Scientific Research and Essays Vol. 7(14), pp. 1521-1526, 16 April, 2012
Available online at http://www.academicjournals.org/SRE
DOI: 10.5897/SRE11.2214
ISSN 1992-2248 ©2012 Academic Journals

Full Length Research Paper

Recent trends on artificial neural networks for prediction of wind energy

Yasmin Zahra Jafri[1], Amir Waseem[2]*, Lalarukh Kamal[3], S. Masood Raza[4] and Muhammad Sami[5]

[1]Department of Statistics, University of Balochistan, Quetta, Pakistan.
[2]Department of Bio Sciences, COMSATS Institute of Information Technology, Islamabad, Pakistan.
[3]Department of Mathematics, University of Balochistan, Quetta, Pakistan.
[4]Department of Physics, Federal Urdu University of Arts, Science and Technology, Gulshan-e-Iqbal Campus, Karachi, Pakistan.
[5]Emerging Sector Wing, Ministry of Production, Islamabad, Pakistan.

Accepted 30 March, 2012

A variety of Artificial Neural Network models for prediction of hourly wind speed (which a few hours in advance is required to ensure efficient utilization of wind energy systems) is studied and the results are compared. Results in terms of simulation and prediction are obtained with Feed Forward Back Propagation Neural Networks (FFBPNN) which shows its performance better than other neural networks. Empirical relationship is developed which shows the Gaussian profile for the number of neurons which varies with lag inputs, that is, $n_n = k \exp(-i_l^2)$ where n_n shows the number of neurons, i_l the lag inputs, and k the sloping ratio. Feed Forward Neural Networks (FFNNs) can be corrected with optimization of our suggested relationship for simulators followed by back propagation technique.

Key words: Artificial Neural Network, McCulloch-Pitts neurons, Feed Forward Back Propagation Neural Networks, Empirical relationship for neurons, Markov Transition Matrix, Artificial Neural Fuzzy Information System.

INTRODUCTION

Stochastic time series model such as ARMA (p, q), non-seasonal ARIMA and seasonal ARIMA (SARIMA) models were developed previously to simulate and forecast hourly averaged wind speed and average annual and monthly rate of dust fall sequences on five year data. Stochastic Time Series Models take into consideration numerous fundamental features of wind rate including autocorrelation, non-Gaussian distribution and non-stationary (Sami et al., 2012; Jafri et al., 2012a, 2012b, 2012c). A critical analysis with emphasis on logical deduction, pedagogic approach, intuition and rationale pertaining to what have been over emphasized and left unattended in all existing literature on NN was made.

The McCulloch-Pitts neuron is the earliest artificial neuron described with fixed weights, a threshold activation function and a fixed discrete (non – zero) time step for the transmission of a signal from one neuron to the next McCullouch and Pitts (1943). A processing unit is termed as a neuron or node. The ANN is an information processing paradigm that is inspired by the biological nervous system such as the brain and the processing information. A biological neuron has three types of components which are of particular interest in understanding an artificial neuron: its dendrites, soma and axon. The dendrites receive signals from neighboring

*Corresponding author. E-mail: waseemq2000@hotmail.com.

Abbreviations: ANN, Artificial neural network; **FFBPNN,** feed forward back propagation neural networks; **MTM,** markov transition matrix; **ANFIS,** artificial neural fuzzy information system.

neurons. The signals are electric impulses that are transmitted across a synaptic gap by means of a chemical process. The synapse is a connection amongst neurons where their membranes almost touch and signals are transmitted from one to the other by chemical neurotransmitters. The soma or cell body sums the incoming signals, fixes signals when sufficient input is received and transmits signals over its axons to other cells. The axon is a long fiber over which a biological neuron transmits its output signals to other neurons. Neural networks are computer algorithms following the information processing exactly in the same manner as in the nervous system. The future is predicted from learning of the past with appropriate simulation. The neural network representative data is gathered and training algorithms are invoked to automatically learn the structure of data. Networks ranging from simple Boolean networks (perceptron), to complex self-organizing networks (Kohonen Networks), to networks modeling thermodynamic properties (Boltzmann machines) (Haykins, 1994) exist. There are nearly as many training methods as there are network types but some of the more popular ones include back propagation, the delta rule and Kohonen learning. A standard network architecture consists of several "layers" of neurons.

