Analyzing the probabilistic distribution of the predicted wind speed

Vigneswaran.T,
Electrical and Electronics Engineering,
Velammal Engineering College,
Chennai, India
vigneswaran1985@gmail.com

Dhivya.S
Electrical and Electronics Engineering,
Velammal Engineering College,
Chennai, India
dhivyaee@gmail.com

Abstract—**Wind power generation is characterized by its variability and uncertainty in the wind speed. Due to the irregular nature of wind power production, accurate prediction of wind speed is essential for its application in optimum power flow, transmission congestion, economic load dispatch and electricity market clearing prices. Wind speed of a wind farm is affected by conditions of the environment in which the wind farm is built, such as temperature, humidity, dew point, atmospheric pressure and wind direction. In this paper, five ANN techniques namely FFBP, CFBP, PNN, GRNN and KNN are considered to predict the wind speed. The feasibility of the proposed techniques is evaluated using the performance measures such as MSE, MAPE and linear regression and it is observed that GRNN is superior amongst the other methods that are used. Also, probability distribution of the predicted wind speed is found using Weibull probability distribution with various scale and shape factors.**

Keywords-Wind speed prediction; ANN; FFBP; CFBP; PNN; GRNN; KNN; Wind speed;Weibull probability distribution

I. INTRODUCTION

Wind energy being clean and inexhaustible is considered one of the most rapidly growing energy resources all over the world. Wind forecasting is essential for the power utility to schedule the connection and disconnection periods of wind farms or the conventional generators with the utility network for an optimal operational cost and a low environmental impact. But wind speed depends on temperature, pressure differences, and the terrain. At present, a number of different approaches have been applied to forecast wind speed produced by wind farms such as Neuro-fuzzy systems [1] and neural networks [2].Wind power prediction is also analysed using improved artificial neural network techniques such as probabilistic neural network (PNN) and other techniques [3,4].The forecast ability of medium wind speed and long term wind speed [7] can be analysed using the performance measures such as regression and mean absolute percentage error (MAPE) [5,6]. For the analysing the parameters of the predicted wind speed Weibull [11] wind speed distribution is used.

II. ARTIFICIAL NEURAL NETWORK TECHNIQUES

A. Feed Forward Backpropagation Algorithm (FFBP)

A FFBP [8] is an artificial neural network where connections between the units do not form a directed cycle. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network. This network can be used as a general function approximator. It can approximate any function with a finite number of discontinuities arbitrarily well, given sufficient neurons in the hidden layer.

A FFBP consists of two layers. The first layer, or hidden layer, has a tan-sigmoid (tan-sig) activation function, and the second layer, or output layer, has a linear activation function. Thus, the first layer limits the output to a narrow range, from which the linear layer can produce all values. The output of each layer can be represented by (1)

$$Y_{Nx1} = f\,(W_{N*M}\, X_{M,\,1} + b_{N,\,1}) \qquad (1)$$

where Y is a vector containing the output from each of the N neurons in a given layer, W is a matrix containing the weights for each of the M inputs for all N neurons, X is a vector containing the inputs, b is a vector containing the biases and f(·) is the activation function. The network was created using the neural network toolbox from Matlab 7.8.

B. Cascade Forward Backpropagation Algorithm (CFBP)

Cascade Forward Backpropagation Algorithms are similar to feed-forward networks, but include a weight connection from the input to each layer and from each layer to the successive layers. For example, a three-layer network has connections from layer 1 to layers 2, layer 2 to layer 3, and layer 1 to layer 3. The three-layer network also has connections from the input to all three layers. The additional connections might improve the speed at which the network learns the desired relationship.

C. Probabilistic Neural Networks (PNN)

Probabilistic Neural Networks [3] are used for classification problems because of ease of training and a sound statistical foundation in Bayesian estimation theory. When an input is presented, the first layer computes distances from the input vector to the training input vectors and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a complete transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 for that class and a 0 for the other classes.

