a valuable alternative to MLPNN's since it requires less computing power and time. These tasks include chaotic time-series prediction, speech recognition, and data classification [12]. Furthermore, given a sufficient number of hidden units a RBFNN is considered as a universal approximator for any continuous functions. [13]. The construction of a RBFNN in its most basic structure (Fig.2) involves three layers with entirely different roles. The input Layer is made up of a source node that connects the network to its external environment, the second layer which is the only hidden layer in the network, applies a non linear transformation from the input space to the hidden space. In most applications the hidden space is of high dimensionality, which is directly related to the network capacity to approximate a smooth input-output mapping. The output layer is linear, supplying the response of the network to the pattern applied to the input layer [9].

If we consider a RBFNN with a single output node that computes a linear combination of the hidden units outputs, parameterized by the weights w between hidden and output layers, the function computed by the network is therefore expressed as:

$$f(\xi, w) = \sum_{b=1}^{k} W_b S_b \tag{3}$$

Where ξ is the vector applied to the input units and S_b denotes the basis function b, each of the N components of the input vector ξ feeds forward to K basis functions whose outputs are linearly combined with weights $\{w_b\}_{b=1}^{k}$ into the network output $f(\xi, w)$. The most common choice for the basis functions is the Gaussian, in this case the function computed becomes:

$$f(\xi, w) = \sum_{b=1}^{k} W_b \left(\frac{-\|\xi - m_b\|^2}{2\sigma_B^2}\right) \tag{4}$$

Where each hidden node is parameterized by two quantities: the center m in input space, that corresponds to the vector

defined by the weights between the node and the input nodes, and the width σ_B.
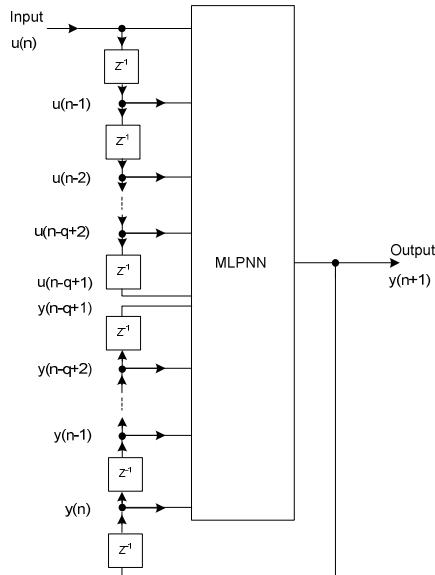
C. Recurrent Neural Network

Fig. 3. Architecture graph of the proposed RNN

Recurrent networks are neural networks with one or more feedback loops, the feedback could be of a local or global type. Starting from a MLPNN as a basic building structure, the global feedback can take various forms: like a feedback from the output neurons to the input layer or from the hidden neurons of the network to the input layers or even booth (when the MLPNN has two or more hidden layers, the possible forms of feedback are expanded). The application of feedback enables RNN to acquire state representations, which make it a suitable device for nonlinear prediction and modeling [9].

In this study, a nonlinear autoregressive with exogenous inputs (NARX) RNN is developed (Fig.3). The network has input applied to a tapped-delay-line memory of q units, with a single output that is fed back to the input via another similar tapped-delay-line memory, the content of these two memories are used to feed the input layer of the MLPNN. The present input is denoted by $u(n)$, and the corresponding output is denoted by $y(n+1)$. Thus, the signal vector applied to the input layer of the MLPNN consists of:

- Present and past values of input, namely u(n), u(n-1)….., u(n-q+1), which represent exogenous inputs originating from outside the network.

- Delayed values of the output, namely, y(n), y(n-1),…., y(n-q+1), which the model output is regressed.

The recurrent network of the Fig.3 is referred to as a nonlinear autoregressive with exogenous input, this model is described by the following relation:

$$y(n+1) = F(u(n),...,y(n-q+1),...,u(n-q+1)) \qquad (5)$$

where F is a nonlinear function of its arguments.
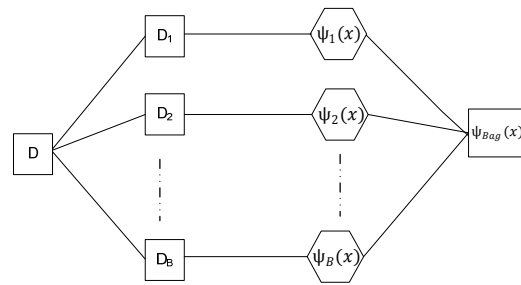
D. Neural Network Ensemble

