Univariate short term forecasting of solar irradiance using modified online backpropagation through time

Animesh Anant Sharma
Dept. of Electrical Engineering
IIT Roorkee
Roorkee, India
anantanimesh36@gmail.com

***Abstract*— In this paper an attempt has been made to predict the solar irradiance values for multiple look ahead time predictions with time intervals as small as fifteen minutes. The recurrent neural networks in the past have been implemented on datasets with an interval of at least 30 minutes. The recurrent neural network was trained using backpropagation through time and the prediction was done using only the past solar irradiance values. A sliding window implementation of the network was achieved and the training time for over 20000 data points was less than 5 minutes which is an improvement over the execution time of other deep learning architectures. The online form of back propagation through time was implemented with the modification that the network took into account both the past mistakes and the current direction to which it is moving. The performance of the proposed network was tested comprehensively using two years of data and it was compared with the performance of persistence model and the normal recurrent network. The modified backpropagation network outperformed the baseline models for different time intervals. The proposed network also showed improvement over results computed for dataset with 15 minutes interval which was not achieved by the earlier state of the art architectures.**

Keywords—time series forecasting; solar irradiance; sliding window; deep learning; recurrent neural networks; online backpropagation through time.

I. Introduction

The solar irradiance prediction can lead to an improvement in the power quality of electric power delivered to the consumers [1]. It can also lead to more efficient energy management in smart grid [2]. One of the approaches used for solar power prediction involves the use of artificial neural networks (ANNs). Many methodologies have been developed over the years which are based on ANNs.

Using backpropagation (BP) neural network, the solar radiation data from the past 24-h was used to predict the value for the next instance in [3]. The mean daily solar radiation data and air temperature values were used to predict future values up to 24-h and ANN was implemented in [4]. In [5], the prediction was made for three different days (sunny, partly cloudy and overcast) using ANN. Bayesian neural network was used in [6] and the input parameters that were considered were relative humidity, temperature, sunshine duration and irradiation. In [7] it was proved that wavelet based ANN can produce better results as compared to classical ANN. Apart from neural networks based networks, there have been numerous attempts at predicting solar energy. Meteorological parameters were considered as input parameters and Radial Basis Function based ANN (RBFANN) was used in [8]. Solar transmissivity was predicted in [9] using Support Vector Machines (SVM) and it outperformed RBFANN. A fuzzy based approach was proposed in [10] and the future sky conditions and temperature were considered as different fuzzy sets. Two day ahead predictions were made in [11] using wavelet based neural networks. A modified form of ANN, generalized neural network (GNN) has been proposed in [12] and its performance was compared with that of ANN and fuzzy logic based networks. In this setup, weather based parameters from several stations in India were used as input features.

Recurrent neural network has also been proposed for the prediction of solar energy. Elman neural networks were compared with adaptive neuro-fuzzy inference system (ANFIS), multi-layer perceptron (MLP) and neural network auto-regressive model with exogenous model (NNARX) in [13]. Gamma test (GT) was used to select input data and training dataset length. Both univariate and multivariate input feature models were developed in [14] with the help of recurrent neural networks trained by cooperative neuro-evolution algorithm.

The prediction with multiple look-ahead times is helpful in various operation and control activities like power scheduling and grid regulation [15]. The methodology associated with the proposed approach is that an easy to implement sliding window based implementation of BPTT has been proposed and in the testing phase, the online version has been modified to give better results for this particular case. Also the input features used here are only the previous values of solar irradiance. The novelty associated with this approach is that for the online implementation of BPTT, the method has been modified to take into account the past mistakes and the current scenario while calculating weight updates. All the implementations up until now have approached the problem with datasets containing the time interval of 30 minutes or more since it is difficult to beat the persistent model for predictions smaller than 1 hour but in this approach the time interval for the dataset of first case is 15 minutes. The training approach used here is similar to the one mentioned as truncated BPTT in [16] but the online version with this modification has never been proposed for solar irradiance prediction.