PNN predicts the value of one or more dependent variables, given the value of one or more independent variables. PNN can take an input vector X of length n ($X=[x_1, x_2, \dots, x_i, \dots, x_n]$) and generates an output vector Y' of length K ($Y'=[y_1, y_2, \dots, y_k, \dots, y_K]$), where Y' denotes the prediction of the actual Y. PNN does this by comparing a new input pattern X with a set of K stored patterns X_k (K pattern units) for which the output y_k is known. In each pattern unit, a normalize Gaussian function is applied to the distance measure (Euclidian norm) between the unknown input pattern X and the training pattern X_k, which gives a measure of the distance or dissimilarity between two patterns. The property of h_k is that its magnitudes for a stored pattern X_k can be inversely relative to its distance from the input pattern X, if the distance is zero the h_k is a maximum of unity. In the summation units, one unit must calculate N, the sums of the products of H_k and associated known output y_k. Another also must compute D, the sum of all H_k. Lastly, the output unit divides N by D to produce the output Y.

D. Generalized Regression Neural Network (GRNN)

A generalized regression neural network [9] is often used for function approximation and regression problems. It has a radial basis layer and a special linear layer. The probability density function used in GRNN is the Normal Distribution. Each training sample, X_j, is used as the mean of a Normal Distribution.

$$D_j^2=(X-X_j)^T.(X-X_j) \tag{2}$$

$$Y(X)=\frac{\sum_{i=1}^{n} Y_i \exp\left[{-D_i^2}/{2\sigma^2}\right]}{\sum_{i=1}^{n} \exp\left[{-D_i^2}/{2\sigma^2}\right]} \tag{3}$$

The variables X and Y are vectors. In system identification, the dependent variable, Y, is the system output and the independent variable, X, is the system input where σ is the smoothing parameter. The distance, D_j, between the training sample and the point of prediction, is used as a measure of how well the each training sample can represent the position of prediction, X. If the Distance, D_j, between the training sample and the point of prediction is small, $\exp\left[{-D_i^2}/{2\sigma^2}\right]$ becomes big. For $D_j=0$, $\exp\left[{-D_i^2}/{2\sigma^2}\right]$ becomes one and the point of evaluation is represented best by this training sample. The distance to all the other training samples is bigger. A bigger distance, D_j , causes the term $\exp\left[{-D_i^2}/{2\sigma^2}\right]$ to become smaller and therefore the contribution of the other training samples to the prediction is relatively small. The term $Y_i \exp\left[{-D_i^2}/{2\sigma^2}\right]$ for the j^{th} training sample is the biggest one and contributes very much to the prediction.

The spread value plays a major factor in the contribution of the neurons in the GRNN network. A larger spread leads to a large area around the input vector where the radial basis layer neurons will respond with significant outputs. Therefore if spread is small the radial basis function is very steep, so that the neuron with the weight vector closest to the input will have a much larger output than other neurons. The network tends to respond with the target vector associated with the nearest design input vector. As spread becomes larger the radial basis function's slope becomes smoother and several neurons can respond to an input vector. As spread becomes larger more and more neurons contribute to the average, with the result that the network function becomes smoother.

E. K^{th} Nearest Neighbour Network (KNN)

The KNN [10] is a popular choice for many real life applications. In the KNN algorithm, the Euclidean distance is computed between the new feature vector and each feature vector from the training set. K-closest neighbours (K being the number of neighbours) are then found by analyzing the distance matrix. Euclidean distance metric can be used to compute nearness. If the output variable is a categorical variable, then the KNN algorithm takes a vote among the K-nearest neighbours and chooses the class voted for by the majority of the neighbours. If the output variable is a continuous variable, the output value is the average of the K-nearest patterns. KNN imputes missing values by the average value of the K nearest patterns, as presented in (4).

$$x_{ij}=\frac{\sum_{k=1}^{K} x_{kj}}{k} \tag{4}$$

where x_{ij} represents a missing value in the j^{th} variable of the i^{th} instance. K is the number of nearest neighbours and x_{kj} is the value of the j^{th} variable of the k^{th} nearest neighbour.

III. MODEL DEVELOPMENT

A model is developed with the data such as temperature, humidity, dew point, pressure and wind direction as inputs and the wind speed as the target using the artificial neural network techniques. To evaluate the proposed model for wind speed prediction, data sets are collected in an Automatic Weather Station and the study is carried out for 18 hours ahead. The sampled time series used in the model consists of 1000 data in total, corresponding to 30 minutes mean data. The time series is divided into two folders, one is the training set with 964 samples used for the model's training and the other is the test set that contains the rest namely 36 samples totally which is used to verify the accuracy during the prediction period.

