Short Term Wind Forecasting Using Logistic Regression Driven Hypothesis in Artificial Neural Network

Sheshnag Chitlur Sreenivasa
Centre for Energy & Environment
Malaviya National Institute of Technology
Jaipur 302017, India
sheshnagcs@gmail.com

Saurabh Kumar Agarwal
Computer Science and Engineering
Malaviya National Institute of Technology
Jaipur 302017, India
saurabhagarwal1331@gmail.com

Rajesh Kumar
Centre for Energy & Environment
Malaviya National Institute of Technology
Jaipur 302017, India
rkumar.ee@gmail.com

***Abstract* – The share of wind power is increasing significantly all over the world. The ever increasing wind power integration poses new issues due to its variability and volatility. Good forecasting techniques are thus important to address these challenges. In this paper, few time series forecasting models like artificial neural networks, adaptive neuro fuzzy interface systems are used for short term prediction of wind speeds and further a new hypothesis for better estimation of wind speed is proposed. The results obtained from a real world case study of a wind farm in the state of Karnataka are presented. In this experimental study, a thorough investigation is carried out, considering the results obtained from the mentioned techniques, the accuracy of the proposed model is found to be better by 13.53% than the existing techniques.**

Keywords - Wind forecasting, artificial neural networks, adaptive neuro fuzzy interface system, Time series wind prediction, Fuzzy logic

I. INTRODUCTION

Intermittency of wind flow is a huge challenge to implement wind energy as a reliable source of electricity. For a high wind penetration scenario, answers to a lot of issues such as competitive energy markets, real time grid operations, interconnection standards, ancillary service requirements and costs, power quality, stability and reliability of power system are required [1], [2].

Accurate wind prediction is an efficient method to overcome many issues. Example, for the competitive energy markets, precise wind prediction is always attractive for different reasons. To start with, appropriate incentives for agreeable market price are given on energy imbalance charges which are mainly based on the average pooled market price. Thereafter, a faithful forecast definitely helps in developing well-functioning day ahead or hour ahead markets [1], [3].

The probability of using the energy from a wind farm for ancillary services such as- supporting reactive power, nominal frequency support, support to damp power oscillation, etc., during un-utilization of wind energy from a farm due to prediction errors like as reported in [4] and results have shown that the errors due to wind forecasting might lead to higher payment to the wind farms for their increased lost opportunity cost.

Most of the methods use supervised learning algorithm and to optimize weights of neural network by gradient descent algorithms like perceptron learning, Back Propagation (BP) [10] but immature convergence to a local minima is seen to be a problem in generalization of this approach. Many heuristic based methods have also been proposed but none could compete with gradient descent in terms of speed. Artificial Neural Network (ANN) has been realized as a successful application in many typical problems [5-8]. For the wind forecasting problem with a higher degree of complexity, ANN provides good predication than other statistical methods like exponential smoothing, Autoregressive integrated moving average (ARIMA), Autoregressive moving average (ARMA) [11] etc. The hybrid neuro fuzzy model like Adaptive neuro fuzzy interface systems (ANFIS) provides improvised predication than their ANN counterparts [11]. The choice of model or the algorithm to be incorporated in the Neural Network, ANFIS is chosen according to the complexity of the problem. These are useful for solving non-linear system or system where in relation between inputs and resultant output is either not established or complex such as for a wind forecasting problem. Main component of the neural network consists of adaptability, capability, learning and generalization [10]. Along with optimal selection of parameters like learning rate, initial weights, number of membership functions, type of membership function, rules associated with the function, number of neurons, etc. major challenge is the selection of a weight updation rule driven by the training algorithm.

This paper discusses about the wind speed forecasting using ANN with the regular squared error cost function, established hybrid model ANFIS and also a new hypothesis as the cost model of the artificial neural network. The main drawback of the derivative based approach used to update the weights of the neural network might often end up in the local minima and it is hard to notice one. It is advised to run many times and in the process tune the parameters for optimum accuracy. This procedure is cumbersome and time consuming. In the hybrid neuro fuzzy model-ANFIS, the training can also happen in a hybrid manner where in the algorithm uses a combination of steepest descent and least squares method to alter the parameters in the adaptive network [13-14]. For a given fixed values of premise (nonlinear) parameters, the consequent (linear) parameters found by this approach are guaranteed to be a unique minimum point [15] which might not be a global minima. This suffices a part of the problem and ensures global minima only if the nonlinear parameters are also at the global optimum point which is again not easy to determine.