An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological system involves adjustments to the synaptic connections that exist between the neurons. This is true of ANN's as well. We shall emphasize only on ANN simulations which appear to be a recent development. This discipline of knowledge was established before the advent of computers. Many important advances in ANNs reported during five decades since its discovery in 1943, recently, the NN enjoy resurgence of interest and have the modeling of complex and non-linear phenomena (Candill and Butler, 1993); Very few relatively recent literature emerged on NN modeling for wind speed prediction (Cadenas et al., 2007, 2009; Piers et al., 2010; Zhang et al., 2012). NN is particularly useful when problems are driven rather by data than by concept or theory. To date NNs have yielded many successful applications in areas, as diverse as finance, medicine, engineering, geology, and physics indeed. ANN models have been applied to problems (Andrawis et al., 2011) involving runoff forecasting and weather predictions (Kang et al., 1993). ANNs have been applied to groundwater reclamation problems (Ranjethan and Eheart, 1993), predicting average air temperature Cook and Wolfe (1991), and for forecasting of price increments (Castiglioue, 2002) and predicting indoor air temperature in modern building (Kemajou et al., 2012). Prediction of time series is an important application of NNs. Since 1995 the time series prediction by NNs have been exhaustively studied, Detecting trends and patterns in financial data is of great interest to the business world to support the decision making process through time series forecasting, that is, with neural networks (Lin et al., 1995). Generally wind speed is a highly non-linear phenomenon (Kamal and Jafri, 1997, 1999). ANNs which are trained on a time series are supported with Artificial Neuro Fuzzy Information System (ANFIS) to achieve firstly to predict the time series many time steps ahead and secondly to learn the rule which has produced.

The prediction and learning are not necessarily related to each other especially for chaotic time series.

Theory on time series forecasting with Neural Networks

A time series is a sequence of vector, x (t), t = 0, 1, ------, where't' represents elapsed time. NN has been widely used as time series forecaster: most often these are feed forward networks which employ a sliding window over the input sequence (Patterson et al., 1993). Feed forward Neural Networks (FFNNs) are composed of layers in which the input layer of neurons is connected to the output layer of neurons through one or more layers of intermediate neurons. The training process of the neural network involves adjusting the weight till a desired input/output relationship is obtained. The majority of adaptation learning algorithms are based on the Widrow-Hoff back propagation algorithm (Widrow and Winter, 1988). The neural network forecast can be described as follows:

$$z_{k+1} = NN\ (z_k, z_{k-1}, \ldots z_{k-l}, e_k, e_{k-1}, \ldots\ldots\ldots\ldots\ e_{k-d}) \text{----} \quad (1)$$

Where z is either original observation or processed data, and e_{k1} ..... e_{k-d} are residuals. Usually, a three-layered feed forward neural network is accomplished with the Quick Prop algorithm (Goodman et al., 1993). All simulations pertaining to ANN have been done on an ULTRIX4.2 in the past. ULTRIX4.2 is an old version of algorithm especially for NN simulations. A feed forward multilayer neural network consists of an input layer and output layer with some number of input and outputs neurons, respectively. There are also one or more hidden layers in between the output and the input layers with some number of neurons on each. The values x_i of the ith neuron of the first hidden layer is given by:

$$x_i = f\ (\Sigma_i\ w_{ji}\ p_j + b_i) \text{-----------------} \quad (2)$$

Where w_{ji} are weights connecting the jth input neuron (whose value is p_j) to the i^{th} hidden neuron whose activation threshold (basis) is b_i and f is the smooth bounded and non linear function called activation function. A similar rule applies to the neuron: for the output layer. The network learns the relationship between the input-output in the training set via NN training in which the weights are modified until a prescribed error criterion is fulfilled. The term $(\Sigma_i\ w_{ji}\ p_j + b_i)$ in Equation (2)