Fig. 4. Framework of a Bagged Neural Network Ensemble

Although an ANN is providing a relevant methodology for solving several types of nonlinear problems, it is still considered as an unstable learning model [14]. In fact, the changes on training data and the architecture of the network which incorporate: the number of hidden layer, the number of neurons and the initial connections weights has an effect on the network training and predicting performance. Moreover the use of a single neural network in case of noise presence in training data set can constraint to over-fitting problems so that the training process always fail into a local optimum unstable solution with low forecasting accuracy. On the other hand, there are no systematic investigation for those issues, mostly researchers has adopted trial and error methodology to deal with these inconsistencies [15]. In this study a neural ensemble is proposed to improve the learning model performance effectively. In fact, when building bagged neural network ensemble (Fig.4), each network $\psi_b(x)$ is associated to a training set D_b that belongs to the original training set D, so that a networks ensemble $\{\Psi(x)\}_{b=1}^B$ is obtained. Each network among the ensemble will constitute a base predictor, the forecasting result of each network is aggregated basing on their average values to get the grand total predictor ensemble $\psi_{Bag}(x)$. Indeed, as the noise is varying among the ensemble networks, the averaging ensemble tends to mitigate the noise parts by retraining the fitting to the regularities of the data.

In the present study, an ensemble of bagged neural networks consisting in: MLPNN, RBFNN and RNN is proposed. The NNE provides a forecasting based on the average of the MLPNN, RNN and RBFNN outputs (Fig.5). The NNE could be also composed of similar types of neural networks, where the framework and training data set related to each network would be coherently differs from network to another.
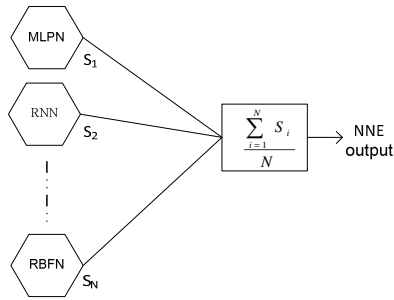
Fig. 5. Architecture of the proposed neural network ensemble

III. METHODOLOGY

Input training data are composed of hourly and daily weather parameters recorded during 2007 and 2008. Vapor pressure, humidity, cloud coverage and sunshine duration was collected by Japan Meteorological Agency in the local area of TUAT, while the maximum, minimum and average temperature, irradiation and the SPG output was recorded in TUAT University (Fig.6). Validation samples were used to avoid over fitting problems by setting up stopping points for the training process. In order to evaluate the reliability of the developed ANN's for different ranges of climatic conditions, the testing data set were split into 4 seasons namely: Winter (December, January and February), Spring (March, April, May), Summer (June, July and August) and Fall (September, October and November) where five days from each month were selected randomly for the testing of each ANN.
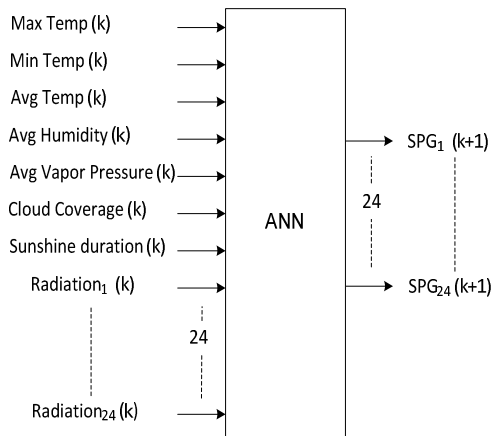

Fig. 6. Inputs and outputs for the proposed ANN's

The rescaling (normalization) of the input training data is important to improve the training convergence of an ANN [16]-[18], mean 0 and standard deviation 1 based Across Chanel Normalization [19] was used for the input training set rescaling basing on the following relations:

$$S_i = \frac{x_i - mean}{\sqrt{\dfrac{\sum\limits_{i=1}^{N}(x_i - mean)^2}{N-1}}} \qquad (6)$$

where

$$mean = \frac{\sum\limits_{i=1}^{N} x_i}{N}$$

X_i is the raw input variable X in the ith training case

S_i is the standardized value corresponding to X_i

N is the number of training case

The target variables were linearly normalized in order to force the network values to be within the range of output activation functions using upper (Y_{max}) and lower bounds (Y_{min}) for the values:

$$Z_i = \frac{Y_i - \dfrac{Y_{max} - Y_{min}}{2}}{\dfrac{Y_{max} - Y_{min}}{2}} \qquad (7)$$

where

Y_i is the raw target variable Y for the ith training case

Z_i is the standardized value corresponding to Y_i

The network framework of the MLPNN, RNN and RBFNN was set out basing on trial and error approach. In fact, the networks were trained for a fixed number of epochs, performance of the MLPNN and the RNN was evaluated by changing the number of the hidden nodes, while no significant decrease of the error was noticed above 27 hidden nodes and as referred previously, only one single hidden layer was sufficient for the proposed forecasting task. On the other hand, RBFNN was evolved to get round over fitting problem in relation to the choice of the network framework, the proposed RBFNN was build up of 300 hidden neurons.

IV. RESULTS AND DISCUSSION

Several performance criteria are reported in the ANN literature as: the training time, the modeling time and the forecasting error. In the present study, as the training process is in offline mode, the first two criteria are not considered to be relevant. Thereby, the forecasting performance will be evaluated only in term of forecasting error, defined as the difference between the actual and the forecasted values basing on statistical and graphical approaches.

Mean Average Deviation (MAD) and Mean Absolute Percentage Error (MAPE) defined respectively in (8) and (9), were applied as statistical error test criteria. While a correlation graph between the forecasted and the actual values (Fig.7), a 2-D error prediction form (Fig.8), and 24 hour SPG comparative forecasting (Fig.9) are presented as graphical error performance criteria.

$$MAD = \frac{1}{N}\sum_{1}^{N}\left|S_{d,i} - S_{f,i}\right| \qquad (8)$$

$$MAPE = \frac{1}{N}\sum_{1}^{N}\left|\frac{S_{d,i} - S_{f,i}}{S_{d,i}}\right| * 100 \qquad (9)$$

where

$S_{d,i}$ is the ith desired value (actual),

$S_{f,i}$ is the ith forecasted value,
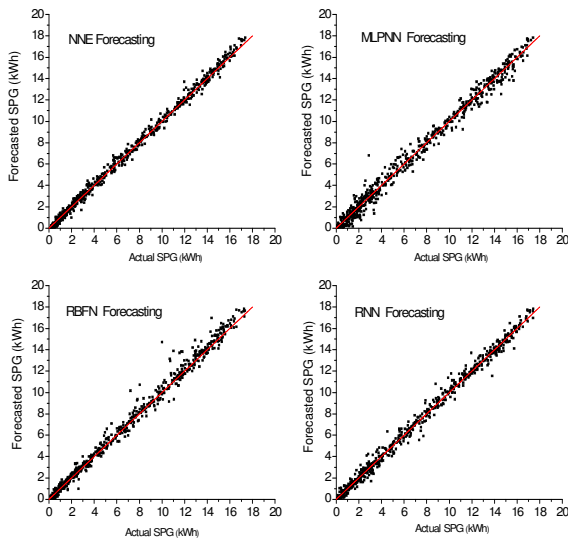
N is the total number of observations.

Fig. 7. Correlation between the actual and the forecasted SPG

Fig. 9. Prediction error for testing data set

Figure 7 presents a correlation analysis for the testing data set, so as to the four ANN's. It is clear that the proposed NNE gives the best forecasting matching with the actual data along the diagonal axis. In fact, the SPG forecasting error performance differs from network to network in relation to the current meteorological conditions (power generation). The RBFNN seems to have a narrower scatter along the matching diagonal axis then the MLPNN or RNN for low SPG levels. However for higher SPG levels, the RNN have the best accuracy. Thereby, for a larger interval of SPG level, the NNE combination based on the average forecasting provided by the
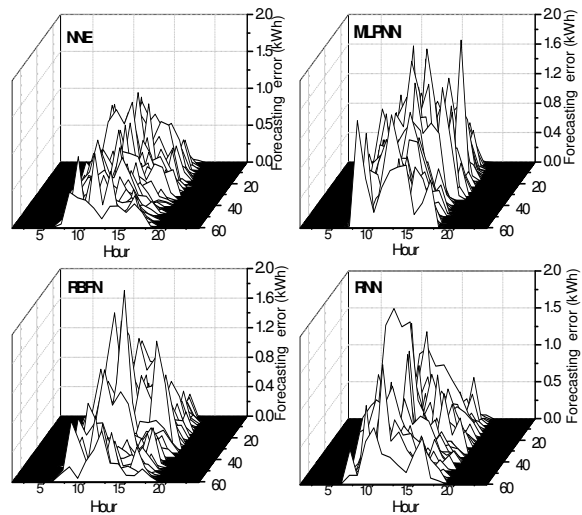
conventional ANN's. Indeed, the scatter along the whole matching diagonal line of the NNE correlation curve is the narrowest and the most stable comparing to the other ANN's.