The new hypothesis based on logistic regression addresses the drawback of the previously mentioned approaches to predict wind with higher degree of accuracy. This proposed hypothesis stresses on changing a non-deterministic model into a deterministic model giving an upper hand to the derivative based approach which results in better performance compared to the regular ANN, ANFIS models. Experiments were performed and comparison has been done with the established ANN and ANFIS models and the results are tabulated. High accuracy of result supports the use this new hypothesis. As the solution domain contains a single minima, this results in a unique solution after every iteration adding to the reliability in system.

This paper is discussed as follows: Section 2 describes the different forecasting models used. Section 3 provides the wind data used in our experimental set up. Section 4 Compares and tabulates the result of considered techniques. Section 5 concludes.

II. Forecasting Models

The forecasting models that are considered for short term wind prediction are as below,

A. *Artificial Neural Network (ANN)*

ANN's are structures are inspired by the natural intelligence and the human ability to adapt by its way of thinking about complex problems by learning from experience and generalizing capabilities. It is composed of interconnected neurons in which each of the neuron is a separate processing unit which is given by Equation 1 and is represented in [16]

$$y = f[\Sigma(x_1 w_1 + x_2 w_2 + \ldots + x_n w_n) + \beta)] \quad (1)$$

Where, y is the output neuron output; x_1,x_2,..., x_n are the input values; w_1,w_2, ..., w_n are the interconnection weights; β means the bias value; the transfer function, f. The sigmoidal function is generally given by,

$$f[\Theta] = \frac{1}{1+e^{-\Theta}} \quad (2)$$

The representative neural network is shown in the Figure 1. This is the feed forward type neural network having input, hidden & output layers and having connection weights between successive layers. The random weights are fixed by the training process. Here, the objective of the training algorithm is to reduce the error or cost function E which is defined as below,

$$E = \frac{1}{2}\Sigma(y_n - t_n)^2 \quad (3)$$

Where, y_n, t_n are network output, target output respectively for the n^th output node.

BP method calculates error 'E' as per Equation 3 and dispenses it backward from the output to input nodes covering all the hidden nodes. This is performed using the steepest gradient descent convention in which change in weight is guided towards negative error gradient as given by Equation 4.

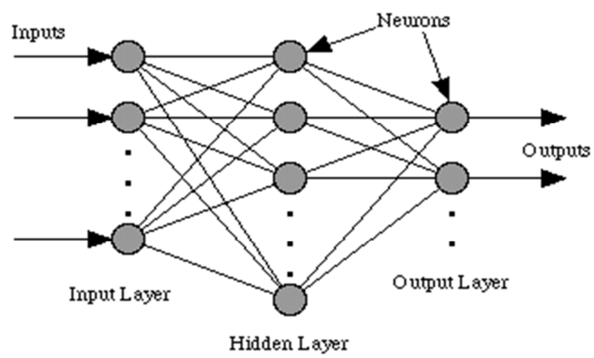

Fig 1. Multilayer neural network architecture

$$\Delta w_n = \alpha \Delta w_{n-1} - \eta \frac{\partial E}{\partial w} \quad (4)$$

Where, w is the weight between two successive nodes; $\Delta w_n, \Delta w_{n-1}$ are the weight change at the n^th and n-1^th iteration; α, η is the momentum factor and learning rate respectively. In the present study these values are varied from 0.1 to 0.9 and then one value is selected for which the convergence is reached.

B. *Adaptive Neuro Fuzzy Interface Systems (ANFIS)*

ANFIS and fuzzy interface systems are similar in nature, but the error is minimized using a back-propagation algorithm. Flavors of both ANN and Fuzzy logic (FL) have been incorporated. The input passes through the input layer and output is obtained at the output layer where in both of them are associated with their corresponding membership functions.