978-1-5090-4420-7/16/$31.00 ©2016 IEEE

The rest of the paper is organized as follows: Section II introduces the problem statement, then in Section III implementation of BPTT and the proposed change is discussed. In Section IV, the experimental setup is given and it is followed by the discussion on results in Section V.

II. Problem Statement

A time series: SI = [SI_1, SI_2, SI_3, SI_4, ….. , SI_n] , is given for n data points of solar irradiance. These data points have been recorded at an interval of 15 minutes, 30 minutes and 1 hour for three different experimental conditions.

The goal here is to predict the multiple look ahead time interval values for the different setup conditions using the previous irradiance values. The multiple look ahead time steps are considered in such a way that predictions are made from the range of 1-h ahead values to 5-h ahead values. In such a setup, very short term predictions can be made which are useful for PV, storage control and electricity market clearing. Also short term predictions are covered which are useful for economic dispatch and unit commitment in the context of electricity market and power system operation [17].

The implementation of BPTT is explained here and the sliding window approach is introduced.

TABLE I. Parameters For Psuedo Code

Parameter	Description
η	learning rate
E1	error calculated using least mean squared function and then E_1 is calculated by differentiating this function w.r.t the predicted value
E2	error calculated through backpropagation
E3	error calculated using BPTT in the unfolded window
W_S	number of times the network has to be unfolded
n_ep	number of iterations of training
n_ts	number of training samples
TH1	bias unit to output layer
TH2	bias unit to hidden layer
W	weight from hidden layer to output layer
V	weight from hidden layer to hidden layer
U	weight from input layer to hidden layer
S(t)	Hidden unit activation value at time instance t
I(t)	Input value at time instance t