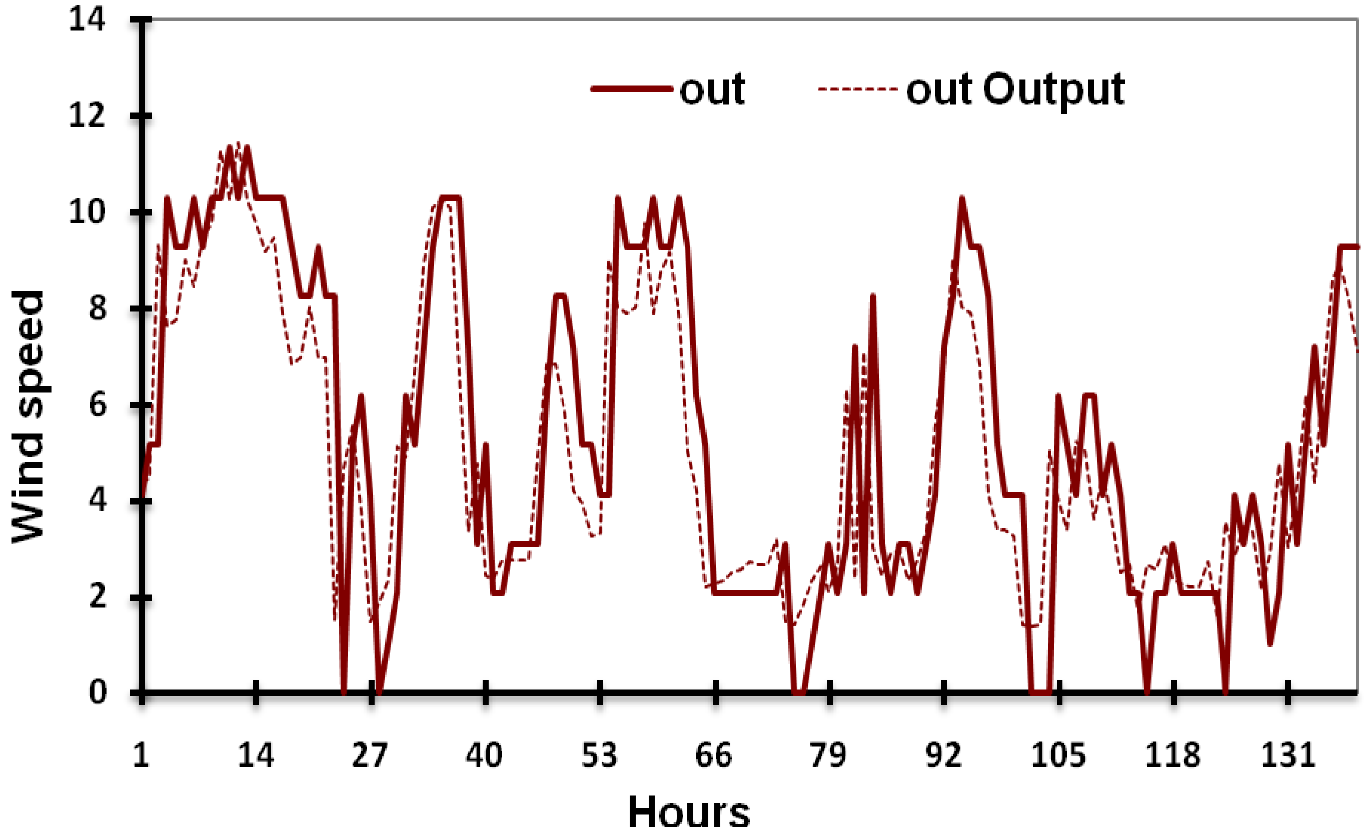

Figure 1. NN prediction modeling of wind data with two lag inputs 1, 24 and one hidden layer (March, 2004).

is known as internal activation level and produced by every neuron in the network as the sum of its weighted inputs. The threshold b_i is treated as a normal weight with input clamped at -1. The learning algorithm is the back-propagation which is a supervision iteration training method for multilayer feed forward nets, with sigmoidal non linear threshold units.

For adjusting the weight (w_{ij}) from the jth input unit to the ith output, the derivative of the error function with respect to any weight in the network is computed as shown as follows:

$$\Delta\ \mathrm{wji} = -k \frac{\delta E}{\delta \mathrm{wji}} = w_{ji}(t+1) - w_{ji}(t) = \eta\, \delta_i\, x_1 + \alpha\, (w_{ji}(t) - w_{ji}(t-1)) \text{ -------} \quad (3)$$

where $w_{ji}(t)$ is the weight from hidden node j or from an input node at time t, x1 is either the output of node i or is an input, η and α are the learning rate (gain) and momentum of the net, respectively and δ_i is the error term for node í. The error function that the back propagation algorithm minimizes is the average of the square difference between the output of each neuron in the output layer and the desired output, that is,

$$E = \frac{1}{2p} \quad \Sigma_p\, \Sigma_k \quad (d_{pk} - o_{pk})^2 \text{ -------} \quad (4)$$

where p is the index of the ‘p’ training pair of vectors, k is the index of elements in the output vector, d_{pk} is the k^{th} element of the p^{th} desired pattern vector and o_{pk} is the k^{th} element of the output vector when pattern p is presented as an input to the network. Detailed analysis and description of the multilayer feed forward neural networks and the back propagation algorithms may be found (Haykins, 1994).