In this paper, Grid partitioning method [18] is considered to generate initial Fuzzy Interface system (FIS) with a random set of initial parameters. The membership functions' parameters are updated by the learning process called vector gradient method and the updated parameters are then injected to the optimization system in order to lessen the error. Generally, if y_t is present value of time, t and F_t is the prediction for the same time, then the error is,

$$e_t = y_t - F_t \quad (5)$$

ANFIS merges minimum squares method and back-propagation to estimate the activation function parameters [17]. It substantiates that ANFIS resort to the advantage of Fuzzy Logic and ANN to update the parameters and to find partial optimal solutions. Here, the model uses Univariate time series models wherein current and past data of wind speeds are used for training the model [18].

Consider that the FIS has x and y as two inputs, and f as an output. For a first order Sugeno fuzzy model, two if-then fuzzy rule sets are, [13]:

$$Rule\ 1: If\ x\ is\ A_1\ and\ y\ is\ B_1, then\ f_1 = p_1x + q_1y + r_1$$
$$Rule\ 2: If\ x\ is\ A_2\ and\ y\ is\ B_2, then\ f_2 = p_2x + q_2y + r_2$$

Let the membership grades be μA_i, μB_i. For the fuzzy set A_i , B_i. For all, i=1, 2.

Where, μ is the membership function. This can be a bell shaped or a Gaussian function. Some of the steps in classifying the rules include,

1. Evaluation of the rule bounds or premises will bring about:
$$w_i = \mu A_i(x)\ \mu B_i(y),\ i-1{,}2. \quad (6)$$

2. Evaluation of rule consequences gives:

$$f(x,y) = \frac{w_1(x,y)f_1(x,y)+w_2(x,y)f_2(x,y)}{w_1(x,y)+w_2(x,y)} \quad (7)$$

After leaving out arguments and separating into phases, defining the following:

$$\overline{w_i} = \frac{w_i}{w_1+w_2} \quad (8)$$

f can further be written as:

$$f = \overline{w_1}f_1 + \overline{w_2}f_2 \quad (9)$$

All computations described in the Equations 6, 7, 8 & 9 can be seen in Figure 2 & 3.

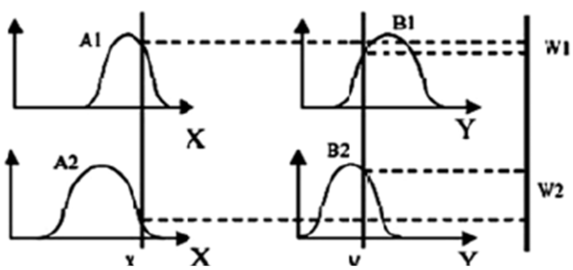

Fig 2. Takagi-Sugeno First-order Fuzzy Model.

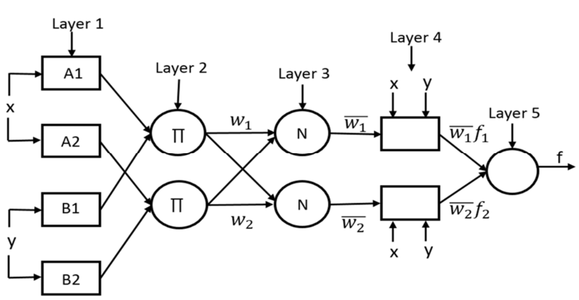

Fig 3. Two input, one output equivalent ANFIS architecture

C. Proposed model of ANN with modified cost function

The overall architecture remains same as discussed in the model 1 but the cost function has been modified. The gradient based method suffers from its premature convergence to local minima and does not promise global minima. Use of proper cost function can suppress this deficiency. An improved cost function that is based on logistic regression can be used to land up with better solution space having single minima.

Let $x^{(i)} = \{x_1, x_2 \ldots, x_{Ni}\}$ be the input pattern,

$y^{(i)} = \{y_1, y_2 \ldots, y_{Ni}\}$ be the corresponding output. Fitness function (ψ) for optimization could be formulated as follow:

$$\psi_{min}(\theta) = Min\left(\frac{1}{m}\sum_{i=1}^{m} cost\left(x^{(i)}, y^{(i)}, \theta\right) + \varphi(\theta)\right) \quad (10)$$

Where m = $\{N_i\}$, number of patterns

$\theta = \{w_1, w_2 \ldots, w_{N_0 * N_i}\}$, weight vector to be optimized.