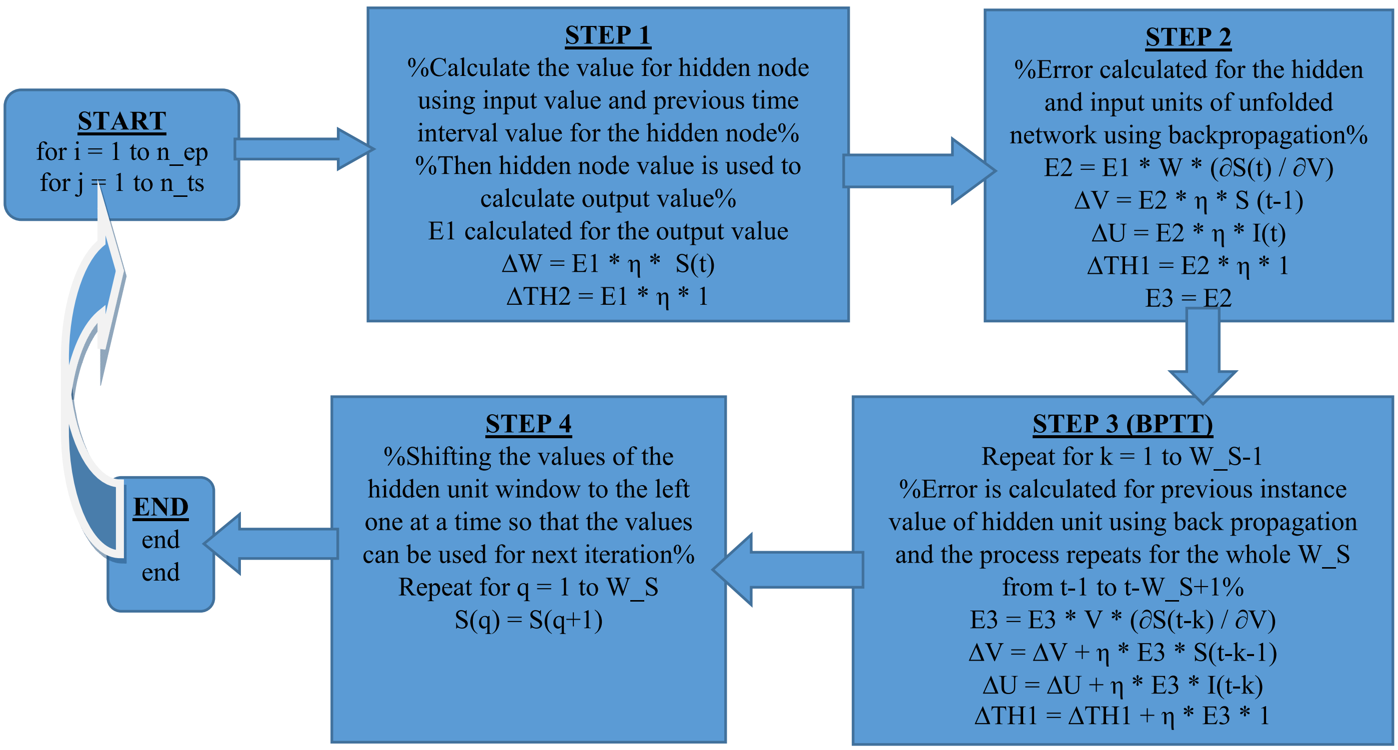

Fig. 1 Psuedo code

III. Proposed Approach

The implementation of BPTT is given first and the sliding window approach is explained along with the pseudo code for the same. Then this implementation was modified to make the network adaptive to testing dataset and the normal implementation was taken as a benchmark model for validating the claim that the adaptive network performs better than the normal network.

In fig.1 the pseudo code for the training phase is described. The same code is used for testing phase in online version with the proposed changes. In signal processing, the adaptive filters function on the same technique as that used in the training of neural networks. But the closed loop adaptive filters continue to train using the error signal as the feedback signal to refine its transfer function [18]. Similarly here in the testing phase, the predicted values are used to calculate the error signals when the target values are available. This approach might be similar to the online BPTT but the weight updating is done differently here. Since it is the testing phase and the correct value is not known, so no updating of weights is done until the desired value is available and network works according to normal BPTT until then. As soon as the target value is available, the error E1 is calculated using the difference in values of the present input and the value predicted in the past for the current time instance. In this way the past error values are taken into account and the technique is similar to online method up until now. But unlike the conventional online BPTT method, in the steps 2, 3 and 4 the errors E2 and E3 are calculated using the present hidden nodes' values or the present hidden layer window. In this way the present direction to which the network is headed is also taken into account. In the conventional online method, the whole state of the architecture is transferred back to the past instance including the hidden layer window values. So error E1 is calculated using past mistake and E2 and E3 are calculated using present values.This ideology arises from the fact that people tend to learn from their past mistakes and try to correct them while preforming the tasks at the present time instance. The updating of weights is done in testing part again but for one iteration only because the system is doing this evaluation in real time.

IV. Experimental Settings

A. Datasets

The dataset used here is publicly available at [19]. The solar energy resource data is available for 12 sites and out of these 12 sites, Elizabeth City State University, Elizabeth City, North Carolina is selected. The exact location of the site at which data is collected is: latitude given by 36.28° North and longitude given by 76.22° West. The solar irradiance is measured with the help of pyrometer and the elevation of the pyrometer is 26 m above the ground level. The unit for the solar irradiance measured is Watts per square meter (W/m^2). Global Horizontal Irradiance (GHI) is selected for estimating solar energy as it provides a good estimate for determining solar energy. The data points are available at an interval of 5 minutes and these data points are averaged over to get data values at an interval of 15 minutes, 30 minutes and 1 hour. The experiments are conducted for all the three cases and multiple look ahead time interval predictions are done. The data points are considered only from 8 AM to 4 PM for the period of January 2001 to December 2002. These data points have been divided into training set, validation set and testing set. The training set is used for training the network, the validation set is used for determining network parameters and the testing set is used for determining performance of the system.