RESULTS AND DISCUSSION

Figures 1 and 2 show the NN prediction modeling of wind data in terms of desired output and actual network output (simulated results) for the months of March and August, 2004 respectively with lag inputs 1 and 24 and one hidden layer Figures 3 and 4 show the same NN prediction modeling with lag inputs 1, 2, 3, 4, 5, 24 and one hidden layer. The results are obtained on trainee data. We also checked the testing data, too. With 24 lag inputs (1, 2, 3… 24) we, however, get a relatively better

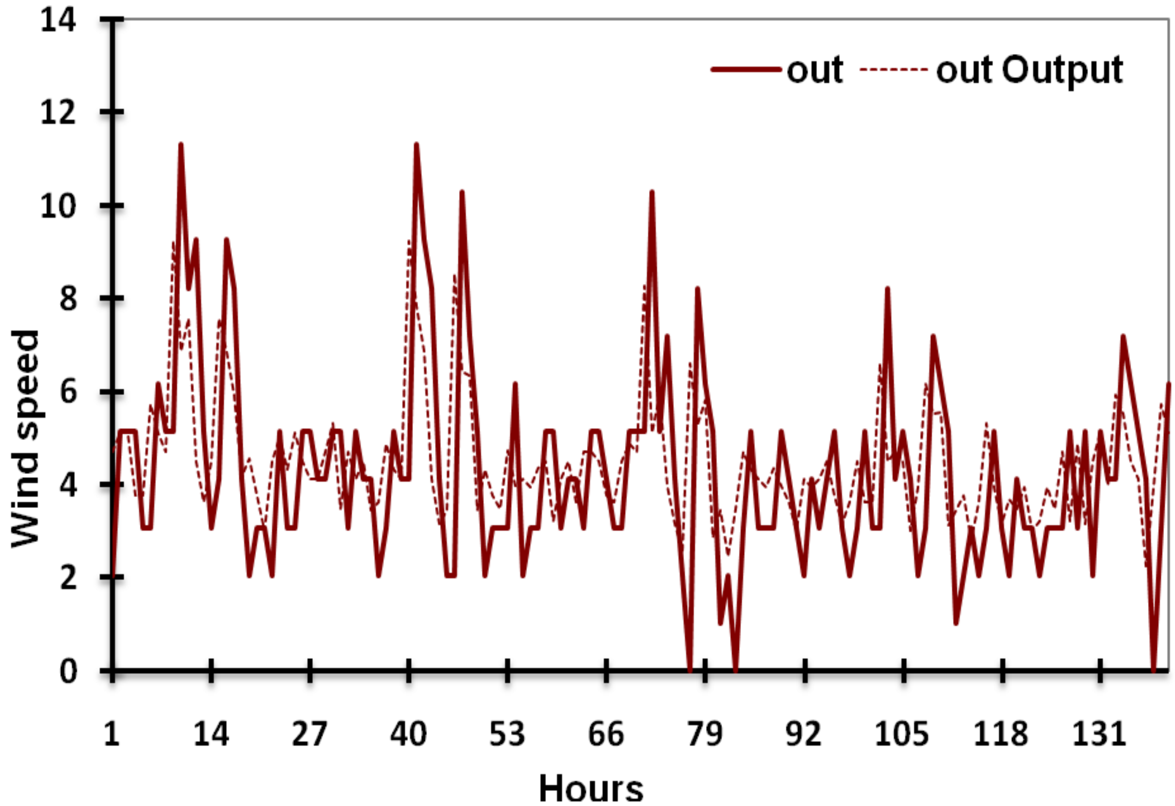

Figure 2. NN prediction modeling of wind data with two lag inputs 1, 24 and one hidden layer (August, 2004).

