Hybrid Approach for Short Term Wind Power Forecasting

Vasanth Reddy[a], *Student Member, IEEE*, Samidha Mridul Verma[b], *Student Member, IEEE*, Kusum Verma[c], *Senior Member, IEEE*, Rajesh Kumar[d], *Senior Member, IEEE*

Department of Electrical Engineering, Malaviya National Institute of Technology, Jaipur, India

reddyvasanth41@gmail.com[a], 2015uee1434@mnit.ac.in[b], kverma.ee@mnit.ac.in[c], rkumar.ee@mnit.ac.in[d]

***Abstract*— Wind is one of the most important parts of renewable energy sources and optimal scheduling of wind power in wind farms is essential. Therefore, accurate prediction is a necessary task to be done in order to have a clear picture of how this wind energy can be utilized to its maximum potential. In this paper, wind speed is forecasted using it as a signal for wavelet transformation and the coefficients are predicted to obtain the forecasted wind speeds. The geographical location under study is taken at Jodhpur in Rajasthan, India. The performance evaluation of the proposed method is calculated using the different statistical error measures like RMSE, MAPE and MAE.**

Keywords—Wavelet; wind forecasting; wind farm; wind power.

I. INTRODUCTION

India has the fourth largest wind capacity in the world. Wind energy has a consistent scale for a long period of time but has a significant variation over shorter time-scales. It constitutes a significant amount of portion in its target of achieving world's largest renewable expansion of generating 175 GW by 2022. The state of Rajasthan has a 4.032 GW wind power installed as per [1] in the year 2016. Wind farms are costly to setup and therefore, short and long term forecasting of wind power is crucial. Long-term prediction doesn't give us the most accurate perspective as the wind can suddenly change at any point of time. Short term wind energy forecasting is therefore most important for optimal scheduling of the wind farms.

A lot of research has been carried out for wind speed forecasting using various methods. Numerical Weather Prediction (NWP) models [2] are mathematical models of atmosphere that are created to predict the future variables. Time series models [3] which comes under statistical methods [4] are models in which the relationship between the features is found out and used to predict data. Other approaches are based on Artificial Intelligence (AI) [5, 6] which predict the data without any predefined mathematical model. All these methods take various meteorological parameters into account to forecast wind power.

The wavelet transformation method has been explored in this paper because it gives us a clear picture of the given signal (in this case, wind speed data). The transformation provides us with two kinds of coefficients: approximate and detail coefficients, which can be assumed that they provide us information when the wavelet is "approximately" seen and the "details" in the change of the amplitude of the particular wave respectively, hence the name of the coefficients.

In this paper, the potential wind power for a short-term period (day-ahead hourly wind generation forecasts for an individual wind farm) has been predicted using a Hybrid Model. At the first stage, wavelet transformation is used to transform the wind speed variability over a specific period and coefficients are obtained by the wavelet decomposition. At the second stage, the coefficients obtained by wavelet transform are predicted for the next 24 hours using Markov Chain and Smoothing Spline and then, the wave is reconstructed. The proposed prediction model depends on one variable factor - wind speed, for a specific wind turbine. The autocorrelation was found out between all the meteorological variables and it was eventually concluded that there is a high correlation between them. Therefore, only one variable (wind speed) is taken into account. The trained model is verified using a part of real data and calculating the error between the predicted and the actual value. The geographical location under study is Jodhpur in Rajasthan, India. The meteorological data of wind speed for one year is obtained from Dark Sky API [7].

Section II gives a basic overview of the concepts used in implementing this hybrid approach. Section III includes the sequential methodology used in the paper and Section IV includes the simulation and results obtained. The paper is concluded in Section V.

II. TERMINOLOGY

A. Wavelet

Wavelet is a pulse that begins (visualizing it in Cartesian coordinates) at origin and oscillates to and fro, from the origin point. Wavelet transform is described in terms of wavelets. As there is lot of variability in wind speed throughout the year, these wavelets provide us a clear information about the wind speed data. These wavelets are similar in nature but differ in dilation and translation. The wavelet basis functions are self-similar: scaled in time to maintain the same number of oscillations and scaled in amplitude to maintain energy. This transformation helps in filtering inconsistent data and the data points can be observed clearly. [8]

Wavelet transformation can be of two types: Continuous Wavelet Transformation (CWT) and Discrete Wavelet Transformation (DWT).