Figure8 represents a daily 2-D prediction error graph in kWh for the proposed ANN's during sixty days, where it can be observed that the NNE has the highest prediction success with the smallest forecasting error. Figure 9 shows 24 hour SPG forecasting projected by the four developed ANN's, the forecasted days were selected randomly among each season data sample, comparing to the actual SPG, the NNE have the most accurate forecasting.
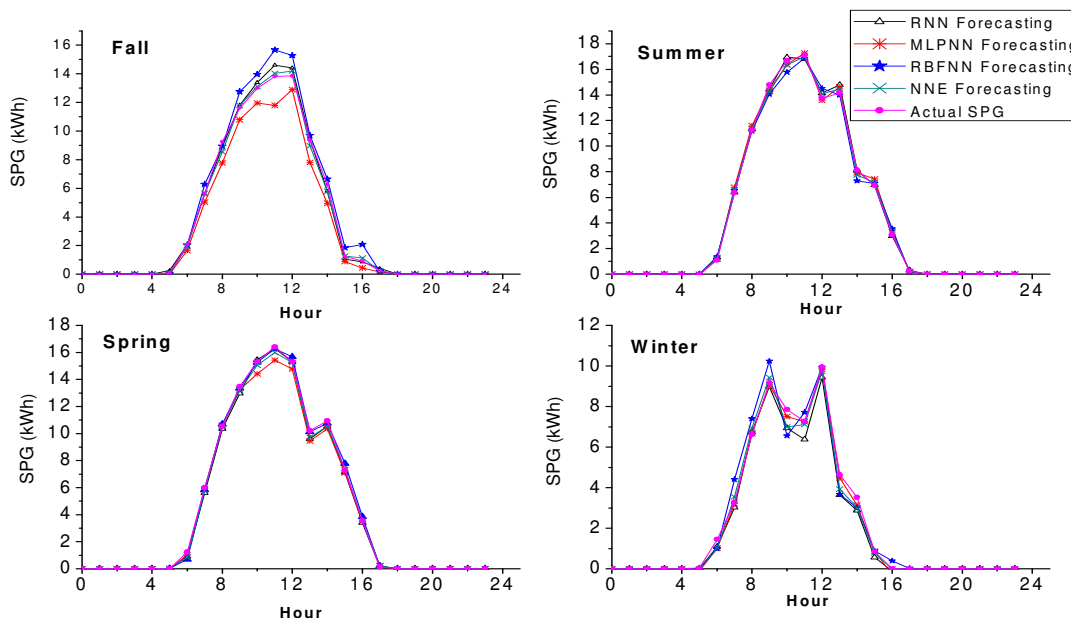

Fig. 8.Comparison of the SPG for the different proposed ANN's

The same conclusion was also carried out basing on statistical error forecasting performances shown in Table 1. In fact, referring to the MAD and MAPE criteria, we can observe that the proposed NNE has the lowest forecasting error including different testing data sets of winter, spring, summer and fall. On the other hand RBFNN, RNN and MLPNN achieved also a reasonable forecasting accuracy, among the previous conventional neural networks the RBFNN presented the best overall results, while the MLPNN achieved the lowest forecasting accuracy. In fact the RBFNN can overcome several limitations of MLPNN and RNN such as a highly non-linear weights update and the slow convergence rate.

TABLE I
STATISTICAL ERROR FOR DIFFERENT ANN'S

		RNN	RBFNN	MLPNN	NNE
Winter	MAPE	3.9832	3.5127	3.9344	2.7867
	MAD	0.2127	0.1844	0.2485	0.1680
Spring	MAPE	6.2676	5.1788	5.2498	4.1313
	MAD	0.3840	0.4557	0.4273	0.3055
Summer	MAPE	7.2161	5.7643	6.6321	4.6816
	MAD	0.4011	0.3705	0.4473	0.2870
Fall	MAPE	5.8278	4.0373	5.9481	3.6387
	MAD	0.2395	0.1730	0.2694	0.1572

V. CONCLUSION

In this study, four neural networks were developed and applied to a 24 hour ahead forecasting of solar power generation for a 20 kW photovoltaic field. The forecasting reliability was evaluated in term of forecasting error basing on graphical and statistical approaches. The experimental results showed that the NNE achieved a higher forecasting accuracy than conventional MLPNN, RBFNN and RNN. In fact, the NNE can improve the generalization and noise tolerance of learning systems effectively through aggregating numbers of neural networks with different models and diverse training data from the original source data set. The conventional neural networks fulfilled also an acceptable forecasting accuracy: in comparison, the RBFNN performed better than MLPNN and RNN while the MLPNN achieved the lowest forecasting accuracy.

