Energy Conversion and Management 52 (2011) 738–745

Very Short-Term Wind Speed Prediction for Wind Turbine Control Applications: A New ANN-based Methodology

M. M. Ardehali[1], S. A. Pourmousavi Kani[2,*]

[1] Power and Energy Management Division, Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Hafez Ave, Tehran, Iran.

[2] Center for Excellence on Power systems, Wind Energy Laboratory, Electrical Engineering Department, Amirkabir University of Technology, Tehran 15914, Iran.

Abstract: Nowadays, wind turbine control in the whole range of wind speed is a very important issue especially for modern wind turbines with high amount of power generation. Due to the importance of short-term decisions such as connection of a load, changing the pitch of the blades and/or any other control action which involves delays, very short-term wind speed prediction plays an important role in wind power production. This paper presents the hybridization of the Markov chain approach with neural networks in order to consider short-term and long-term patterns in very short-term wind speed prediction. First, primary prediction with ANN with limited input data is considered concerning short-term trends. Then, transition probability for predicted values and four other indices are calculated. Eventually, final prediction has been done utilizing another neural network. A wind speed data set sampled every 2.5 seconds from Manjil, Iran is utilized for verification and comparison purposes. We will show that our method is able to improve the performance of the system, reducing the Maximum Prediction Error (MPE) and Mean Absolute Percentage Error (MAPE) better than the ones obtained by the system using single neural networks while uncertainty of prediction is declined with proposed model. On the other hand, limited number of input and length of data set is needed for both artificial neural networks, further avoiding over-training and contributing to the significant reduction of the required CPU time and feasible to be used for on-line application.

[*] Corresponding Author: S. A. Pourmousavi Kani,
Email Address: ali.pourmousavi@gmail.com ; Tel: (+98) 111 367 2332; Cell: (+98) 912 522 6612
Postal Address: Shahid Keshavarz Dormitory, Shahid Monaee Ave, Shahid Bahoonar St, Nyavaran, Tehran, Iran. P.O. Box: 1978711167.

Keywords: Artificial Neural Network, Markov Chain, Mean Absolute Percentage Error (MAPE), Maximum Prediction Error (MPE), Uncertainty.

1. Introduction:

The contribution of wind power in market-driven power systems together with the uncertain nature of the wind resource have led to many research efforts on methodologies to predict future wind speed/power production. Very short-term prediction is a difficult parameter to calculate but if it is done correctly, it can be used as a method for achieving better results at a wind farm and consequently can improve profits for the owner. The main area of application for very short-term prediction horizon is related to wind turbine control issues. Applications such as avoiding voltage and frequency fluctuations due to variation in wind power, reducing unacceptable shocks in the conventional power units because of a sudden cut-off of wind power due to excessive wind speeds would benefit from accurate very short-term forecasts of wind speed/power predictions [1-2]. Also, the importance of short-term decisions in control systems of a wind turbine such as connection of a load, changing the pitch of the blades and/or any other control action which involves delays are the encouragements for very short-term wind speed prediction [3-5].

Also, in large modern wind turbines with high investment cost, it is very important for producers to extract maximum available energy from wind speed. So, different control algorithms for entire operational regions need to be applied. Advanced control theory known as model predictive control (MPC) is one of the most interesting method implemented to control the wind turbines in all the operational regions [6]. The behavior of the model (MPC) is predicted based on past measurements and computed future inputs [6]. Then, it is obvious that very short-term wind speed prediction will find a wide range of applications in MPC. Some researches have been focused on model predictive control. In literature [7], a pitch controlled technique is proposed for grid-connected wind turbine in a small power system by Nanayakkara. The proposed pitch controller incorporated the predicted effective very short-term wind speed to minimize effects on the power system while producing optimum wind-generated power. Another operational optimization strategy is introduced at the wind park control level in [8] by Moyano. There was assumed that for individual wind turbines, very short-term wind speed forecasts are known. This operational strategy was also developed with a concern on the minimization of the connection/disconnection changes of the individual wind generators, for a given time horizon.

The area of short-term wind speed prediction can be generally subdivided into two main groups, depending on the underlying prediction model used. These can be either based on numerical weather prediction (NWP) models, similar to those used by national meteorological agencies, or other alternative approaches [9-10]. The second category encompasses, amongst others, artificial intelligence (fuzzy logic [11-13], artificial neural networks [2, 5, 14-22]) and statistical models [1, 4, 12, 19, 23-25]. Also, some post-processing approaches based on statistical methods are introduced as a hybrid algorithm in conjunction with NWP models [9-10, 12, 23]. While the NWP-based models being usually the best option for longer term predictions (over around days, weeks and months ahead), artificial intelligence and statistical approaches are the most promising methods considering short-term and its subclasses time horizons (over around seconds, minutes and few hours ahead) predictions. Also, statistical approaches found some problem in very short-term wind speed prediction. First, most of them assumed that the wind speed data is normally distributed; whereas it is a well known characteristic of general wind speed series that its variation at a given site can be modeled using the Weibull distribution which is not normally distributed function. So, a transformation of the original wind speed data was required which makes the time series unstable and difficult to predict [1]. Second, highly unstable and variable nature of wind speed series particularly for very short-term prediction time horizon needs more complicated functions between inputs and outputs data to estimate the relations, in spite of simple linear functions used in Bayesian or linear prediction methods. So, artificial intelligence-based methods attract more attention from researchers for accurate wind speed prediction especially for short-term and very short-term prediction horizons. In the literature [2, 11-12, 17], artificial intelligence-based methodologies are found to be more accurate as compared to traditional statistical models.

The majority of studies based on artificial intelligence in wind speed prediction focused on short-term prediction horizon [14-23]. Also, different parameters as input variables have been applied in these studies such as wind speed data, relative humidity, generation hour and etc. Most of these researches utilized various artificial intelligence-based approaches as hybrid models [5, 10-13]. Very short-term wind speed prediction has been considered in the following studies. Riahy in [4] utilized the linear prediction method in conjunction with filtering of the wind speed waveform as a new method for short-term wind speed forecasting. Safavieh et al

[20] proposed a new integrated method utilizing wavelet-based networks and Particle Swarm Optimization (PSO) to predict very short-term wind speed prediction. Pourmousavi et al [21] developed a new model for very short-term wind speed prediction utilizing ANN, Markov Chain and linear regression. In literature [5], Pourmousavi et al, introduced a new ANN-based methodology for very short-term wind speed prediction in conjunction with Markov chain approach.

In this paper, an artificial neural network based predictor is proposed to forecast wind speed in very short-term time scale, which can be mathematically modeled as a highly non-linear random process. In this new predictor, the short-term patterns in wind speed data are grasped by artificial neural networks and the long-term patterns are considered utilizing Markov chain approach and four neighborhood indices.