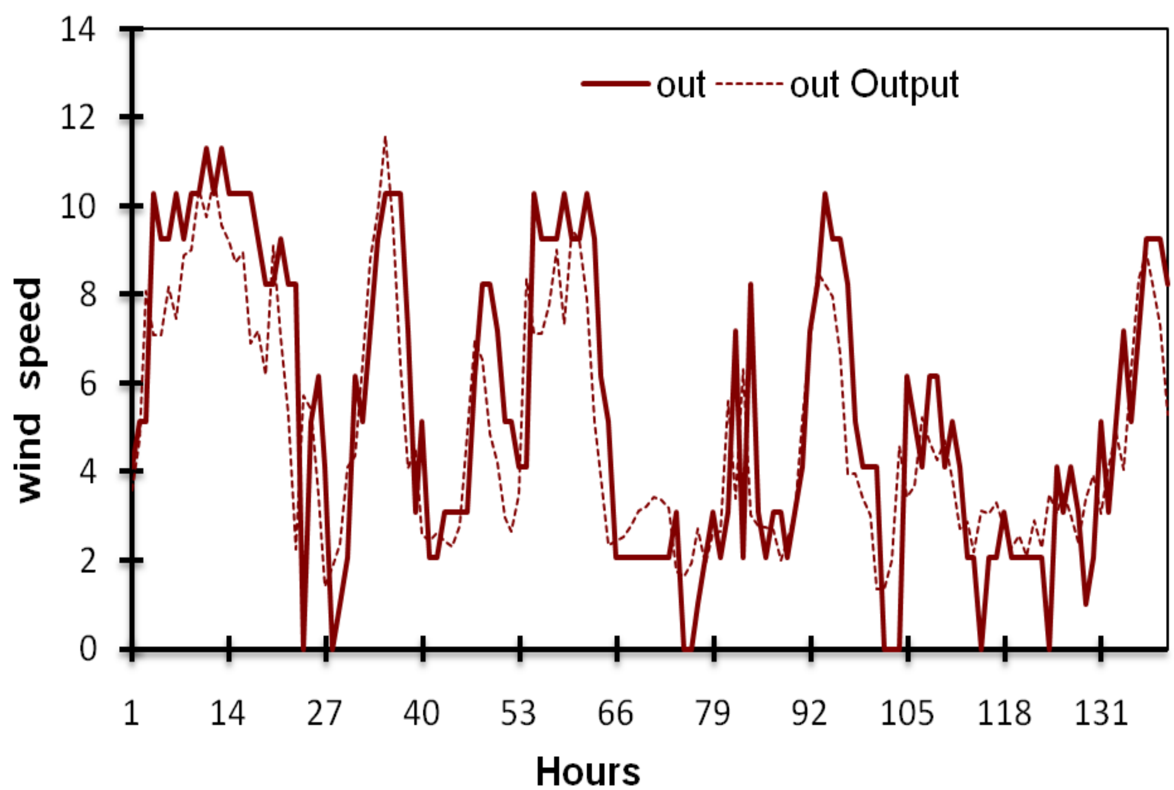

Figure 3. NN prediction modeling of wind data with six lag inputs 1, 2, 3, 4, 24 and one hidden layer (March, 2004).