Further application of the proposed ensemble will include distributed intelligent management system for the cost optimization of a MG. In fact, the knowledge of future available SPG let the system to store energy in advance or inject into the main grid, offering more flexibility to take advantage of real time electricity pricing.

REFERENCES

[1] Reddy K.S. and Manish R, ''Solar resource estimation using artificial neural networks and comparison with other correlation models,'' Energy Conversion and Management, 2003, Vol. 44, pp.2519–2530.
[2] Mohandes M, Balghonaim A, Kassas M, Rehman S, Halawani. ''Use of radial basis functions for estimating monthly mean daily solar radiation;'' Sol Energy, 2000;68(2):161–8.
[3] Mellit, A. and Kalogirou, S.A. ''Artificial intelligence techniques for photovoltaic applications: a review,'' Progress in Energy and Combustion Science, 2008, Vol. 34, pp.574–632.
[4] A. Mellit ''Artificial Intelligence technique for modeling and forecasting of solar radiation data: a review,'' International Journal of Artificial Intelligence and Soft Computing, 2008, Volume 1 , Issue 1, pp 52-76
[5] SA. Kalogirou, ''Artificial Neural Networks in Renewable Energy Systems: A Review,'' Renewable & Sustainable Energy Reviews, 2001, Vol. 5, No. 4, pp. 373-401.
[6] CHAABENE Maher, BEN AMMAR Mohsen, ''Neuro-Fuzzy Dynamic Model with Kalman Filter to Forecast Irradiance and Temperature for Solar Energy Systems,'' Renew Energy, 2008, pages 1435-1443.
[7] J.C Cao,. and, S.H. Cao ''Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis', Energy, 2006, Vol. 3, pp.13435–13445.
[8] YINGNI JIANG, ''Prediction of monthly mean daily diffuse solar radiation using artificial neural networks and comparison with other empirical models,'' Energy policy, 2008, vol.36,n10,pp.3833-3837.
[9] Simon Haykin, Neural Networks. A Comprehensive Foundation, 2nd Edition, Prentice Hall, 1999.
[10] K.M.Hornik, M. Stinchcombe, H.White, ''Multilayer Feedforward Networks are Universal Approximators,'' Neural Networks, 1989, 2(2):pp. 359-366.
[11] R. Fletcher. Practical ''Methods of Optimization,'' 2nd ed. Wiley, Chichester, 1990.
[12] Cornelius T.Leondes, Neural Network Systems Techniques and Applications, Volume 1 ofNeural Network Systems architecture and applications, Academic Press, 1998.
[13] E. J. Hartman, J. D. Keeler, and J. M. Kowalski, ''Layered neural networks with gaussian hidden units as universal approximators,'' Neural Comput, 1990, 2:210-215.
[14] ZHANG Gao, FAN Ming, ZHAO Hongling, ''Bagging Neural Networks for Predicting Water Consumption,'' Journal of Communication and Computer, 2005, Volume 2, No.3 (Serial No.4).
[15] Hansen LK, Salamon P ''Neural network ensembles,'' IEEE Trans Pattern Anal, 1990; 12(10):993-1001.
[16] D., Liew, A.C. and Chang, C.S., ''A neural network short-term load forecaster,'' Electric Power Systems Research, 1994 , 28, pp. 227–234
[17] J. Sola and J. Sevilla, ''Importance of data normalization for the application of neural networks to complex industrial problems,'' IEEE Transactions on Nuclear Science, 1997, 44(3) 1464–1468.
[18] Guoqiang Zhang, B. Eddy Patuwo and Michael Y. Hu, ''Forecasting with artificial neural networks:The state of the art,'' International Journal of Forecasting, 1998, Volume 14, Issue 1, Pages 35-62.
[19] Azoff, E.M., ''Neural Network Time Series Forecasting of Financial Markets,'' John Wiley and Sons, Chichester, 1994.