The proposed model is applied for different time scales. Obtained results, from the proposed model, are compared with their corresponding values obtained using single neural networks. The presented results validate the effectiveness of the new prediction model for wind speed.

In the following, organization of the paper is described briefly. The new predictor configuration, its implementation and the aims and reasons for selected input variables are deeply discussed in section 2. Also, the whole process with details is presented in this section. Section 3 is allotted to results and verification purposes. In the last section, a conclusion has been made from the whole study.

2. Proposed Model:

Applying ANN as a wind speed predictor without combination with other approaches such as statistical or other artificial intelligence methods, increases number of input variables to the network and data sets for training; Because both short-term and long-term trends have to be exist on the training data and hence in the input variables [1-2]. In one hand, learning algorithms for ANN show slow rate of learning in presence of huge input variables and data sets which are not efficient for very short-term wind speed prediction as on-line predictor. On the other hand, increasing number of input variables and data sets for network training, need more sophisticated network with more layers and neurons in each layer which, in turn, increase calculation time. It is obvious that calculation time is an important parameter for very short-term wind speed prediction for on-line applications.

As pointed out by Yin in literature [26], there are two limitations on the use of ANN models, which seriously degrade the prediction performance of ANN models. One is over-training. Over-training occurs when the capacity of the ANN for training is too great because it is too large or is allowed too many training iterations. Therefore, as discussed above, applying ANN as a single approach for prediction may increase the possibility of over-training because of huge input variables and training data sets. The other is that ANN models are not effective for extrapolation [26]. The benefit of ANNs is lost when they are needed to extrapolate beyond available experimental data.

In this study, Markov chain approach is applied to grasp long-term trends in wind speed data to solve over-training problem. Because, a simple structure for ANN is used with minimum number of input variables and data sets for training. Also, Markov chain approach memorizes long-term behavior of the signal so it will reduce the error obtained from extrapolated prediction. Transition probability will be calculated for the primary predicted values which show the transition probability to the calculated value considering previous record of data. To show the efficiency of the proposed model to reduce extrapolation problem, Maximum Prediction Error (MPE) is introduced in section 3. Results show that maximum prediction error is reduced in the new predictor efficiently. As another solution for extrapolation problem, artificial samples, covering the entire range as much as possible, are drawn based on the existing knowledge about the modeled problem, and then used to initialize the ANN to ensure most of the future prediction will be an interpolation.

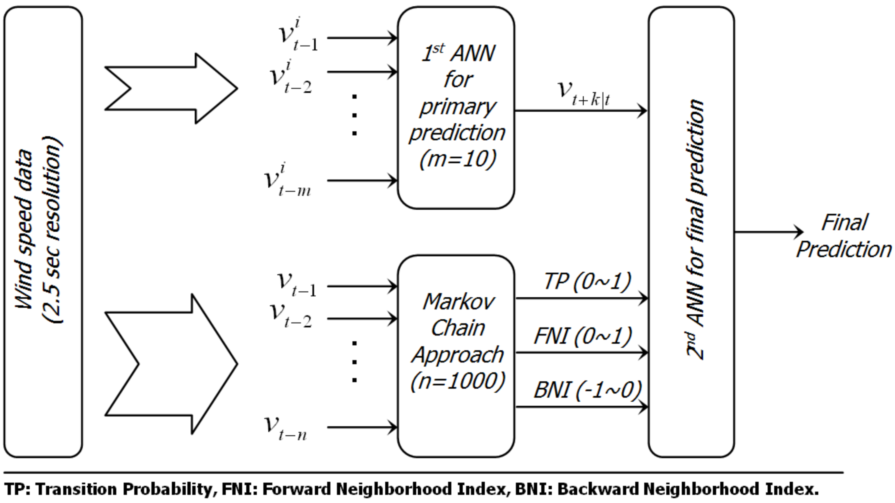
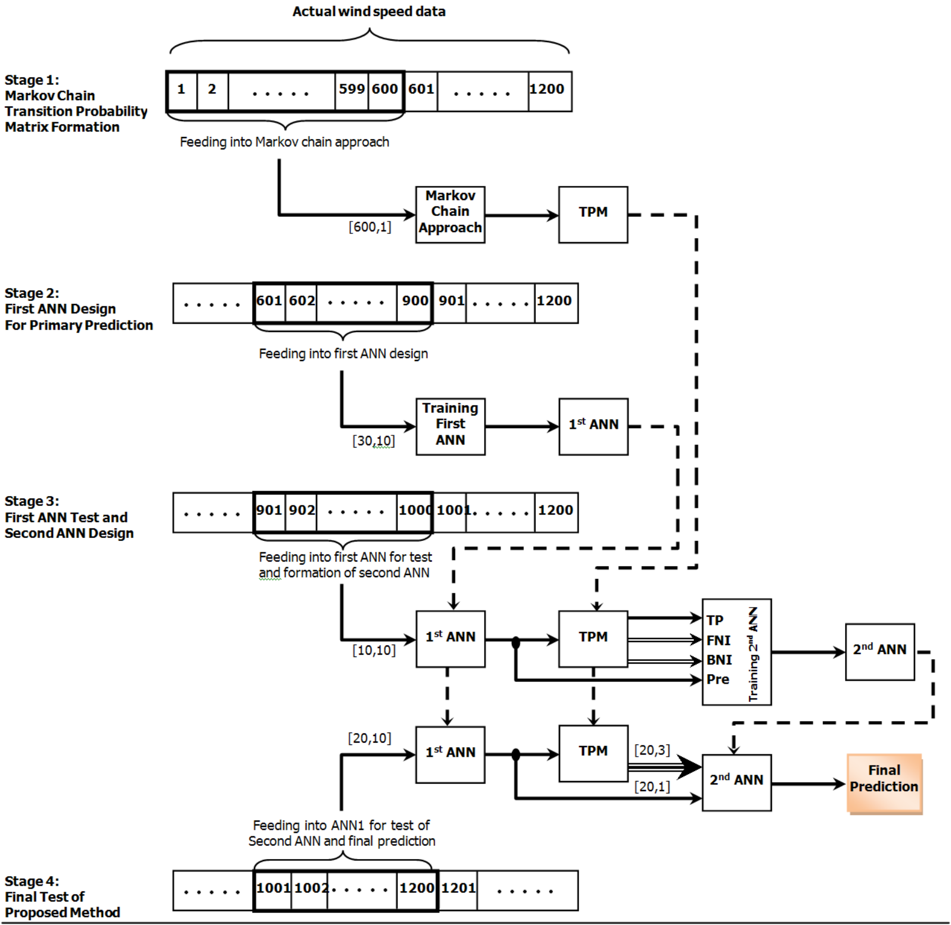


Fig. 1. Outline for the proposed method.

Eventually, the outline of the proposed model in this study is introduced according to Fig. 1.