IV. WIND PREDICTION RESULTS

A. Actual and Predicted Wind Speed Relationship

Fig. 1 to Figure 6 displays the resultant relationships between the predicted and the actual values of wind speed for different ANN techniques.

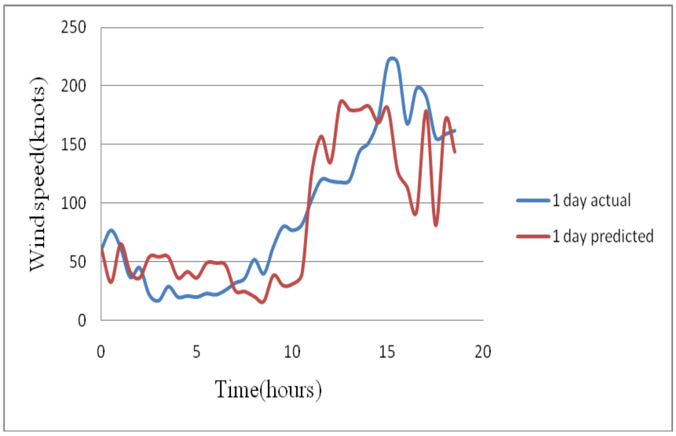

Figure 1. Wind speed vs time using FFBP

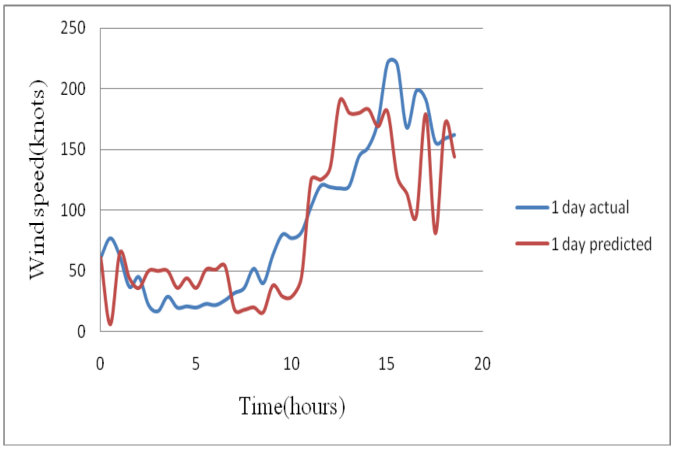

Figure 2. Wind speed vs time using CFBP

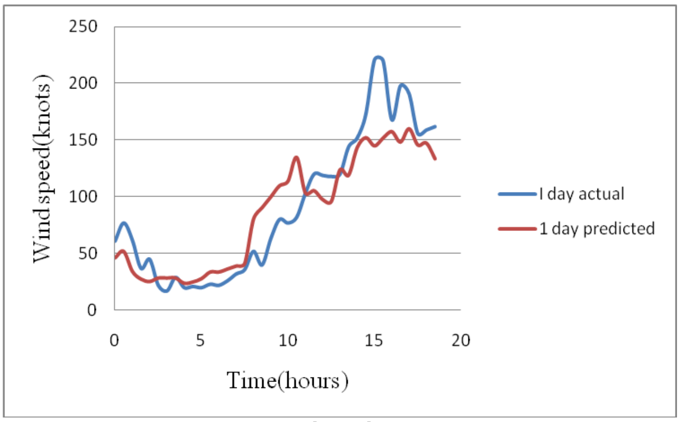

Figure 3. Wind speed vs time using PNN

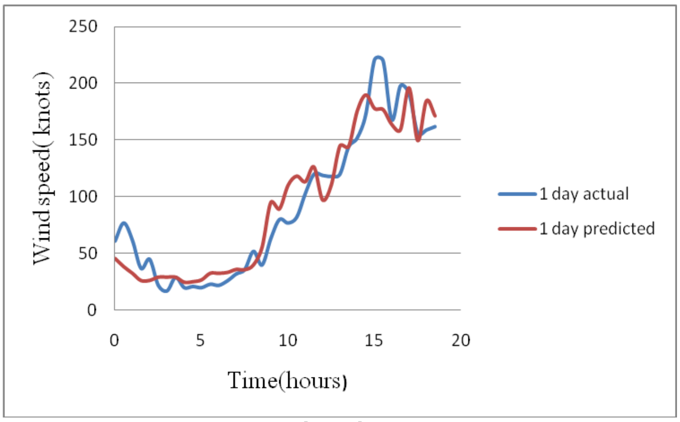

Figure 4. Wind speed vs time using GRNN

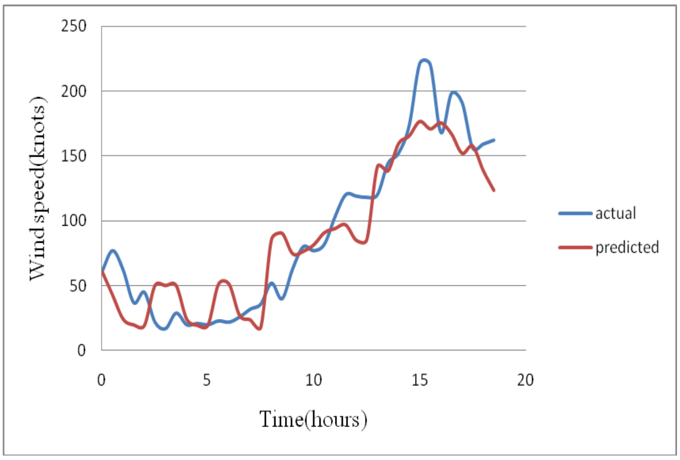

Figure 5. Wind speed vs time using KNN

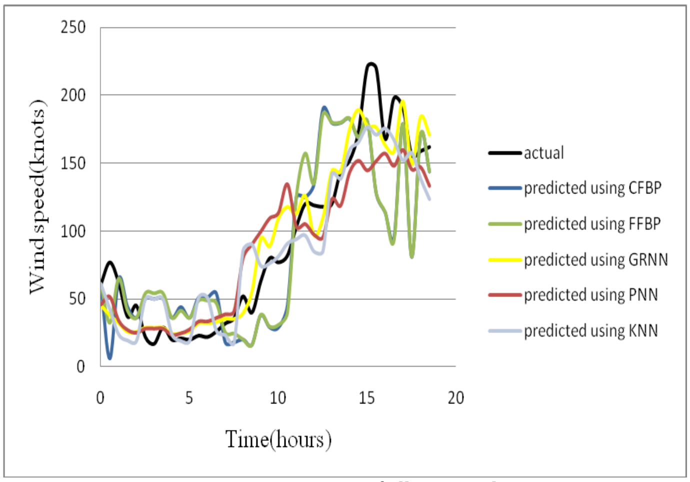

Figure 6. Comparison of all ANN techniques

The performance measures adopted are the Mean Square Error (MSE), the Mean Absolute Percentage Error (MAPE) and the linear regression. These measures can be calculated using the following (5-7):

$$\text{MSE}=\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{5}$$

$$\text{MAPE}=\frac{1}{n}\left[\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| * 100\%\right] \tag{6}$$

$$\hat{y}_i = my_i + c \tag{7}$$

where y_i is the actual wind speed value at time interval i, $\hat{y}_i$ is the predicted wind speed value at time interval i, m and c are the scaling factor and the y-axis (prediction-axis) intercept of the linear relation, respectively. The performance measures are calculated for various ANN techniques and presented in Table I.