The Continuous Wavelet Transformation (CWT) of a signal is given by

$$F(a,b) = \int_{-\infty}^{\infty} f(t)\varphi^*_{(a,b)}(t)dt \qquad (1)$$

where * is the complex conjugate function, function φ is some function, a is the dilation variable, b is the translation variable which determines the central position. The F(a,b) coefficient represents how well the original signal *f(t)* and the dilated/translated mother wavelet match. Thus, these wavelet coefficients associated to a particular signal, is the wavelet representation of the signal with respect to the mother wavelet. This transformation is attained by continuous translation and dilation of mother wavelet which results in generation of abundant information. As more coefficients are generated which are not required in recovering signal, so Discrete Transformation is chosen which is more efficient than and as accurate as CWT. DWT is attained by transforming:

a = 2^{-x}

b = y.2^{-x}

where x, y are scale and translation respectively, they are discrete in nature and can only take integer values.

Mallat Algorithm is an efficient way to implement the wavelet transform. This algorithm is based on four filters (low pass decomposition, high pass decomposition, low pass reconstruction, high pass reconstruction) [8]. The original signal is made to pass through low pass filters and high pass filters, which gives approximate and detail coefficients. Approximate coefficients are obtained when the signal is passed through the low pass filters whereas detail coefficients are obtained when passed through high pass filters.

Approximate coefficients are further broken down into lower resolution components. Thus the multilevel decomposition is achieved. In this paper, Daubechies wavelet function is used for decomposing till level 3. Here, one level of approximate coefficients A3 and three high level detail coefficient D1, D2, D3 are obtained.

B. Markov Chain

Markov Model is a stochastic model which is used to capture randomly changing systems on the basis of Markov Property. The Markov property holds when the future state is dependent only upon the current state of the system and not on any of the historical states. Markov chain process is the simplest of the Markov Models. Markov chain models the state of a system with a single random variable that changes through time. In this context, the Markov property suggests that the distribution for this variable depends only on the distribution of previous state.

Let S_t $\{t = 0, 1, 2, 3, \ldots, n\}$ be a stochastic process with a finite number of values of t. Let the total number of states be N. If $S_t = i$, it means that the process is in state i at a time t. Let's say in general that if the process is moving from one state i to another state j, then the probability of changing from one state to another state is given by $P_{i,j}$. More specifically, $P_{i,j}$ denotes the probability of the system changing from state i at at time t to state j at time $t+1$.

Transition probability is given by:

$$P\{S_{t+1} = j \mid S_t = i, S_{t-1} = i_{t-1}, \ldots, S_1 = i_1, S_0 = i_0\} = P_{i,j} \qquad (2)$$

where $t \geq 0$ for all states, this process is known as Markov chain.

The state at S_{t+1} does not depend upon states at S_0, S_1, …, S_{t-2}, S_{t-1} but only on previous state at S_t. Transition matrix or Markov matrix is a square matrix which is used to represent the transitions of the Markov process. Each entry in the transition matrix is a non-negative number. Since the total of transition probability from a state i to all other states must be 1, the sum of each row is 1.

$$\sum_{j=1}^{N} P_{i,j} = 1 \qquad (3)$$

$$P_{i,j} = \begin{pmatrix} P_{1,1} & \cdots & P_{1,N} \\ \vdots & \ddots & \vdots \\ P_{N,1} & \cdots & P_{N,N} \end{pmatrix} \qquad (4)$$

Higher order Markov chains are Markov chains that use more than one preceding states in the process of Markov chain modelling. For example, a second order Markov Chain would use two preceding states to obtain the transition probability matrix. Higher order Markov chain is useful for efficient prediction in the wind data as they provide additional memory regarding the data, which is same as increasing the features in training the model [9].