B. Baseline Models

Two baseline models are selected for evaluating the performance of the proposed network. The performance indices are calculated for all the three baseline models and the performance of the proposed network is compared with them in each case.

B_1 is the baseline model given by the normal implementation of BPTT network. This is the model initially formulated for the problem but it was observed that there is scope for improvement and so it was taken as the baseline model.

B_2 represents the persistence model. This is a naïve predictor which is useful as benchmark model in meteorology-related forecasting [20]. This model states that the future value for the next desired time instance will be same as the latest measured value. Suppose that the time interval for which predictions are made is η and the prediction is being made for some variable p, then this model states that:

$$p_{t+\eta} = p_t \tag{1}$$

V. Results and Discussions

The values recorded from 8 AM to 4 PM were considered and the 5 minute interval values were converted by taking the average of values. The values from January 2001 to December 2001 were taken as training set, the values from January 2002 to June 2002 were taken as validation set and the values from July 2002 to December 2002 were considered for testing.

P represents the proposed model. B1 and B2 represent the two benchmark models defined earlier.

Percent improvement indicates the improvement in performance of proposed model over the benchmark models. This is calculated by comparing MAE of the models. Mean Relative error (MRE) is used so that the error can be normalized and the advantages of the proposed model can be compared with other state of the art models which were evaluated on a dataset with different range.

A. 15 min case

23360 data points were generated for this case by taking the average of the values provided in [19]. The number of hidden units were 25 in this case and predictions were made for t+4 and t+5 case. The results are tabulated for these two cases. The proposed model was able to perform well as compared to other benchmark models for look ahead predictions of time interval greater than 3 but due to space constraint, the performance indices for these two cases is tabulated.

TABLE II. T+4

Model	**MAE (W/m^2)**	**% Improvement MAE**	**MRE (%)**
P	109.41	-	9.7
B_1	112.97	3.1	10.02
B_2	110.19	0.7	9.78

TABLE III. T+5

Model	MAE (W/m^2)	% Improvement MAE	MRE (%)
P	124.82	-	11.07
B_1	127.74	2.2	11.33
B_2	126.09	1	11.18

B. 30 min case

11680 data points were generated for this case by taking the average of the values provided in [19]. The number of hidden units were 50 in this case and predictions were made for t+3, t+4 and t+5 case. The results are tabulated for these three cases. The proposed model was able to perform well as compared to other benchmark models for look ahead predictions of interval greater than 2 but due to space constraint, the performance indices for these three cases is tabulated.

TABLE IV. T+3

Model	MAE (W/m^2)	% Improvement MAE	MRE (%)
P	133.62	-	12.01
B_1	143.63	6.9	12.91
B_2	135.38	1.3	12.17

TABLE V. T+4

Model	MAE (W/m^2)	% Improvement MAE	MRE (%)
P	158.34	-	14.23
B_1	168.89	6.22	15.18
B_2	162.12	2.33	14.57

TABLE VI. T+5

Model	MAE (W/m^2)	% Improvement MAE	MRE (%)
P	177.76	-	15.98
B_1	188.31	5.6	16.93
B_2	184.56	3.6	16.6

C. 1 hr case

5840 data points were generated for this case by taking the average of the values provided in [19]. The number of hidden units were 100 in this case and predictions were made for t+2, t+3, t+4 and t+5 case. The results are tabulated for these four cases. The proposed model was able to perform well as compared to other benchmark models in multiple look ahead predictions but due to space constraint, the performance indices for these four cases is tabulated.