TABLE I. COMPARISON OF DIFFERENT ALGORITHMIC PRECISION

	FFBP	CFBP	PNN	KNN	GRNN
MSE	6.748866	6.81772	3.795714	5.882085	2.856825
MAPE (%)	5.109752	5.26715	2.972917	0.91048	2.300082
R^2	0.5127	0.5316	0.8491	0.7741	0.8584

The scope of the MAPE for the five algorithms changes from 5.109752 to 0.91048, the MSE changes from 6.748866 to 5.882085 and the regression changes from 0.5127 to 0.7741. According to the forecasted errors, it is clearly seen that GRNN model is the optimal model because of the lowest errors, MSE (2.856825) and also R^2 (0.8584); Secondly, the PNN model; And then comes the KNN. The backpropagation models are not the ideal models because of the poor forecasted errors. Thus, the GRNN is found to be the superior technique for predicting the wind speed among the other ANN techniques used.

V. WIND ENERGY PROBABILITY FUNCTION

A. *Wind speed probability function*

The PDF for Weibull distribution is given by (8).

$$f_v(v) = (k/c)(v/c)^{k-1} e^{-(v/c)^k} \tag{8}$$

where $f_v(v)$ is Weibull wind speed PDF, v is wind speed, c is scale factor at a given wind generator, k is shape factor at a given wind generator. The Weibull distribution function with a shape factor of 2 is also as the Rayleigh distribution. The advantages of the Weibull distribution are as follows.