C. Smoothing Spline

Smoothing splines are function estimates obtained from a set of noisy observations of the target in order to balance a measure of goodness of fit of with a derivative based measure of the smoothness. This method smoothens the edge curves of the signal and fit the data.

Let us consider the time series X = $\{x_1, x_2, \ldots, x_i, \ldots, x_n\}$, where x_i is the wind speed at a time t_i and each x_i has two components; smoothing function and normally distributed noise.

$$x_i = g(t_i) + e_i \qquad (5)$$

where g is the unknown smoothing function and these must be the best approximation to minimize

$$\int_{x_1}^{x_n} g''(x)^2\, dx \qquad (6)$$

Smoothing Function is given by

$$F(g) = \sum_{i=1}^{n} \left(\frac{xi - g(t_i)}{\sigma_i}\right) + \frac{\lambda}{2} \int_{t_1}^{t_n} g''(t)^2\, dt \qquad (7)$$

λ is the smoothing parameter, chosen such that it minimizes the least square error. This function smoothens the signal and fits the data.

III. Proposed Methodology

A. Categorizing the Wind Data into Groups

The data here used is collected from a reliable website called Dark Sky [7] for the year 2015. It is observed from the data for one year that there is a lot of variability and inconsistency. This problem can be sorted out by grouping the consecutive months into one category (i.e based on the average wind speed of the month).

B. Wavelet Transform

The wind speed data is passed through wavelet transformation filter. Approximate coefficients and detail coefficients are obtained from low pass and high pass filters respectively. In this paper, decomposition is done till third level. Thus, one level of approximation coefficients and three different level detail coefficients are obtained. Let's say, *n* values of wind speed are taken. In the first level of decomposition, *n/2* approximate coefficients and *n/2* detail coefficients are obtained. During further decomposition, *n/4* approximate and *n/4* detail coefficients are obtained and in the last decomposition stage, *n/8* approximate and *n/8* detail coefficients are obtained. At the end of multilevel decomposition *n/8* approximate coefficients and *n/2*, *n/4*, *n/8* details coefficients are obtained at first, second and third levels respectively. In later stages, these coefficients are passed through different models for prediction.

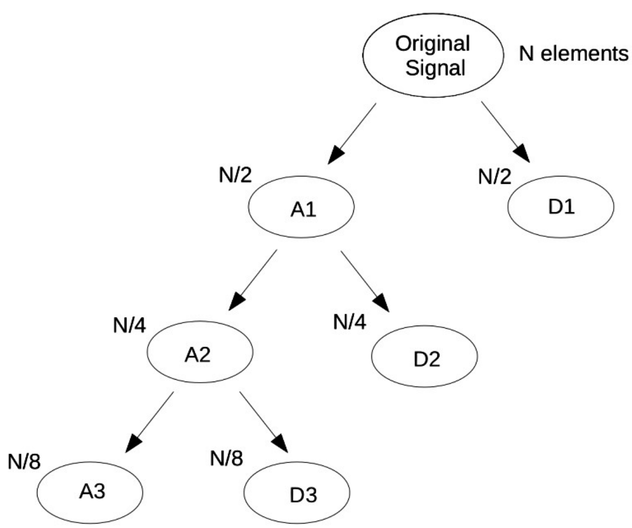

Fig.1. Multi-dimensional decomposition

C. Markov Chain for Prediction

1. Categorization of States (States based on mean and standard deviation)

Let *m* be the mean and *s* be the standard deviation of the wind speed data. Then the first state starts for the values that lie in the range (0, *m* - *s*], then state as 2 for the values lying in the range (0, m] and so on… Thus, 6 states are obtained in doing the above process.

2. *Transition probability matrix estimation*

For a given time series of the wind speed V_t (t = 0, 1, 2, …, T), estimation of the transition probability matrix/Markov matrix of a matrix of order 1 is a simple process. Let $R_{i,j}$ be the number of speeds/data that are in state, S_i during the time period t and in state S_j during the time period t+1, for t between 0 and $T-1$. Then the $P_{i,j}$ is given by the mathematical expression

$$P_{i,j} = R_{i,j} / \sum_{j=1}^{T} R_{i,j} \quad (8)$$

It is known that above expression is likelihood estimator which tends to zero for large sample [10].