TABLE VII. T+2

Model	MAE (W/m^2)	% Improvement MAE	MRE (%)
P	154.30	-	14.8
B_1	161.79	4.62	15.52
B_2	155.89	1.01	14.96

TABLE VIII. T+3

Model	MAE (W/m^2)	% Improvement MAE	MRE (%)
P	188.38	-	18.07
B_1	197.96	4.8	18.99
B_2	193.81	2.8	18.59

TABLE IX. T+4

Model	MAE (W/m^2)	% Improvement MAE	MRE (%)
P	203.34	-	19.51
B_1	213.93	4.9	20.52
B_2	213.53	4.77	20.49

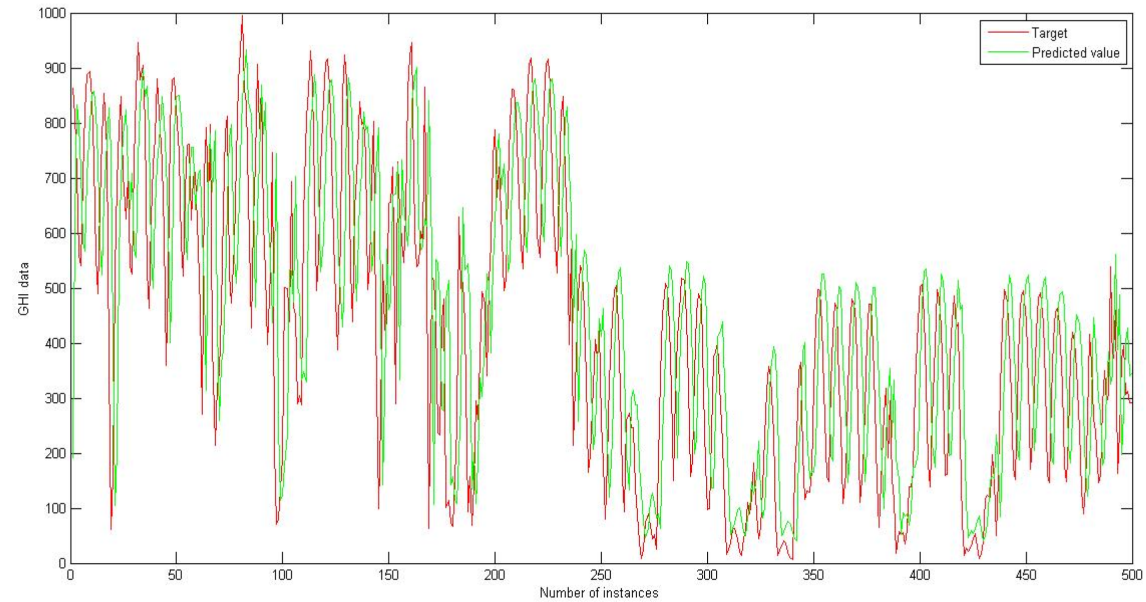

Fig. 2 Output for 1 hr case with t+2 prediction given by proposed method

TABLE X. T+5

Model	MAE (W/m^2)	% Improvement MAE	MRE (%)
P	202.64	-	19.44
B_1	212.81	4.77	20.42
B_2	211.56	4.21	20.3

D. Discussions

In all the above cases the proposed method outperformed the baseline methods and the performance index MRE indicates that the method performs well on the dataset. The performance in the case of 15 min interval is comparable to the one with MRE 7.45% (univariate model with dataset of 30 min time interval) given in [14]. So it can be said that the proposed method performs satisfactorily given that the proposed algorithm is very simple and basic and a minor modification can equal the performance of the state of the art models.

Also the training time for the case of t+4 with 15 min time interval is 220 seconds on a CPU with intel i5 processor and computed using MATLAB. This computation is done for 23360 data points. This statistic points towards a major advantage over other deep learning architecture which take hours to train and give a satisfactory result. This also implies that this online version can satisfactorily be applied in real time and multiple iterations can be run over incoming data to get even better results.

The multiple look ahead time predictions are done with just predicting increasing the time interval for the output without using any iterative approach to use the output as input n-1 times to get t+n prediction.