A set of data that extends to 50 minutes is used in investigating the accuracy of the model for predicting wind speeds up to 7.5 seconds ahead.

In the proposed predictor, two artificial neural networks are applied. First ANN is used for primary prediction according to short-term trend in wind speed signal, so 10 actual wind speed data are utilized from t to $t-10$ as input variables to the network. Primary prediction can be carried out for different time horizon by first ANN. 30 data sets with 10 measured wind speed in each set are selected for training. The structure of the network will be discussed in section 2-1. After primary prediction using first ANN, transition probabilities for predicted values and other four indices together with primary predictions are fed to the second ANN as input variables. Methods for transition probability calculation and other four neighborhood indices will be discussed in subsection 2.2. Finally, the new predictor including two ANNs and Markov chain approach can be used for different time horizon prediction. In Fig. 2, the whole process to form ANNs and Markov transition probability matrix by measured data toward final prediction is illustrated in details.



TPM: Transition Probability Matrix, TP: Transition Probability, FNI: Forward Neighborhood Index, BNI: Backward Neighborhood Index.

Fig. 2. Different stages for implementation of the proposed predictor.

According to Fig. 2, the steps taken to predict wind speed in different time horizons are as follows:

Stage1: Markov chain transition probability matrix (TPM) is calculated. 600 points of wind speed is used to form this matrix. In subsection 2.2.1, the whole process for TPM calculation has been discussed deeply.

Stage 2: Another 300 points of wind speed data is applied to design first ANN for primary prediction. The structure and type of network are surveyed in section 2-1. Number of inputs and length of each data set are illustrated in Fig. 2.

Stage 3: To test the first ANN and design second ANN, another 100 data set is selected from the rest of data. According to Fig. 2, First ANN designed in the previous stage is utilized for primary prediction. Then, TPM which has been calculated in stage 1 is applied to calculate the required coefficients. These coefficients will be discussed in subsection 2.2. The second ANN involves six input variables-primary wind speed prediction, its transition probability value, transition probability to the 1 and 2 forward states from current prediction state and transition probability to the 1 and 2 backward states from current prediction state (neighborhood indices).

Stage 4: In the last phase, both ANNs and TPM were calculated in the previous stages have been applied for final prediction. 200 wind speed data is considered in this stage. In comparison with primary predicted values, the effectiveness of proposed model will be illustrated.

All stages above have to be applied for different prediction time horizons.

2.1- The ANNs used: type and structure

In literatures [14-23], detail descriptions and applications of ANN with different structure and learning algorithms are reported. In this study, the well-known multi-layered perceptron (MLP) is used for both ANNs. This structure is one of the simplest and fastest ANN types which are used in different studies. The used ANN consists of one input layer, one hidden layer, and one output layer. In the output layer, only one neuron is needed, $\hat{v}(t+k|t)$, where k is the time step and $\hat{\mathrm{v}}$ is the anticipated wind speed at time $\mathrm{t}+\mathrm{k}$ which is calculated at time t. Since the number of neurons in each layer influence the speed and stability of network, sensitivity analysis have been carried out for different number of neurons in input and hidden layers. Fig. 3

shows the mean absolute percentage error obtained for different number of neurons in input and hidden layers while the output layer has one neuron.

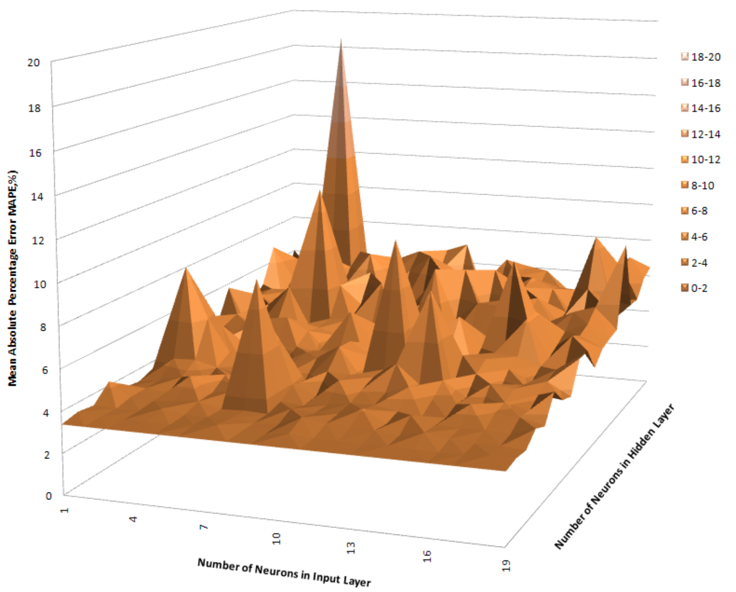


Fig. 3. Sensitivity analysis for number of neurons in input and hidden layers.

The best structure for the first ANN according to Fig. 3 is as follows:

- Number of neurons in input layer: 5
- Number of neurons in hidden layer: 2
- Number of neurons in output layer: 1
- Number of training vector: 30
- Value of the learning rate: (0.01-0.08)

Number of neurons in input layer is selected considering required time for calculation and obtained error. Since the second ANN has six input variables and one output variable, number of neurons in each layer should be in the range of variables. So, the best structure for the second ANN which is used for final prediction is as follows:

- Number of neurons in input layer: 3
- Number of neurons in hidden layer: 0
- Number of neurons in output layer: 1
- Number of training vector: 10

- Value of the learning rate: (0.01-0.05)

2.2- Markov chain approach:

In literatures [23-25], some methods based on Markov chain approach are introduced for short-term wind speed prediction. First-order Markov chain approach is applied in [23] for wind speed modeling directly. It is observed that for short periods, the parametric results were close to the measured values. Also, in [25], second-order Markov chain approach is utilized for short-term wind speed prediction. The results obtained in this study in comparison with observed wind speed data shown that the statistical characteristics are satisfactorily preserved, but the accuracy of prediction were not good enough. Kantz et al. [24] used Markov chain approach to model turbulent wind speed data. Besides, in literatures [5, 21], Markov chain is used in hybrid algorithms for very short-term wind speed prediction in conjunction with ANN. Different algorithms are applied in these studies for final prediction.