φ = regularization term to prevent over fitting of the model.

$h_\theta x^{(i)}$= the transfer function that is used to map between successive layers. The transfer function or the activation function is given by Equation 11.

$$h_\theta x^{(i)} = \frac{1}{1+e^{-\sum_{j=1}^{m}\left(\theta_j^i x^i\right)}} \quad (11)$$

Cost function is chosen as,

$$cost\left(x^i, y^i, \theta\right) = -(y^i * log\left(h_\theta x^i\right) + \left(1-y^i\right) * log\left(1-h_\theta x^i\right)) \quad (12)$$

The regularization term φ is formulated as

$$\varphi(\theta) = \frac{\lambda}{2m}\sum_{j=1}^{m}\theta_j^2 \quad (13)$$

Where λ is a regularization term. For a neural network, an optimization method seeks to determine optimal set of weights (θ) such that the error of forecasting can be reduced. i.e., cost function could be maximizing or a minimizing function. Here the Mean square error (MSE) which is mentioned in ANN model section is not chosen as a cost function as it can have more than one minimum. Only one among all the minima will be the global minima, rest will be local minima. Cost function as given in Equation 12 is chosen to have only one minimum so that gradient based approach can be efficient. Regularization term has been added to original cost function to suppress any possibility of over fitting. The proof about this hypothesis with Mathematical and graphical support having a single minimum which is also a global one is described in [12].

Here, Back-Propagation (BP) algorithm is used for updating weights in which the error is propagated from backside i.e., from the output layer to front side i.e., to the input layer. A generic formula of the error propagation is given in Equation 14.

$$
\begin{aligned}
E_l &= \widehat{Y}_l - Y_l && if \quad l = L \\
E_l &= (\theta_l^T * E_{l-1}) * (h_\theta{}'(x))_l && if \quad l < L
\end{aligned} \tag{14}
$$

Where E_l is the error at the layer l with total layers being L. The weights are now updated as per Equation 4 with the derivative term as defined in Equation 15.

$$
\frac{\partial}{\partial \theta_l} cost(\theta, x_i, y_i) = h_\theta x(i) * E_l + \lambda \theta_l \tag{15}
$$

III. Experimental Setup

The dataset used during the experimental procedure is described in detail in this section. Wind forecasting is a time series problem and the usually historical data of wind turbine power output or wind speed will be available. In the current setup, historical wind data is considered for training and testing the model. Two years of hourly wind data is used both for the model training and testing purpose. Because of data limitation for more number of years, the first year data which is 8760 data points having all the different seasons are used for training and tuning purpose as it captures all the essential features of the climatic data. The same data is used for training all the three models and further separately discussed in this paper. The second year data is used to check the prediction accuracy of each of the models. The wind data considered is obtained from a wind farm measured at the hub height of the wind turbine from the state of Karnataka, India.

The three models-ANN, ANFIS and Proposed are compared on the basis of the best results each model can deliver. Since the configuration and data processing in each of the models is different, the optimum results obtained from each model need not use the same configuration of input-output data pair for the supervised training purpose. For instance in the experimental setup, it is found that the neural network having 10 inputs gave its best prediction for its architecture whereas ANFIS having only 3 inputs of historical data gave its best prediction.

Table 1 gives the details of the datasets like number of training, testing samples and model's architectural parameters like number of input, hidden and output nodes for the ANN and the proposed model. Number of membership functions for an input node is specifically mentioned for the ANFIS model. Because of the supervised learning and the constraint on the available data sets, the maximum number of samples that can be fed for training the model is limited to 8760-n, where n corresponds to the total input nodes.

Figure 4 pictorially represents the data being considered for training the model in which the wind speed at a given instance of time, t along with the previous n wind speeds are taken as the input for the network model and the wind speed at a future instant, t+P is taken as the target for the model training. In the experimental setup, the n value is set to 9 for ANN and proposed model but the n value is set to 2 for the ANFIS model. The P value is taken as 1 for all the three models as we are concentrating on the short term hour ahead forecasts. The ANN, proposed model has been designed to have a single hidden layer. In ANFIS model, for an input node, three Gaussian membership functions are chosen and for an output node, a linear membership function is chosen. Time shifting method is employed for the training of the networks in all the models.