But as observed in the prediction of t+5 case with 1 hr interval data (in fig.5), the results were obtained with a slight shift towards left which indicates that the gradient is vanishing as is often seen in the case of BPTT. So future work can entail implementing LSTM on the same scale to get an efficient yet simple solution to multiple look ahead time predictions of solar irradiance.

VI. Conclusion

In this paper, the deep learning algorithm BPTT was implemented for the RNN. The output was studied and simple modification was suggested which helped in improving the performance of the network.

The network was able to beat the baseline models for multiple look ahead time intervals and its performance was comparable to that of the latest state of the art models proposed in this field. Also the proposed method is very suitable for implementation in real time due to the fact that the computation time is very less.

The time series prediction was done for solar irradiance for three different cases with time interval 15 min, 30 min and 1 hr. It was observed that it is difficult to beat the persistent model for small look ahead time prediction for all the time intervals but look ahead time predictions for time interval greater than 3 were done successfully for all the three cases.

A lag was observed when the look ahead predictions were done for a time interval of greater than 5. Due to this reason

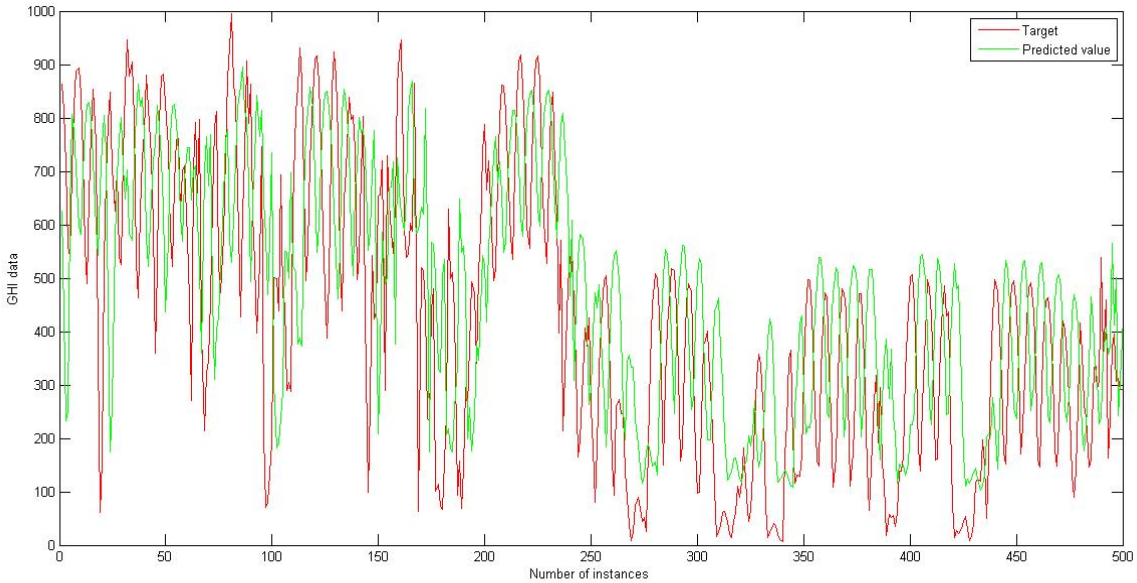

Fig. 3 Output for 1 hr case with t+5 prediction given by proposed method

future work can entail implementing LSTM along the same lines to get a more effective yet fast implementation.

REFERENCES

[1] H.-T. Yang, C.-M. Huang, Y.-C. Huang, and Y.-S. Pai, “A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output,” IEEE Transactions on Sustainable Energy, vol. 5, no. 3, pp. 917–926, Jul. 2014.

[2] D. Yang, C. Gu, Z. Dong, P. Jirutitijaroen, N. Chen, and W. M. Walsh, “Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging,” Renewable Energy, vol. 60, no. 0, pp. 235–245, Dec. 2013.