All Markov chain models are based on the transitional probability matrices of various time steps. Most often, a first-order Markov chain implies preservation of statistical parameters and especially the first-order autocorrelation coefficient in the synthetic sequences. In order to calculate the Markov chain transitional probabilities, initially the wind speed variation domain is divided into many states. Such a state categorization may be rather arbitrary depending on the purpose, but herein, it is determined with upper and lower limit difference of 0.1 m/s. [23]

For the Markov process, the probability of the given condition in the given moment may be deduced from information about the preceding conditions. A Markov chain represents a system of elements moving from one state to another over time. The order of the chain gives the number of time steps in the past influencing the probability distribution of the present state, which can be greater than one. Many natural processes are considered as Markov processes [23]. In fact, the probability transition matrix is a tool for describing the Markov chains' behavior. Each element of the matrix represents probability of passage from a specific condition to a next state. [25]

Let $\mathrm{X(t)}$ be a stochastic process, possessing discrete states space $S = \{1,2, \dots, K\}$. In general, for a given sequence of time points $\mathrm{t_1 < t_2 < \cdots < t_{n-1} < t_n}$ the conditional probabilities should be [25]:

$$Pr\{X(t_n) = i_n | X(t_1) = i_1, \dots, X(t_{n-1}) = i_{n-1}\} = Pr\{X(t_n) = i_n | X(t_{n-1}) = i_{n-1}\} \qquad (1)$$

The conditional probabilities $Pr\{X(t) = j|X(s) = i\} = P_{ij}(s,t)$ are called transition probabilities of order $r = t - s$ from state i to state j for all indices $0 \le s \le t$, with $1 \le i$ and $j \le k$.They are denoted as the transition matrix $P_{transition}$ For k states, the first order transition matrix P has a size of $k \times k$ and takes the form [25]:

$$P_{transition} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,k} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,k} \\ \vdots & \vdots & \vdots & \vdots \\ p_{k,1} & p_{k,2} & \cdots & p_{k,k} \end{bmatrix} \quad (2)$$

The state probabilities at time t can be estimated from the relative frequencies of the k states. If n_{ij} is the number of transitions from state i to state j in the sequence of speed data, the maximum likelihood estimates of the transition probabilities is:

$$p_{ij} = {n_{ij}}\Big/{\sum_j n_{ij}} \quad (3)$$

A second order transition probability matrix for k state can be shown symbolically as below:

$$P_{transition} = \begin{bmatrix} p_{1.1,1} & p_{1.1,2} & \cdots & p_{1.1,k} \\ \vdots & \vdots & \vdots & \vdots \\ p_{1.k,1} & p_{1.k,2} & \cdots & p_{1.k,k} \\ p_{2.1,1} & p_{2.1,2} & \cdots & p_{2.1,k} \\ \vdots & \vdots & \vdots & \vdots \\ p_{2.k,1} & p_{2.k,2} & \cdots & p_{2.k,k} \\ p_{3.1,1} & p_{3.1,2} & \cdots & p_{3.1,k} \\ \vdots & \vdots & \vdots & \vdots \\ p_{k.k,1} & p_{k.k,2} & \cdots & p_{k.k,k} \end{bmatrix} \quad (4)$$

In this matrix, the probability $p_{j.k,l}$ is the probability of the next wind speed state l if the current wind speed state is k and the previous wind speed state were j. It has a size of $k^2 \times k$. This is how the probability of making a transition depends on the current state and on the preceding states [23]. The following properties of the transition matrix are valid by definition. Any state probability varies between zero and one. Notationally,

$$0 < p_{j.k,l} < 1.0 \quad (5)$$

On the other hand, the row summation in the transition probability matrix is equal to 1 and hence notationally,

$\sum_{l=1}^{n} p_{j.k,l} = 1.0$ (6)

2.2.1- second order Probability Transition Matrix (TPM) formation

Initially, the wind speed time series were converted to wind speed states, which contains wind speeds between certain values. In [11] wind speed states determined according to the average $\hat{v}$ and standard deviation S_v of the available wind speed time series. In [12] the wind speed states have been adopted with an upper and lower limit difference of 1 m/s of wind speed, Based on the visual examination of the histogram of the wind speed data. Such a state categorization may be rather arbitrary depending on the purpose and the real wind speed data used for verification [12]. In this study, sensitivity analysis has been carried out for different strategies. Categorization with an upper and lower limit difference of 1 m/s, 0.1 m/s and according to the average, $\hat{v}$, and standard deviation, S_v, of wind speed is considered. Finally, the best results are obtained from categorization with an upper and lower limit difference of 0.1 m/s. Based on state matrix, it is possible to find the number of transition from two preceding states in the sequence of wind speed data to another state at time $t + k$. Finally, the transition probabilities are calculated according to Eq. 4.

According to stage 1 from Fig. 2, transition probability matrix is formed by 600 preceding wind speed data. Calculated matrix has been used for primary predicted values. First, Markov state for primary predicted values by first ANN is calculated for one step ahead. Then, according to Markov transition probability matrix, the probability of predicted value in the next step is calculated (TPM in Fig. 2). This is the process which is carried out for all primary predictions. Notice that predicted values are produced in the previous step by first ANN. For longer prediction horizon, the above transition probability matrix will multiply to itself according to number of time steps in the future.

Primary predicted values and their states in comparison with real data states show that more than 83% of actual data are in the state corresponding to the predicted values or two upper and twp lower states. This survey is carried out for 330 wind speed data points. Primary predictions have been done for this data set. Then, Markov state for predicted values is calculated. Finally, the state of real values has been compared with state of predicted values and other four upper and lower states. In the rest of the paper, upper states and their corresponding probabilities are called Forward Neighborhood Index (FNI); while the lower states and their corresponding probabilities are called Backward Neighborhood Index (BNI). Table 1 illustrates number of observation in each state. It is obvious that there is a logical relation between different states in comparison with real states. To find this relation, second ANN is applied between predicted values, transition probability corresponding to predicted values and FNIs and BNIs. The transition probability of the predicted values' state, transition probabilities toward two next states (FNIs) and transition probabilities toward two backward states (BNIs) are explanatory inputs to the second ANN in order to have an acceptable performance. BNIs are negative values between -1 to 0 to represent transition probability of lower states; while FNIs are positive values in the range of 0 to 1. Notice that current study has been carried out for one step-ahead prediction horizon. All simulations have been repeated 5 times, and the results are reported in Table 1.

Table 1

Number of observations in each state in comparison with state of actual wind speed.

No. Iteration	**Predicted values state**		**1FNI**		**2FNI**		**1BNI**		**2BNI**	
	No. Obs.	% of total	No. Obs.	% of total	No. Obs.	% of total	No. Obs.	% of total	No. Obs.	% of total
1	72	21.8	51	15.5	21	6.36	84	25.5	46	13.9
2	84	25.5	41	12.4	23	6.97	78	23.6	50	15.2
3	83	25.2	44	13.3	21	6.36	80	24.2	48	14.5
4	82	24.8	48	14.5	22	6.67	78	23.6	43	13
5	79	23.9	46	13.9	22	6.67	78	23.6	48	14.5
Mean	80	24.2	46	13.9	21.8	6.6	79.6	24.1	47	14.2

It is obvious from Table 1 that there is a relation between five different states with real states. In fact, predicted values can be improved considering other states in neighboring the predicted state. So, BNI and FNI are introduced in this study as input variables to the second ANN to improve the predicted values. Also, this fact is shown in Fig. 4. Which is illustrated number of observations in each pre-defined stated.