TABLE 1. DATASET AND MODEL PARAMETERS

	ANN	ANFIS	Proposed
Training samples	8750	8757	8750
Testing samples	8750	8757	8750
Input nodes	10	3	10
Output nodes	1	1	1
Hidden Nodes	50	-	50
Membership Functions/ input node	-	3	-

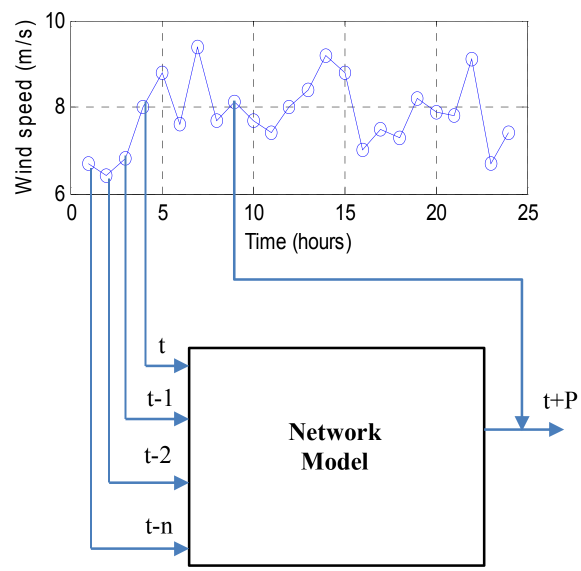

Fig 4. Input-Output mapping of the prediction model

IV. Result & Discussion

A. Evaluation of forecasting accuracy

To assess the forecasting accuracy of a given network in predicting short term wind speeds, different benchmarks are available but the ones considered in the present paper are Mean square error (MSE) and sum of squared error (SSE) which is further defined as follows,

$$
MSE = \frac{1}{n} \sum_{i=1}^{n} (\widehat{Y}_i - Y_i)^2 \tag{16}
$$

$$
SSE = 100 \sum_{i=1}^{n} (\widehat{Y}_i - Y_i)^2 \tag{17}
$$

Where, $\widehat{Y}$ is a vector of the forecasted wind is speeds and Y is a vector of the actual wind speeds.

B. Analysis & Discussion.

In this section, the results of the various models considered for the experiment has been discussed. The results are compiled in the Table 2 and it is clearly seen that the proposed model out performs ANN and ANFIS models. As per the results obtained, Proposed, ANFIS & ANN models stand in the increasing order of the ranks based on the overall performance.

TABLE 2. COMPARISON OF RESULT – MSE CRITERION

MSE		ANN	ANFIS	Proposed
Summer	Training	0.48	0.44	0.30
	Testing	0.65	0.62	0.39
Monsoon	Training	0.37	0.32	0.25
	Testing	0.51	0.35	0.30
Autumn	Training	0.29	0.22	0.19
	Testing	0.45	0.19	0.21
Winter	Training	0.49	0.34	0.29
	Testing	0.63	0.32	0.29
Annual	Training	0.43	0.35	0.27
	Testing	0.58	0.41	0.31

TABLE 3. COMPARISON OF RESULT – SSE CRITERION

√(SSE)		ANN	ANFIS	Proposed
Summer	Training	37.68	35.89	29.64
	Testing	43.63	42.61	33.79
Monsoon	Training	28.46	26.58	23.49
	Testing	33.56	27.80	25.74
Autumn	Training	20.75	17.95	16.60
	Testing	25.67	16.68	17.53
Winter	Training	32.67	27.10	25.03
	Testing	36.89	26.29	25.03
Annual	Training	61.05	55.47	48.63
	Testing	70.97	60.27	52.11

Table 2, 3 gives the MSE and SSE error values respectively for the four seasons- summer, monsoon, autumn & winter. It is seen that the error from the proposed model is the least for the summer, monsoon and winter seasons but for autumn, ANFIS model is performing better than the proposed model marginally by 4.8%. It is observed from the error figures that the accuracy of forecast is the best during the months of autumn season and the accuracy is the worst during the months of summer season. The overall testing accuracy of the proposed model considering the root of SSE criterion is found to be better than that of ANN