1. It is a two parameter distribution, which is more general than the single parameter Rayleigh distribution, but less complicated than the five-parameter bivariate normal distribution.
2. It has been previously shown to provide a good fit with the observed wind speed data and
3. The k and c parameters are known at one height. The characteristic of the wind depend on various factors like geography, topography and is estimated by the observed frequency wind speed in the target region.

B. *Weibull probability function for various c and k values*

Mean is given as

$$\mu = c\Gamma(1+k^{-1}) \tag{9}$$

Variance is given as

$$\sigma_v^2 = c^2\Gamma(1+2k^{-1}) - \mu^2 \tag{10}$$

From (9) and (10) it is seen that as the c-factor increases the mean and standard deviation also increases. Thereby we can infer that higher values of wind power are probable at higher values of c-factors.

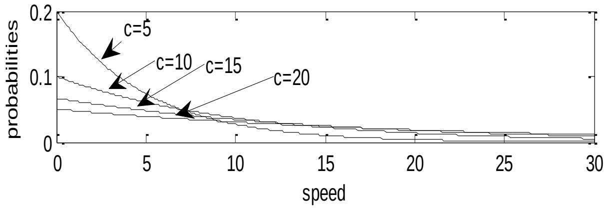

Figure 7. Weibull pdf with k=1 and c= {5, 10, 15, 20}

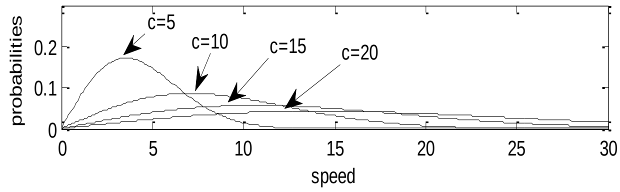

Figure 8. Weibull pdf with k=2 and c= {5, 10, 15, 20}

The Fig. 7 and Fig.8 shows the Weibull pdf of the predicted wind speed for various values of the shape factor and scale factor.

C. Wind power output at various speed

The output of the WECS with a given wind speed input may be stated as

$w = 0$, for $v < v_i$ and $v > v_o$

$$w = w_r \frac{(v - v_i)}{(v_r - v_i)} \text{ , for } v_i \le v \le v_r$$

$w = w_r$,for $v_r \le v \le v_o$

where w -WECS output power, w_r -WECS rated power,

v_i -cut-in speed, v_r -rated speed, v_o - cut-out wind speed. Thus, it seen that the WECS has

1. No power up to cut-in wind speed
2. A linear power output relationship between cut-in and rated wind speed
3. A constant rated power output between rated wind speed and cut-out wind speed and
4. No output power for wind speeds greater than cut-out wind speed.

D. Weibull PDF of wind power

The Weibull probability density function (PDF) of wind power output variable in the continuous range is given in (11).

$$f_w(w) = (k.l.v_i / c)((1 + \rho.l)v_i / c)^{k-1} \exp(-(1 + \rho.l)v_i / c)^k \quad (11)$$

where $\rho = w / w_r$, $l = (v_r - v_i) / v_i$

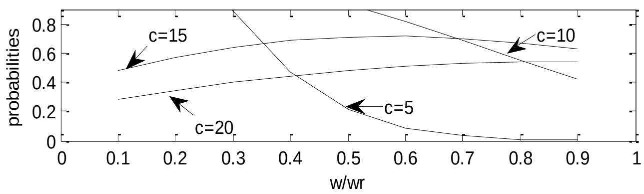

Figure 9. Variation of wind power output probabilities with k=2, c= {5, 10, 15, 20}

Fig.9 shows the variation of the wind power output probabilities for various values of shape factor and scale factor.