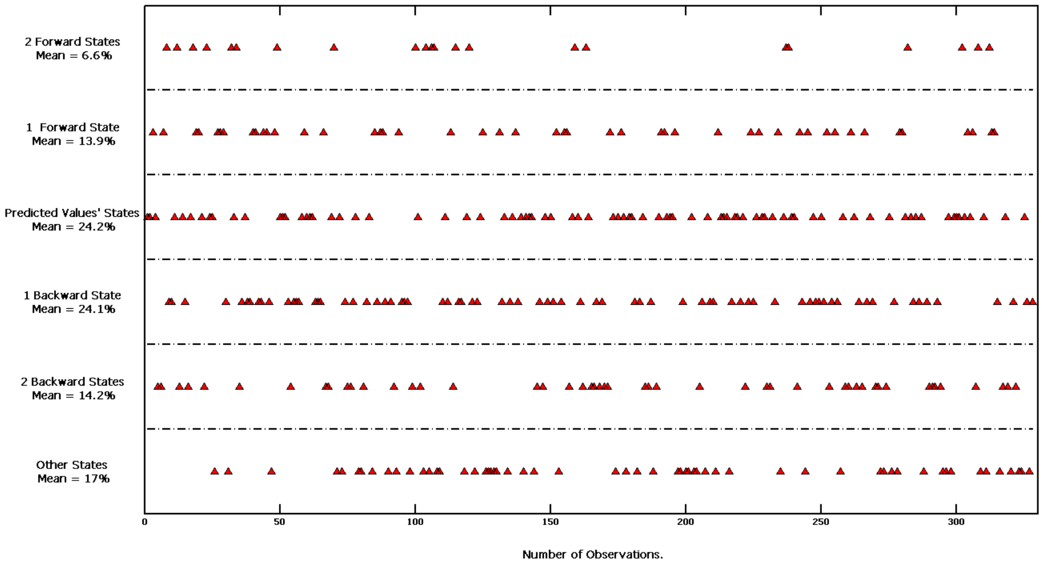


Fig. 4. Number of observation for different Neighborhood indices.

The relationships between the primary prediction and these factors obtained from Markov chain are difficult to determine. Because ANNs can encode complex, non-linear relations, second ANN is used to capture the relationships between the primary prediction and obtained probabilities.

3. Results:

In forecasting, determining which forecasting method is the best is one of the forecaster's prime concerns. In most cases, model performance is determined by examining model accuracy. To determine model accuracy, though, is also a complex issue, given that several model accuracy measures exist.

A model's prediction error is classically defined as the difference between the measured and the predicted value. A horizon dependent model error $e(t+k|t)$ should be considered as:

$$e(t+k|t) = v(t+k) - \hat{v}(t+k|t) \qquad (7)$$

where $v(t+k)$ is the measured wind speed at time *t* + *k*, $\hat{v}(t+k|t)$ is the wind speed forecast for time *t* + *k* computed at time *t*. The most commonly used evaluation criterion is the Mean Absolute Percentage Error (MAPE). The MAPE can be defined as:

$$MAPE(k) = \frac{1}{N}\sum_{t=1}^{N}\left(\left|\frac{e(t+k|t)}{v(t+k)}\right| \times 100\right) \qquad (8)$$

where |.| denotes the absolute value. Also, *k* represents the prediction horizon. Another criterion which is called Maximum Prediction Error (MPE) is also used in this study. In wind turbine control applications, it is very important to reduce maximum error of prediction because; a great prediction error may create an unstable condition for wind turbine because of wrong control command. So, a sudden cut-off of wind power due to unstable condition may cause unacceptable shocks in the conventional power units. Since the prediction has been done for a specific number of values with different time steps, the maximum error is calculated as follows:

$$MPE(k) = max\left\{\left|\frac{e(t+k|t)}{v(t+k)}\right| \times 100\right\} ; t = 1, \cdots, N \qquad (9)$$

As suggested in [36], the skewness and kurtosis permit a distribution-oriented analysis of the forecast error. The skewness can be estimated using Fisher's formula: [27]

$$\hat{v}_e(k) = \frac{N}{(N-1)(N-2)} \sum_{t=1}^{N} \left(\frac{e(t+k|t) - \hat{\mu}_e(k)}{\hat{\sigma}_e(k)}\right)^3 \qquad (10)$$

The skewness is linked to the third moment of a distribution; it indicates the degree of symmetry of the error distribution. If the skewness is null then the distribution is symmetrical, if the skewness is negative then the distribution is left-skewed (the left tail of the distribution is the "longest") and, if the skewness is positive, the distribution is right-skewed (the right tail of the distribution is the "longest").

The excess kurtosis can be defined as: [27]

$$\hat{\kappa}_e(k) = \frac{N.(N-1)}{(N-1)(N-2)(N-3)} \sum_{t=1}^{N} \left(\frac{e(t+k|t) - \hat{\mu}_e(k)}{\hat{\sigma}_e(k)}\right)^4 - \frac{3(N-1)^2}{(N-2)(N-3)} \qquad (11)$$

The excess kurtosis is linked to the fourth moment of the distribution. It provides information on the shape of the error distribution with respect to a Gaussian distribution. If the kurtosis is positive, the distribution presents a sharper peak around the mode and longer tails than the Gaussian distribution. Conversely, if the kurtosis is negative the distribution presents a flatter peak around the mode and shorter tails. In two previous Eq (10) and (11), $\hat{\mu}_e(k)$ is the mean of the errors and can be defined as follows:

$$\hat{\mu}_e(k) = \overline{e(k)} = \frac{1}{N} \sum_{t=1}^{N} e(t+k|t) \qquad (12)$$

Also, $\hat{\sigma}_e(k)$ is standard deviation of the errors. It can be computed as follows:

$$\hat{\sigma}_e(k) = \sqrt[2]{\frac{1}{N} \sum_{t=1}^{N} \left(e(t+k|t) - \overline{e(k)}\right)^2} \qquad (13)$$

Wind speed data on a horizon of seconds is an arbitrary non-linear non-stationary stochastic process whose memory is short-ranged. [24] This random behavior known as turbulence which does not exist in wind speed data on a horizon of minutes and hours. So, very short-term wind speed prediction is a highly random process prediction; hence the errors are more here in comparison with longer prediction horizon. Here, a set of data that extends to 50 minutes is used in investigating the accuracy of the model for predicting wind speeds up to 7.5 seconds ahead. Fig. 5, illustrates actual wind speed data with 2.5 seconds resolution. The maximum, minimum, mean and standard deviation of wind speed data illustrated in Fig. 5 are 7.2664 (m/s), 2.2121 (m/s), 4.0887 (m/s) and 0.9042 respectively. The highly random short-ranged memory characteristics of wind speed data are obvious in Fig. 5.