[3] Z. Wang, F. Wang, and S. Su, “Solar irradiance short-term prediction model based on BP neural network,” Energy Procedia, vol. 12, no. 0, pp. 488–494, Dec. 2011.

[4] A. Mellit and A. M. Pavan, “A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy,” Solar Energy, vol. 84, no. 5, pp. 807–821, May 2010.

[5] A. Mellit, A. Massi Pavan, and V. Lughi, “Short-term forecasting of power production in a large-scale photovoltaic plant,” Solar Energy, vol. 105, no. 0, pp. 401–413, Jul. 2014.

[6] R. Yacef, M. Benghanem, and A. Mellit, “Prediction of daily global solar irradiation data using Bayesian neural network: A comparative study,” Renewable Energy, vol. 48, no. 0, pp. 146–154, Dec. 2012.

[7] J. C. Cao and S. H. Cao, “Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis,” Energy, vol. 31, no. 15, pp. 3435–3445, Dec. 2006.

[8] M. Benghanem and A. Mellit, “Radial basis function Network-based prediction of global solar radiation data: Application for sizing of a stand-alone photovoltaic system at Al-Madinah, Saudi Arabia,” Energy, vol. 35, no. 9, pp. 3751–3762, Sep. 2010.

[9] J. Zeng and W. Qiao, “Short-term solar power prediction using a support vector machine,” Renewable Energy, vol. 52, no. 0, pp. 118–127, Apr. 2013.

[10] S. X. Chen, H. B. Gooi, and M. Q. Wang, “Solar radiation forecast based on fuzzy logic and neural networks,” Renewable Energy, vol. 60, no. 0, pp. 195–201, Dec. 2013.

[11] G. Capizzi, C. Napoli, and F. Bonanno, “Innovative second-generation wavelets construction with recurrent neural networks for solar radiation forecasting,” IEEE Transactions on Neural Networks and Learning Systems, vol. 23, no. 1, pp. 1805–1815 Nov. 2012.

[12] M. Rizwan, M. Jamil, and D. Kothari, “Generalized neural network approach for global solar energy estimation in India,” Sustainable Energy, IEEE Transactions on, vol. 3, no. 3, pp. 576–584, July 2012

[13] A. Moghaddamnia, R. Remesan, M. H. Kashani, M. Mohammadi, D. Han, and J. Piri, Comparison of LLR, MLP, Elman, NNARX and ANFIS Models—with a case study in solar radiation estimation,” Journal of Atmospheric and Solar-Terrestrial Physics, vol. 71, no. 8–9, pp. 975– 982, Jun. 2009.

[14] Mashud Rana, Rohitash Chandra and Vassilios Agelidis, “Cooperative Neuro-evolutionary Recurrent Neural Networks for Solar Power Prediction”, Proceedings of the IEEE Congress on Evolutionary Computation, Vancouver, July 2016, In Press.

[15] Elman, J. L.. “Finding structure in time,” Cognitive Science, 14:179–211, 1990.

[16] A. Sfetsos and A. H. Coonick, “Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques,” Solar Energy, vol. 68, no. 2, pp. 169–178, Feb. 2000.

[17] C. Wan, J. Zhao, Y. Song, Z. Xu, J. Lin, and Z. Hu, “Photovoltaic and solar power forecasting for smart grid energy management,” CSEE J. Power Energy Syst., vol. 1, no. 4, pp. 38–46, 2015.

[18] J.J. Shynk, “Frequency-domain and multirate adaptive filtering,” IEEE Signal Processing Mag., vol. 9, no. 1, pp. 14–37, Jan. 1992.

[19] Cooperative Networks For Renewable Resource Measurements (CONFRRM) solar energy resource data. Available: http://rredc.nrel.gov/solar/new_data/confrrm/

[20] S. D. Campbell and F. X. Diebold, “Weather forecasting for weather derivatives,” Journal of the American Statistical Association, vol. 100, no. 469, pp. 6–16, Mar. 2005.