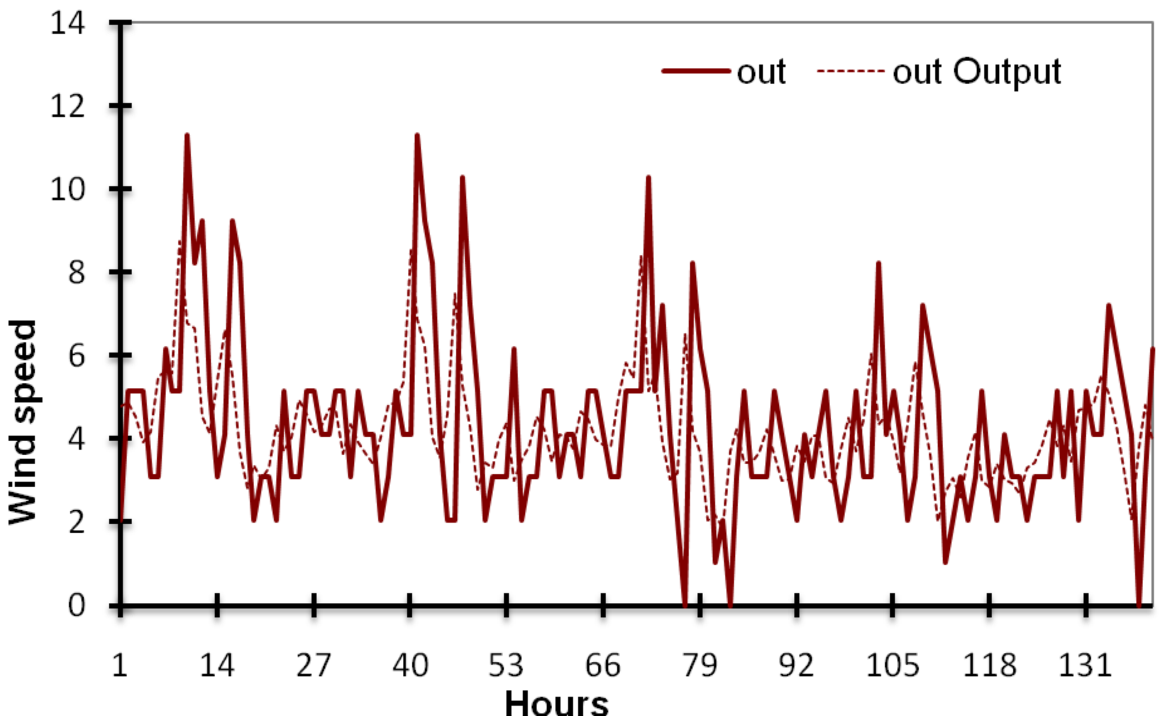

Figure 4. NN prediction modeling of wind data with six lag inputs 1, 2, 3, 4, 24 and one hidden layer (August, 2004).

simulation of wind data only for March and August, 2004. This confirms restriction of simulation for a large data set. Prediction for remaining months of 2004 and hourly averaged monthly wind data of 2003 and 2004 are, of course, recorded but simulations are badly affected. We used 720 or 744 values of wind data. We considered 500 values for trainee (69.4%), 80 values for cross validation (11.1%) and 140 values for testing (19.4%). With appropriate choices of trainee, cross validation and testing data but following the standards set by NN algorithm for multi-step ahead forecasts (Hill et al., 1996), the predicted values did not increase more than 250 values.

We used Neuro-solution, Version 5.0 (Copyright 1997-2005, Neuro-dimension, Inc.) which worked with feed forward back propagation neural networks algorithms and observed the same behavior.

We then tested the parameters of the algorithm such as the relationship between the number of neurons and the lag inputs in FFNN. On the basis of observations we established the empirical relationship, that is,

$$n_n \propto \exp(-i_l^2)$$

$$n_n = (n_1^{/}/ n_1) \exp(-i_l^2) = k \exp(-i_l^2) \text{ ---------------} \quad (5)$$

Where n_n shows the number of neurons, i_l the lag inputs, $n_1^{/}$ the data value of neuron at lag1, n_1 the extrapolated values of the number of neurons at lag1 and k the sloping factor or ratio. Equation (5) shows the Gaussian profile for the number of neurons which varies with lag inputs. With minimum number of hidden layers, the numbers of neurons are relatively larger. The neurons are the neurotransmitters for either prediction or neural network modeling; therefore, its appropriate number would make a significant affect on prediction modeling. With increasing numbers of lag inputs, the number of neurons also decreases. This decrease is more pronounced with increasing number of hidden layers. This is perhaps a severe limitation in neural network modeling.

We observe in our study, that the mean square error is relatively small for two number of lag inputs as compared to six (Table 1). Monte Carlo simulation can resolve anomalies in the relatively unreliable values of mean square error, coefficient of determination and efficiency coefficient for remaining months of the year 2004 other than March and August. We conjecture that there is a need to optimize our empirical relationship (Equation 5) in neural network back-propagation algorithm especially pertaining to simulation.

From the above discussion we infer the following conclusions:

1. There is a need to correct feed forward back propagation neural network (FFBPNN) algorithm to accommodate more predicted values as a simulator.
2. The number of neurons should increase following the geometric progression with increasing lag inputs. Neuron

Table 1. Statistical analysis.

Date	No of lags inputs, i_l	Hidden layer	Mean square Error (MSE)	Coefficient of determination R^2 (%)
March 2004	Two: 1,24	One	4.0283809	98
	Six: 1,2,3,4,5 and 24	-	4.165101044	87
August 2004	Two: 1,24	One	4.080314087	96
	Six: 1,2,3,4,5 and 24	-	4.520709573	85

absorption with hidden layers follows the Gaussian profile with a negative slope. There should remain a parallel competition between neuron increase or breed and neuron absorption with hidden layers having both positive and negative slopes of Gaussians profiles.