3. Prediction

The first and foremost step is to calculate the transition probability matrix (both First order and second order). For any random number from the dataset, the following state can be determined. First the wind speed is checked. If the wind speed is not zero, then the random number from the specific state interval is used to generate a value ([10]). For the intermediate state, a random uniform value is generated, it is taken from that specific state interval range ([11], [12]). For higher state values, gamma parameter [13] is used to generate the wind speed.

D. Prediction of Approximate and Detail Coefficients

In this step, the approximate and detail coefficients are for the next 24 hours (= n). For multilevel transformation (db1, level = 3), there is a need to predict 3 (= n/8) approximate coefficients and 3 (= n/8) detail coefficients at level 3. At level 2, 6 (= n/4) detail coefficients are predicted and at level 1, 12 (= n/2) detail coefficients is predicted. Markov chain is an accurate method to predict short term data. Even small variation in predicted approximate coefficients can result in large changes in the original signal. Therefore, Markov Chain is used in this paper to predict level 3 approximate coefficients (markov chain is explained above). For predicting the detail coefficients, the smoothing spline method is used, which predicts the future data based on past values. After the prediction, the wavelet is reconstructed with the above predicted values to get forecasted wind speed for 24 hours.

IV. Simulation and Results

There is a lot of variability and inconsistency in the wind speed data throughout one year. It would be inefficient to train the whole year's data all at once. It's better to group it into seasons or categories and then train them individually. It can be done by taking consecutive months into one category based on the similarity of average wind speeds of the month.

The meteorological data here used is collected from Dark Sky [7] for the year 2015. The hourly data for wind speed for the location of Jodhpur (26.2389° N,73.0243° E) for the year 2015 is extracted. The box plot of the wind speed data is given below. In can observed from the graph from fig. 2 that some of the months should be categorized into groups. The five groups according to this:

Group 1 - January, February, March, April
Group 2 - May, June
Group 3 - July
Group 4 - August, September
Group 5 - October, November, December

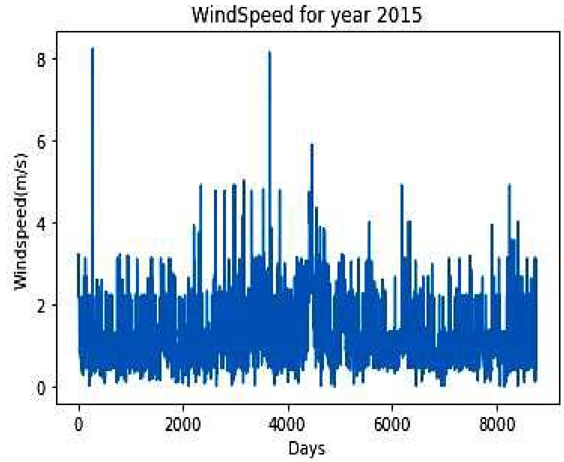

Fig.2. Wind speed for the year 2015

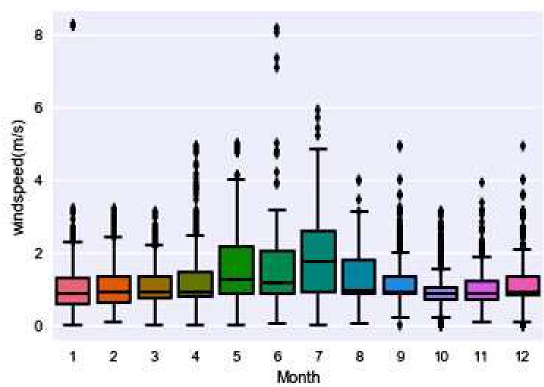

Fig.3. Average wind speed for year 2015

Each group is trained separately for prediction. For numerical accuracy assessment of the wind speed forecasts, different error measurements are used. Root-Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are employed for evaluating the forecasting accuracy. The prediction error of a model is defined as the difference between the predicted values and the measured data. Let $G(t)$ be the predicted value and $M(t)$ be the measured value. N be the length of total forecast horizon at time t.