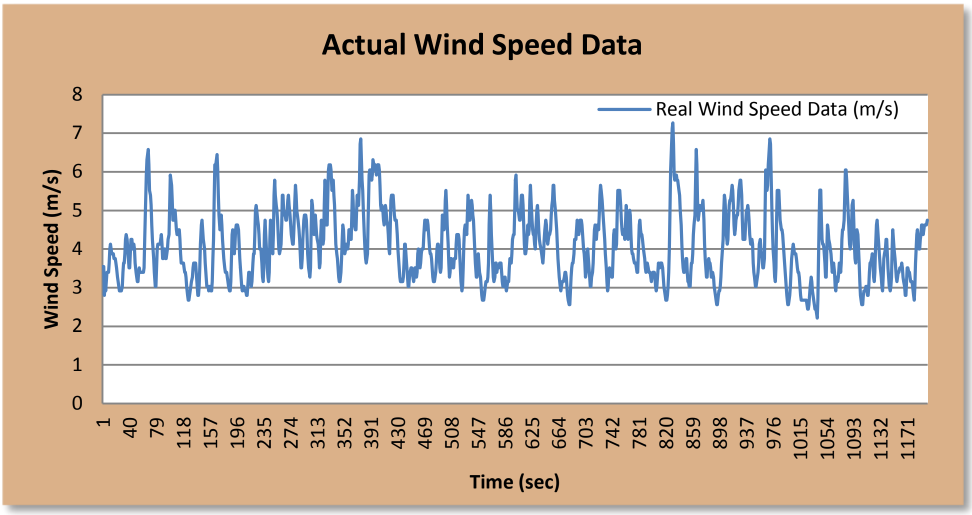
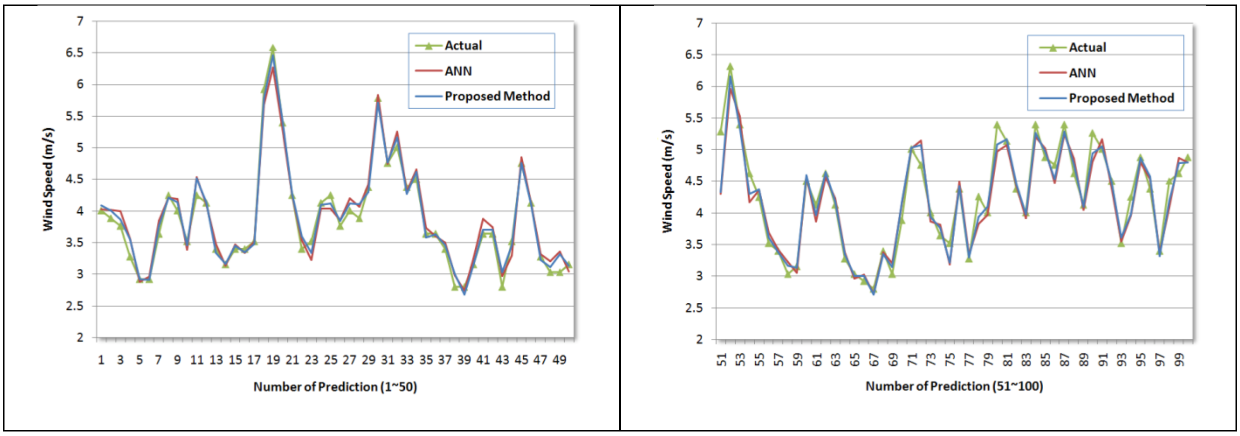


Fig. 5. Actual wind speed data which are applied in this study.

According to Fig. 2, 200 points of wind speed data is applied from the rest of the data for verification purposes. Also, predictions have been carried out for three different time horizons. In Fig. 5, predicted values for one step-ahead by both methods in comparison with actual wind speed data are illustrated comparatively. For better resolution, the results are divided into four ranges to sketches in four separate figures.



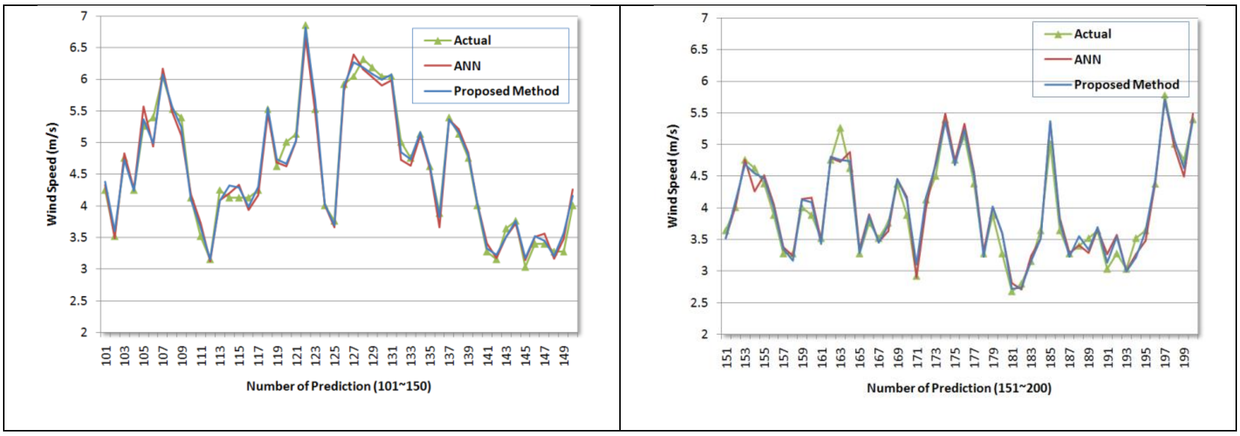


Fig. 6. Wind speed data observed, forecasted with First ANN and proposed method for one step-ahead prediction.

In Table 2, the results are reported for different criterion in comparison with primary results obtained by first ANN.

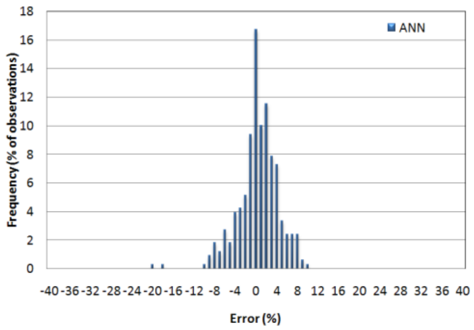
Table 2.