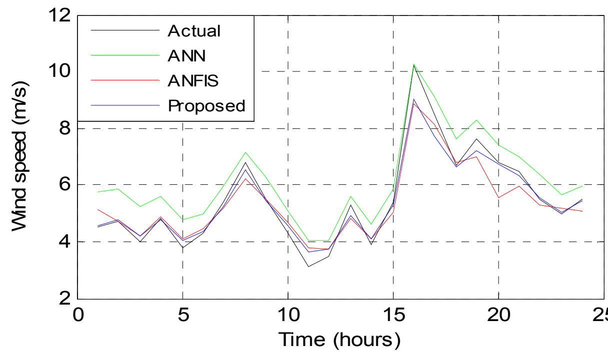

Fig 5. Actual and forecasted wind speeds of a summer day

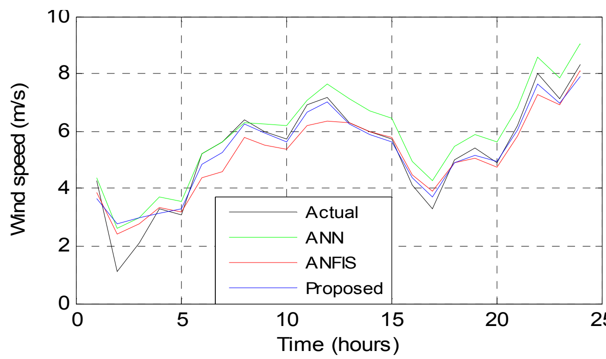

Fig 6. Actual and forecasted wind speeds of a monsoon day

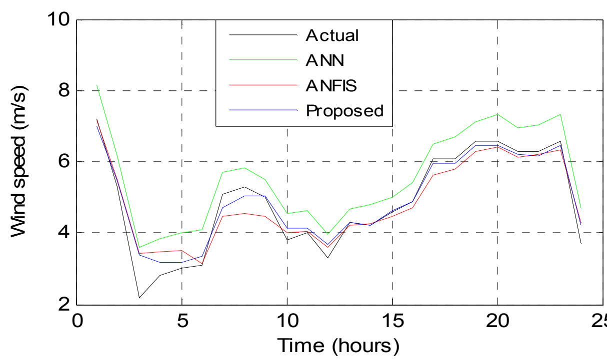

Fig 7. Actual and forecasted wind speeds of an autumn day

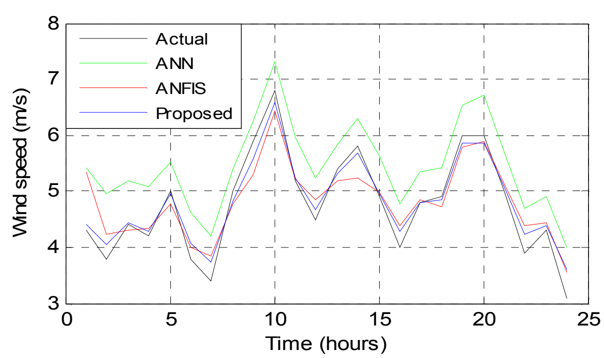

Fig 8. Actual and forecasted wind speeds of a winter day

by 34.84% and better than ANFIS by is 13.53%. This increase in the accuracy is considered to be phenomenal for the wind forecasting industry.

Figures 5, 6, 7 & 8 gives the actual and forecasted wind speeds of a day during summer, monsoon, autumn & winter seasons respectively for all the three models under study. From these figures, it can be seen that the proposed forecasting model clearly captured different features of the site wind data and can predict wind speeds for a new set of data with a better accuracy as seen in the testing error section of the Table 2.

V. Conclusion

The paper compares the proposed model with ANN, ANFIS models for the short term prediction of wind speed and is found that the proposed model outperforms the other two techniques by an accuracy of 13.53% better than its closest counterpart, in which case it is ANFIS.