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N} (G(t) - M(t))^2} \quad (9)$$

$$MAPE = \frac{1}{N}\sum_{j=1}^{N} \frac{|G(t) - M(t)|}{M(t)} \quad (10)$$

$$MAE = \frac{1}{N}\sum_{j=1}^{N} |G(t) - M(t)| \quad (11)$$

Mean Absolute Error (MAE) is the easiest to interpret. In Root Mean Square Error (RMSE), the errors are turned positive by raising it to the square, with the advantage of being easy to handle. The Mean Absolute Percentage Error (MAPE) calculates the percentage error and compares two models. MAPE provides an indicator of percentage deviation of the predicted values with the real values of the time series.

TABLE 1: ACCURACY MEASUREMENTS OF EACH GROUP

Date	RMSE	MAPE	MAE
April 1st	0.06589	0.2699	0.1860
July 31st	0.10414	0.1836	0.3837
Dec 31st	0.06121	0.0848	0.17922

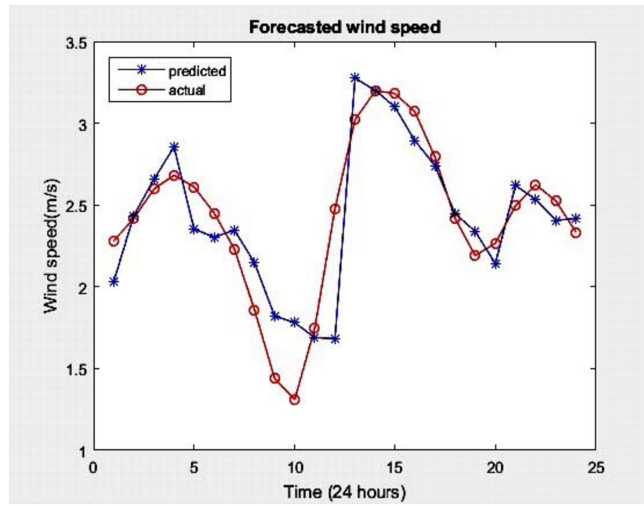

Fig.4. Forecasted and actual wind speeds

- *Wind Turbine Specifications*

For analysing the wind energy potential of the study area the data of a typical wind turbine is taken. Power output from a wind turbine is directly proportional to third power of wind speed. Each turbine model has a unique site-dependent power curve which shows relation between wind speed and power. Due to high variability in wind speed, the turbine power output changes continuously [14]. Table 1 shows the specifications of a SUZLON model S97-2.1 MW wind turbine, where these turbine model is located in present wind farm site.

Using the above specifications, wind power generated can be found out by using the below equation

$$P = 0.5\, \rho A C_p V^3 \quad (12)$$

where P is the power in W, V is the wind speed in m/s, C_p is the dimensionless factor i.e., power coefficient, ρ is the air density in kg/m3, A is the turbine rotor area in m^2. Fig.5. shows the estimated wind power output in kW for the turbine for one day (24 hrs) on 31 December 2015 for the geographical location under study.

TABLE 2: SPECIFICATIONS OF THE WIND TURBINE

Specifications	**S 97**
Rated Power	2100 kW
Rotor Diameter	97 m
No. of Blades	3
Swept area	7386 m^2
Air density	1.225 kg/m^3
Rotational speed	variable 12-15.7 rpm
Cut in wind speed	3.5 m/s
Rated wind speed	11 m/s
Cut out wind speed	20 m/s