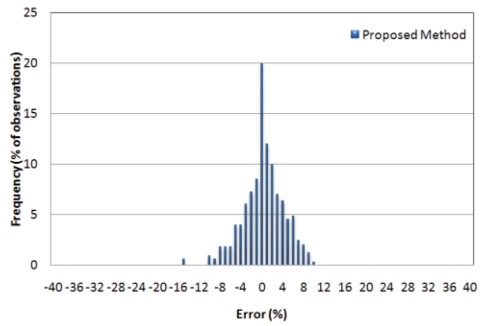
Comparison between different prediction methods for multi-step ahead prediction.

Prediction Horizon	Primary Predicted Values by first ANN				Proposed Method			
	MAPE(%)	MPE(%)	Skewness	Kurtosis	MAPE(%)	MPE(%)	Skewness	Kurtosis
t+2.5 sec	3.5259	20.3222	-1.2905	2.7566	3.0233	15.5862	-0.9610	4.5790
t+5 sec	8.3471	31.6608	-1.1459	1.1289	7.9763	29.4970	-0.8976	1.8802
t+7.5 sec	11.4458	36.8226	-1.2918	0.6357	11.2987	34.2081	-0.8738	1.7191
Mean	7.77	29.61	-1.2427	1.51	7.43	26.43	-0.9108	2.7261

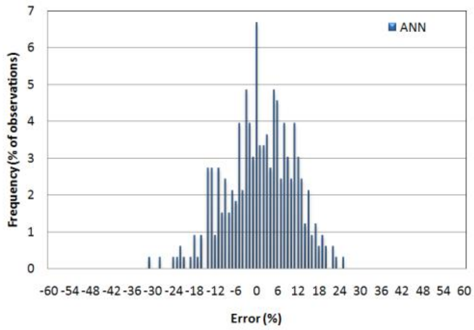
The values in Table 2 are the mean value obtained after 10 times run of algorithm. Three different time horizons are considered in Table 2. It is obvious from Table 2 that the results from proposed method show slightly higher efficiency in comparison with ANN for different time horizons. Greater improvement can be found for smaller horizon. For example, for one step-ahead prediction, mean absolute percentage error is 3.5259 by ANN, whereas for proposed method, MAPE is reduced to 3.0233 which present about 14.25% improvement in prediction. Also, skewness in the same prediction horizon for proposed method is -0.9610, while it is -1.2905 for ANN in the same prediction horizon. So, it can be concluded that error distribution is more symmetrical for proposed method. This conclusion is valid for other prediction horizons according to Table 2. As another criterion, kurtosis coefficient is a greater positive value for proposed method. It shows that the error distribution in proposed predictor illustrates a sharper peak around the mode and longer tails than the Gaussian distribution. Therefore, the proposed predictor leads to a sharper distribution of the errors and lower uncertainty. Also, this conclusion is valid for other prediction horizons according to Table 2. While the improvement in criteria for longer prediction horizons is less than those for lower horizon, but error distribution improved more.

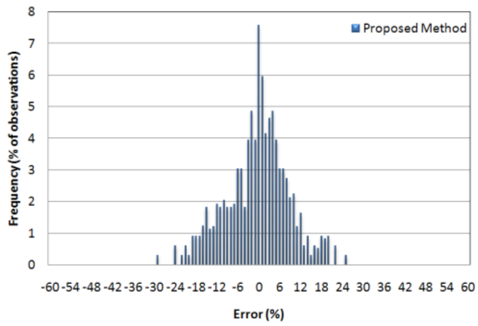
In figs. 7 to 9 the error distributions for the 2.5, 5 and 7.5 seconds-ahead horizons are presented for both errors obtained by primary prediction and proposed method. They clearly show the reduction of the errors as well as lower uncertainty. It can be concluded from more symmetrical and sharper distribution of error histogram for proposed method in comparison with ANN. All histograms are calculated from 200 data points for both prediction methods.



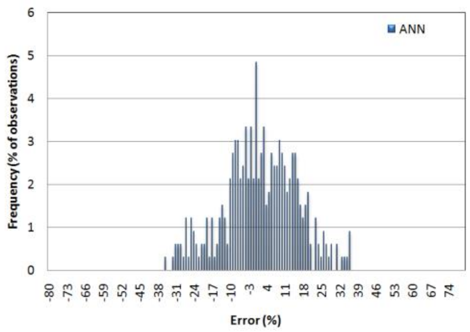


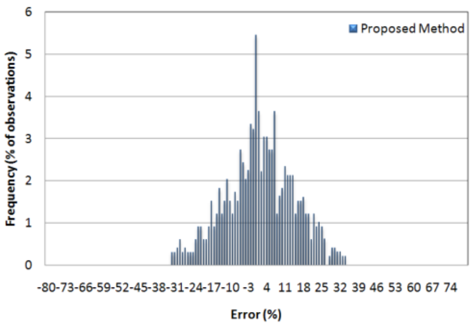
2.5 second-ahead horizon





5 second-ahead horizon





7.5 second-ahead horizon

(a) (b)

Fig. 7. Error distribution histogram for multi-step ahead prediction horizon; (a) ANN, (b) Proposed method.

Also, lower uncertainty can be illustrated by analyzing the number of prediction errors inferior to a certain error margin. Table 3 represents the number of predictions between ±5%, ±15% and ±30% error margins for both methods for comparison. For example, for 5 second-ahead horizon, 34.2424% of the errors are between ±5% measured values for primary predicted values by first ANN, whereas for the proposed predictor, 39.4% of the errors are in the same error margin for the same horizon. This trend is true for 2.5 seconds horizon with smaller improvement in the proposed predictor rather than the other one. It is clear from Table 3 that the proposed predictor reduces the number of error for higher margins. So, the most errors in the new predictor are limited to the first narrower margin. It means that the new predictor is reduced the uncertainty of prediction as well as reducing prediction errors.

Table 3.

Number of wind speed forecast errors between ±5%, ±15% and ±30% for both predictors.

	2.5 sec		5 sec		7.5 sec	
	ANN	Proposed	ANN	Proposed	ANN	Proposed
±5%	76.3636	78.1818	34.2424	39.3939	25.4545	28.1818
±15%	23.0303	21.2121	48.4848	49.3333	44.2424	45.1212
±30%	0.6061	0.6061	16.9697	10.9697	26.3636	21.5455
	% of Errors in each margin					

For different prediction horizons, the required time for training of the network was less than 2 seconds using a conventional PC. So, it is obvious that the proposed predictor can be used in on-line application successfully.

4. Conclusion:

Very short-term wind speed predictor is a necessary tool in wind turbine control applications. In this study, a new predictor is presented involving ANN and Markov chain approach. The results show that the proposed method declined mean absolute percentage error (MAPE) and maximum prediction error (MPE) as well as uncertainty. Also, the proposed predictor successfully declined required time for calculations and so, it can be feasible for on-line applications. In the proposed predictor, over-training and extrapolation as the greatest problems for ANN in prediction application are considered and successfully are avoided. It is clear from diminishing maximum prediction error.