TABLE II. VARIATION IN PROBABILITIES OF SCHEDULED OUTPUT WITH C-FACTOR

w/wr	**Probabilities of scheduled output**			
	C=5	*C=10*	*C=15*	*C=20*
0.1	1.9264	0.9476	0.4774	0.2806
0.2	1.4216	1.0465	0.5681	0.3427
0.3	0.8856	1.0715	0.6377	0.3973
0.4	0.4724	1.0318	0.6850	0.4435
0.5	0.2176	0.9421	0.7100	0.4805
0.6	0.0870	0.8200	0.7138	0.5080
0.7	0.0303	0.6829	0.6986	0.5261
0.8	0.0092	0.5455	0.6673	0.5350
0.9	0.0025	0.4187	0.6232	0.5351

From the Table II it is inferred that as the c-factor increases, probability of output power is high at higher values of w / w_r .

VI. CONCLUSION AND SCOPE FOR FUTURE WORK

Due to the irregular nature of wind power production, wind speed prediction presents to be one of the major criteria to the researchers. Prediction horizons differ with the required application. One hour-ahead prediction is suitable for small power systems operations and one hour electricity markets. However, one day prediction is appropriate for interconnected power system operations such as unit commitments, conventional generators scheduling, as well as one day electricity markets. Therefore, 18 hours (medium term prediction) forecasting model is developed for wind speed using ANN techniques and the best algorithm is chosen from the precision and accuracy of its forecasted value. Also, the probability distribution of the same is found out using Weibull probability distribution.

The future research is as follows:

- With the availability of sufficient data for actual wind farm production, the developed wind speed prediction tools and the wind farm model can be further examined and analyzed for the prediction of large scale wind farm production.
- The paper uses wind speed data series recorded at an Automatic Weather Station. Further evaluation and accuracy investigation of the proposed prediction models can be conducted on wind data collected from other wind farm facilities sites.
- To incorporate the wind speed prediction in one-day ahead electricity markets.
- The PDF of the predicted wind power can be used in calculating the underestimation and overestimation cost of the wind generators in a wind farm.

REFERENCES

[1] Cameron Potter. W, and Michael Negnevitsky, "Very Short-Term Wind Forecasting for Tasmanian Power Generation", IEEE Transactions on Power Systems, vol. 21, no. 2, pp.965-972, May 2006.

[2] Han Xiaojuan, Yang Xiyun and Liu Juncheng, "Short-Time Wind Speed Prediction for Wind Farm Based on Improved Neural Network", Proceedings of the 8th World Congress on Intelligent Control and Automation, Jinan, China, pp.5891-5894, July 2010.

[3] Kang M.S, Chen C.S, Ke Y.L, Lin C.H, and Huang C. W. "Load Profile Synthesis and Wind Power Generation Prediction for an Isolated Power System", IEEE Conference, pp.124-129, 2005.

[4] Shuhui Li, Donald Wunsch.C, Edgar O'Hair.A, and Michael Giesselmann.G. "Using Neural Networks to Estimate Wind Turbine Power Generation", IEEE Transactions on Energy Conversion, Vol. 16, No. 3, pp 276-282, September 2001.

[5] Tarek Hussein Mostafa EL-Fouly, "Wind Farms Production: Control and Prediction", Waterloo, Ontario, Canada, 2007.

[6] Tarek H. M. El-Fouly, Ehab F. El-Saadany, and Magdy M. A. Salama, "One Day Ahead Prediction of Wind Speed and Direction", IEEE

Transactions on Energy Conversion, Vol. 23, No. 1, pp.191-201, March 2008.

[7] Thanasis G. Barbounis, John B. Theocharis, Minas C. Alexiadis, and Petros S. Dokopoulos, "Long-Term Wind Speed and Power Forecasting using Local Recurrent Neural Network Models", IEEE Transactions on Energy Conversion, Vol. 21, No. 1, pp.273-284, March 2006.

[8] M.K. Deshmukh and C. Balakrishna Moorthy, "Application of genetic algorithm to neural network model for estimation of wind power potential", Journal of Engineering, Science and Management Education, Vol. 2, pp.42-48, October 2010.

[9] Donald F. Specht, "A General Regression Neural Network ", IEEE Transactions on Neural Networks, Vol. 2, No. 6, pp.568-576, November 1991.

[10] Dudani, S.A, "A note on distance-weighted k-nearest neighbour rules", IEEE Transactions on Systems, Man, and Cybernetics, Vol.8, pp.311-313, 1976.

[11] J.V Seguro and T.W Lambert, "Modern estimation of the parameters of the weibull wind speed distribution for wind energy analysis", J.Wind Eng. Ind. Aerodyn., Vol.1, No.85, pp.15-84, 2001