This improvement in the accuracy can be a great boon for the wind industry wherein even a small improvement in the forecasting accuracy will pay off handsomely because of various reasons. The improvement in accuracy mean higher confidence on unit commitment of electricity, reducing the risk of penalty on the wind farm owners for not meeting the predicted generation, reduced risk in the competitive electricity bidding strategy. It also helps in pre-emptive maintenance of the wind turbines decreasing the loss of opportunity cost. There is a constant thrust in bettering the forecasting models and the proposed model can be very helpful in this direction. The overall accuracy can be further improved if there is an access to the large amounts of data which help in better generalization and good prediction on long term as well. Further research opportunity can be to develop an hybrid model with other techniques which can have potential to increase the accuracy further and help in the seamless assimilation of the wind farms to the electric power grid.

References

1. Yuan-Kang Wu, and Jing-Shan Hong, "*A literature review of wind forecasting technology in the world*," IEEE Power Tech, July. 2007.
2. Henrik. Lund, "*Large-scale integration of wind power into different enery systems*," Energy, vol. 30, no. 13, pp. 2402-2412, Oct. 2005.
3. M. Negnevitsky, P. Johnson, and S. Santoso, "*Short term wind power forecasting using hybrid intelligent systems*," IEEE Power Engineering Society General Meeting 2007, pp.1-4, June. 2007.
4. N.R Ullah, K. Bhattacharya, and T. Thiringer, "*Wind farms as reactive power ancillary service providers - technical and economic issues*," IEEE Transaction. Energy Conversion., vol. 24, no. 3, pp. 661-672, Sept. 2009.
5. C.M. Bishop, "*Neural Networks for Pattern Recognition*", Oxford University Press, Oxford, 1995.
6. H.A. Bourlard, N. Morgan, "Connectionist Speech Recognition: A Hybrid Approach", Kluwer Academic Publishers, Dordrecht, 1993.
7. P. Cortez, M. Rocha, J. Neves, "*Evolving time series forecasting ARMA models*", Journal of Heuristics 10 (4) (2004) 415429.
8. M. Dirst, A.S. Weigend, "Time Series Prediction: Forecasting the Future and Understanding the Past", Addison-Wesley, Reading, MA, 1994.
9. E. Alba, J.F.Chicano, "*Training Neural Networks with GA Hybrid Algorithms*", Proceedings of GECCO'04, Seattle, Washington, LNCS 3102, pp. 852-863, 2004.
10. Ernesto Cortes Perez, Airel Nunez Rodriguez et al, "*Forecast of wind speed with a backpropagation artificial neural network in the Isthmus of Tehuantepec region in the state of Oaxacana, Mexico*", ENC Marzo, Vol,22(NE-1), 2012.
11. Saurabh S. Soman, Hamidreza Zareipour et al, *"A review of wind power and wind speed forecasting methods with different time horizons"*, North American Power Symposium (NAPS), 2010.
12. Saurabh Kumar Agarwal, Rajesh Kumar, *"Explication of a logistic regression driven hypothesis to strengthen derivative approach driven classification for medical diagnosis"*, IEEE Students' Conference on Electrical, Electronics and Computer Science, 2014.
13. Jyh-Shing Roger Jang, *"ANFIS: adaptive network-based fuzzy inference systems"*, IEEE Transactions on Systems, Man and Cybernetics, vol. 23, no. 3, pp. 665-685 May/June. 1993,.
14. Jyh-Shing Roger Jang, Chuen-Tsai Sun and Eiji Mizutani, *"Neuro Fuzzy and Soft Computing, A Computational Approach to Learning and Machine Intelligence"*, New Jersey: Prentice Hall, pp. 73,74, 86, 95-97, 86-87, 26-28, 74-85, 1997
15. M. Negnevitsky, and C. W. Potter, *"Very Short-Term Wind Forecasting for Tasmanian Power Generation"*, IEEE Transactions on Power Systems, vol. 21, no. 2, pp. 965 – 972, May 2006.
16. Anurag More, M. C. Deo, *"Forecasting wind with neural networks"*, Marine Structures, Elsevier, 2003.
17. Perez, Ernesto Cortes et al, *"Performance analysis of ANFIS in short term wind speed prediction",* International Journal of Computer Science Issues (IJCSI), Vol. 9 Issue 5, pp. 94-102, Sep. 2012.
18. Fernando Castellanos, Nickel James, *"Average hourly wind speed forecasting with ANFIS"*, American Conference on Wind Engineering- San Juan, Puerto Rico, 2009.