5. References:

[1] M. S. Miranda, and R. W. Dunn, *One-hour-ahead Wind Speed Prediction Using a Bayesian Methodology*, 2006.

[2] X. Wang, G. Sideratos, N. Hatziargyriou, L. H. Tsoukalas, *Wind Speed Forecasting for Power System Operational Planning*, In proceeding of 8th International Conference on Probabilistic Methods Applied to Power Systems, Iowa State University, Ames, Iowa, September 12-16, 2004.

[3] Riahy G, Freere P, *Dynamic wind turbine energy extraction showing the benefits of short term wind prediction*. In: Proceedings of the Solar'96 conference. Darwin, Australia, OCT. 1996

[4] G.H. Riahy, M. Abedi, *Short-term wind speed forecasting for wind turbine applications using linear prediction method*, Renewable Energy, Volume 33, Issue 1, January 2008, Pages 35-41.

[5] S. A. Pourmousavi Kani, G. H. Riahy, *A new ANN-based methodology for very short-term wind speed prediction using Markov chain approach*, In Proceeding of 8th IEEE Annual Electrical Power and Energy Conference, October 6-7, 2008, Vancouver, BC Canada.

[6] Lars Christian Henriksen, *Model Predictive Control of a Wind Turbine*, Technical University of Denmark, Information and Mathematical Modeling, 2007.

[7] Nanayakkara N.; Nakamura M.; Hatazaki H., *Predictive control of wind turbines in small power systems at high turbulent wind speeds*, Control Engineering Practice, Volume 5, Number 8, August 1997 , pp. 1063-1069(7).

[8] Carlos F. Moyano, João A. Pec¸ as Lopes, *An optimization approach for wind turbine commitment and dispatch in a wind park*, Electric Power Systems Research, Volume 79, Issue 1, January 2009, Pages 71-79.

[9] P. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. katsafados, I. Pytharoulis, G. Kallos, *Improvements in Wind Speed Forecasts for Wind Power Prediction Purposes Using Kalman Filtering*, Journal of Wind Engineering and Industrial Aerodynamics 96 (2008) 2348-2362.

[10] S. Salcedo-Sanz, A. M. Perez-Bellido, E. G. Oritz-Garcia, A. Portilla-Figueras, L. Prieto, D. Paredes, *Hybridizing the Fifth Generation Mesoscale Model with Artificial Neural Networks for Short-Term Wind Speed Prediction*, Renewable Energy (2008), doi:10.1016/j.renene.2008.10.017.

[11] C. Potter, M. Negnevitsky, *Very Short-Term Wind Speed Forecasting for Tasmanian Power Generation*, IEEE Transactions On Power Systems, VOL. 21, NO. 2, MAY 2006.

[12] P. Pinson, G. Kariniotakis, *On-line Assessment of Prediction Risk for Wind Power Production Forecasts*, Wind Energy. 2004; 7:119-132 (doi: 10.1002/we.114).

[13] M. Monfared, H. Rastegar, H. Madadi Kojabadi, *A New Strategy for Wind Speed Forecasting Using Artificial Intelligent Methods*, Renewable Energy 34 (2009) 845-848.

[14] S. Li, D.C. Wunsch, E.A. O'Hair, M.G. Giesselmann, *Using Neural Networks to Estimate Wind Turbine Power Generation*, IEEE TRANSACTIONS ON ENERGY CONVERSION, VOL. 16, NO. 3, SEPTEMBER 2001.

[15] M.A. Mohandes, T.O. Halawani, S. Rehman, A.A. Hussain, *Support vector machines for wind speed prediction*, Renewable Energy 29 (2004) 939–947.

[16] P. Flores, A. Tapia, G. Tapia, *Application of a Control Algorithm for Wind Speed Prediction and Active Power Generation*, Renewable Energy 30(2005) 523-536.

[17] T.H.M. El-Fouly, E.F. El-Saadany, M.M.A. Salama, *One Day Ahead Prediction of Wind Speed Using Annual Trends*, 2006.

[18] T.G. Barbounis, J.B. Theocharis, M.C. Alexiadis, P.S. Dokopoulos, *Long-Term Wind Speed and Power Forecasting Using Local Recurrent Neural Network Models*, IEEE TRANSACTIONS ON ENERGY CONVERSION, VOL. 21, NO. 1, MARCH 2006.

[19] E. Cadenas, W. Rivera, *Wind Speed Forecasting in the South Coast of Oaxaca, Mexico*, Renewable Energy 32 (2007) 2116-2128.

[20] E. Safavieh, A. Jahanbani Ardakani, A. Kashefi Laviani, S. A. Pourmousavi, S. H. Hosseinian, M. Abedi, "A new integrated approach for very short-term wind speed prediction using wavelet networks and PSO", International Conference on Power Systems, 12-14th December, 2007 Bangalore, India.

[21] S. A. Pourmousavi Kani, S. M. Mousavi, A. Kashefi Kaviani, G. H. Riahy, "A new integrated approach for very short-term wind speed prediction using linear regression among ANN and Markov Chain," International Conference on Power System Analysis, Control and Optimization, 13-15th March 2008, Andhra Pradesh, India.

[22] M.C. Mabel, E. Fernandez, *Analysis of Wind Power Generation and Prediction Using ANN: A Case Study*, Renewable Energy 33 (2008) 986-992.

[23] S. Salcedo-Sanz, A. M. Perez-Bellido, E. G. Oritz-Garcia, A. Portilla-Figueras, L. Prieto, F. Correoso, *Accurate Short-Term Wind Speed Prediction by Exploiting Diversity in Input Data Using Banks of Artificial Neural Networks*, Neurocomputing (2008), doi:10.1016/j.neucom.2008.9.010.

[24] H. Kantz, D. Holstein, M. Ragwitz, N.K. Vitanov, *Markov Chain Model for Turbulent Wind Speed Data*, Physica A 342 (2004) 315-321.

[25] A.D. Sahin, Z. Sen, *First-order Markov Chain Approach to Wind Speed Modelling*, Journal of Wind Engineering and Industrial Aerodynamics 89 (2001) 263-269.

A. Shamshad, M.A. Bawadi, W.M.A. Wan Hussin, T.A. Majid, S.A.M. Sanusi, *First and Second Order Markov Chain Models for Synthetic Generation of Wind Speed Time Series*, Energy 30 (2005) 693-708.

[26] C. Yin, L. Rosendahl, Zh. Luo, *Methods to improve prediction performance of ANN models*, Simulation Modeling Practice and Theory 11 (2003) 211-222.

[27] N. Siebert, *Development of Methods for Regional Wind Power Forecasting*, PhD Thesis, Sciences des Métiers de l'Ingénieur, 2008.