3. An empirical relationship is developed which shows the Gaussian profile for the number of neurons varying with lag inputs.

4. NN modeling is not suitable for chaotic data like wind speed even with application of ANFIS because it shows randomness and non-stationarity as is evidenced in previous studies (Kamal et al., 1997).

5. Neurons if they process the chaotic data will themselves become confused neurotransmitters

REFERENCES

Andrawis RR, Atiya AF, El-Shishiny H (2011). Forecast combinations of computational intelligence and linear models for the NN5 time series forecasting competition. Int. J. Forecast., 27: 672-688.

Cadenas E, Rivera W (2007). Wind speed forecasting in the south Coast of Oaxaca Mexico. Ren. Ene., 32: 2116-2128.

Cadenas E, Rivera W (2009). Short term wind speed forecasting in La Venta, Oaxaca, Maxico, Using artifical neural network. Ren. Ene., 34(1): 274-278.

Candill M, Butler C (1993). "Understanding Neural Networks: Computer Explorations." MIT press, Cambridge.

Castiglioue F (2000). Forecasting Price Increments, Adv. Complex Syst., 1: 1-12.

Cook DF, Wolfe ML (1991). A back-propagation neural network to predict average air temperature, A1 Appl., 5: 40-46.

Goodman P, Rosen D, Plummer A (1993), NevProp algorithm, University of Nevada, USA.

Haykins S (1994). "Neural Network: A comprehensive foundation", McMillan, New York.

Hill T, O'Connor M, Remus W (1996). Neural network models for time series forecasts. Manage. Sci., 42, 1082-1092.

Jafri YZ, Sami M, Waseem A, Murtaza G, Akbar S (2012a). Stochastic approaches for time series forecasting of rate of dust fall: a case study of North West of Balochistan, Pakistan. Int. J. Phys. Sci., 7(4): 676-686.

Jafri YZ, Kamal L, Waseem A, Raza SM (2012b). Rule Based Fuzzy Logic (FL) Time Series Prediction. Int. J. Phys. Sci., 7(11): 1862-1873.

Jafri YZ, Kamal L, Raza SM, Shah SH, Sami M, Waseem A (2012c). Chaotic Time Series Prediction and Mackey-Glass Simulation with Fuzzy Logic. Int. J. Phys. Sci. (in press 2012).

Kamal L, Jafri YZ (1997). Time Series Models to Simulate and Forecast Hourly Averaged Wind in Quetta, Pakistan. Solar Energy, 61: 23-32.

Kamal L, Jafri YZ (1999). Stochastic Modeling and Generation of Synthetic Sequences of Hourly Global Solar Irradiation at Quetta, Pakistan. Renew. Energ,.18: 565-572.

Kang KW, Park CY, Kim JH (1993). Neural networks and its application to rainfall-runoff forecasting, Korean J. Hyd. Sci., 4: 1-9.

Kemajou A, Mba L, Meukam P. (2012). Application of Artificial Neural Network for Predicting the Indoor Air Temperature in Modern Building in Humid Region. British J. App. Sci. Technol., 2(1): 23-34.

Lin F, Yu XH, Gregor S, Irons R (1995). Time Series forecasting with neural networks. Complex. Intern., 7(2): 375-388.

McCullouch WS, Pitts W (1943). A logical calculus of the Ideas imminent in Nervous Activity", Bull. Math. Biophys., 5: 115-133.

Ranjethan S, Eheart JW (1993). Neural Network based screening for ground water reclamation under uncertainty. Water Resour. Res., 29(3): 563-574.

Sami M, Waseem A, Jafri YZ, Shah SH, Khan MA, Akbar S, Siddiqui MA, Murtaza G. (2012). Prediction of the rate of dust falls in Quetta city by using seasonal ARIMA (SARIMA) modeling. Int. J. Phys. Sci., 7(10): 1713-1725.

Widrow B, Winter R (1988). Neural Net for adaptive filtering and adaptive pattern recognition. Computer, 12: 25-39.

Zhang CL (2012). The wind speed prediction based on AR model and BP neural network, Adv. Mater. Res., 450-451: 1593-1596.