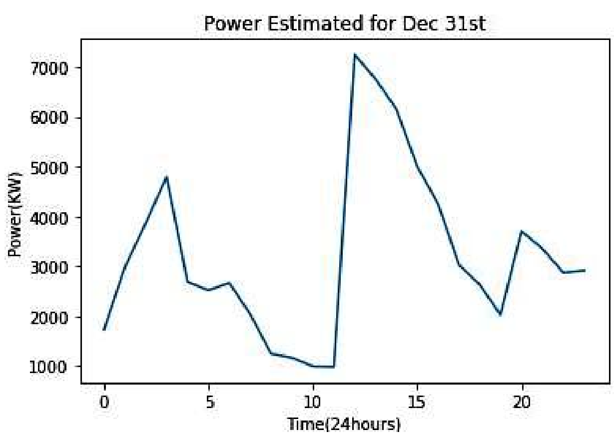

Fig.5. Estimated power for 24 hours

V. CONCLUSIONS

In this paper, Wavelet transformation is used in predicting short term day-ahead wind speeds for wind farms. With proper analysis, the wind speed data can be grouped into different seasons/groups so that it can be trained and processed most efficiently. Only one variable – wind speed, is used for training. The model is validated by using real measured data for the duration of one year (2015) for the region Jodhpur, Rajasthan, India. The results are measured by calculating the error, the difference between the real data and the predicted data, using performance parameters such as RMSE, MAE and MAPE. It is seen from the results that the proposed hybrid methodology gives good performance and can be suitably employed for short term wind power forecasting in the wind farms.

ACKNOWLEDGMENT

This work is supported by the Department of Science and Technology (DST) Rajasthan, India. The authors would like to acknowledge the Department of Science and Technology Rajasthan for the financial support under Grant R&D/2016/2593.

REFERENCES

[1] MNRE (Ministry of New and Renewable Energy, Government of India) http://mnre.gov.in/
[2] L. Landberg, "A Mathematical Look at a Physical Power Prediction Model": Wind Energy Volume: 1, Issue: 1, September 1998, pp. 23 ± 28.
[3] Peiyuan Chen, Troels Pedersen, BirgitteBak-Jensen, Zhe Chen "ARIMA - Based Time Series Model of Stochastic Wind Power Generation", IEEE Transactions On Power Systems, Vol. 25, No. 2, May 2010
[4] Doan Duc Tung and Tuan-Ho Le."A Statistical Analysis of Short-Term Wind Power Forecasting Error Distribution", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 10 (2017) pp. 2306-2311
[5] P.C. Răzuşi, M. Eremia, "Prediction of Wind Power by Artificial Intelligence Techniques", 16th International Conference on Intelligent System Applications to Power Systems (ISAP) - Hersonissos, Greece (2011.09.25-2011.09.28)] 2011
[6] Yuan-Kang Wu, Jing-Shan Hong, "A literature review of wind forecasting technology in the world", Power Tech, 2007 IEEE Lausanne, Lausanne, Switzerland, 1-5 July 2007
[7] Dark Sky API, available online: https://darksky.net/dev
[8] Khan, A. A., and M. Shahidehpour. "One day ahead wind speed forecasting using wavelets." Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES. IEEE, 2009.
[9] Kevin Brokish, James Kirtley, "Pitfalls of Modeling Wind Power Using Markov Chains",Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES
[10] Shamshad A, Bawadi MA, Hussin W, Majid TA, Sanusi SAM. "First and second order Markov chain models for synthetic generation of wind speed time series", Energy 2005;30(5):693e708.
[11] A.Carpinone, M.Giorgio, R.Langella, A.Testa."Markov chain modeling for very-short-tern wind power forecasting", Electric Power Systems Reserch 122 (2015) 152-158
[12] Nfaoui H, Essiarab H, Sayigh AAM. "A stochastic Markov chain model for simulating wind speed time series at Tangiers, Morocco", Renew Energy 2004;29(8):1407e18.
[13] Hafzullah Aksoy, Z. Fuat Toprak, Ali Aytek, N. Erdeml: "Stochastic generation of hourly mean wind speed data". Renewable Energy 29(14), 2004, pp. 2111–2131.
[14] Chinnu Mariam Baby, Kusum Verma, Rajesh Kumar. "Short Term Wind Speed Forecasting and Wind Energy Estimation : a Case Study of Rajasthan". 2017 International Conference on Computer, Communications and Electronics (Comptelix) Manipal University Jaipur, July 01-02, 2017.