PROBABILISTIC MODELING AND FORECASTING OF WIND POWER

ANURADHA M1*, B. K. KESHAVAN1, T. S. RAMU2 and V SANKAR3

1PES Institute of Technology, 100ft. Road, BSK III Stage, Bangalore, Karnataka -560 085
2PES Institute of Technology, 100ft. Road, BSK III Stage, Bangalore, Karnataka -560 085
3Indian Institute of Science, C V Raman Ave, Bangalore, Karnataka 560012

(Received on December 22, 2015, revised on March 24, 2016)

Abstract: Modeling of wind power is essential for an effective management and balancing of a power grid, supporting real-time operations. Forecasting the expected wind power production would help to deal with uncertainties. The data driven approach for forecasting is expected to give detailed information on the system and real time measurements. Wind being a natural phenomenon, probabilistic methods need to be employed in generated wind power, based on previous history of the system. In this paper, the data on the wind speed and power generated from a location in the state of Karnataka, India, has been analyzed for the duration of three months. It has been shown that the probability distribution of wind speed conforms closely to Rayleigh distribution. It is expressly demonstrated that, while the wind speed conforms to Rayleigh distribution, the electrical power developed follows a Weibull distribution in two parameters. Besides using graphical methods for estimating the Weibull parameters, Maximum likelihood equations are set up to estimate the parameters. These parameters have been used in estimating / forecasting of wind power using both Weibull algorithm as well as the Monte Carlo Method.

Keywords: Wind Power Forecasting, Rayleigh, Weibull, Gaussian distributions, Likelihood equations, Monte Carlo algorithms

1. Introduction

A considerable progress has been made in the site location, design and installation of wind energy systems to augment conventional energy sources. An important requirement for this purpose is the knowledge of spatial and temporal wind pattern, its velocity and direction. This information is vital in the location of the wind turbine systems. At the outset, it is to be understood that the wind velocity is a location specific stochastic variable [1]. Therefore, one needs to treat it as a random variable conforming to a known distribution. Perusal of recent literature on the topic indicates that, the magnitude of wind velocity follows an extreme value distribution of the smallest type, the Weibull distribution, in two parameters [2, 3, 4]. This consideration seems to be inappropriate as the random behavior of the wind velocity is not so extreme. Admittedly, the assumption is data driven and substantiation to that effect is essential. Also, more often the distribution of velocity data can be seen by observation, to be a Rayleigh or one similar to it.

*Corresponding author’s email: anuradha@pes.edu
Forecasting the possible wind power generating capacity over a time window subsequent to the observed data is of particular interest to power utilities. This will greatly help load rescheduling besides improving the short term system reliability [5]. The observed wind speed data is location specific and has the characteristics of a time series - a discrete random sequence of natural events in time. Forecasting, therefore, entails a prediction paradigm with a large scatter and often, contradictory uncertainties. In the recent past, proven probabilistic methods have been developed to address this issue [6,7]. One of the methods of forecasting wind power is to use the Weibull parameters estimated from sample data from an actual wind generating system and employing a Monte Carlo type simulation procedure [8,9]. In this paper, the authors have shown that the wind velocity follows the Rayleigh distribution [10]. An important aspect of this contribution is to examine the stochastic nature of the power generated by a wind energy system using the probabilistic wind velocity data. Authors show that, given a random variable of the nature encountered here, conforming to a certain conceived distribution, the corresponding distribution of a function of this random variable is very different and needs extensive and somewhat rigorous mathematical treatment and this has been demonstrated.

A need has been increasingly felt, to try and develop, although empirically, models for forecasting of wind power at an epoch of time, beyond the observation window. However, this is a difficult task, given the degree of scatter in the location specific wind speeds. Wind speeds conform to a nearly stationary time series characterized by a strong seasonality. Notwithstanding the effects of seasonality, it is believed that the wind speed, and hence the power generated over a given time window, can be forecast with a reasonable degree of certainty.

In this paper an attempt has been made to predict the power output in wind energy systems. The exercise in forecasting random events from out of a time series is generally fraught with irreconcilable uncertainties. A phenomenological Weibull model as shall be detailed later in this paper seems to offer a method of circumventing the said complications. Estimation of forecasting uncertainties has also been made to validate the model.

The organization of the paper is as follows:
Section 2 discusses various statistical distributions used to fit the observed wind speed data, modeling of wind power and various methods of parameter estimations. Section 3 deals with the results of the forecast, analysis of the plots of several relations and estimation of error in wind power forecast. In section 4 the discussion on various results are highlighted. Finally section 5 concludes the work carried out and the future scope.

2. Theoretical Aspects

In the development of the mathematical models described in this paper, the direction of the wind is seemingly less important. Hence its scalar counterpart, the wind speed, may conveniently replace the velocity, a vector quantity. This stems from the fact, that, the wind turbine blades are designed to be self – seeking / adjusting, in so far as the direction is concerned. It is therefore sufficient to model the location specific wind speed probabilistically [11, 12].
2.1 Probability Distribution Functions

Let the continuous random variable, \( X \) representing the wind speed, be \( v \), thus, \( X \rightarrow v \), with its domain of definition, \( 0 \leq X \leq \infty \).

Also, by definition, \( \Pr[X = x] = f(v) \) \hspace{1cm} (1)

Where, \( f(v) \) is the probability density function (p.d.f) of and the corresponding cumulative distribution function (c.d.f), is defined by;

\( \Pr[x \leq v] = F(v) = \int f(v)dv \), if it exists,

\( f(v) = F(v) = \frac{d}{dv} F(v) \hspace{1cm} (2) \)

It is well known that, while a p.d.f shall always exist for nearly all values of the random variable, a c.d.f may not be defined over a certain interval and that it may not be possible to represent it in a closed mathematical form. A classic example is the Normal or Gaussian probability distribution.

2.1.1 The Rayleigh Distribution

A Rayleigh distribution is often found valid when the sub-component of a random variable bears a strong functional relationship with others [13, 14]. One example where the Rayleigh distribution naturally arises is when wind velocity is analyzed into its associative orthogonal vector components; the components themselves are related to their directional cosines. The c.d.f of Rayleigh distribution is given by;

\[ F(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}, \quad x \in [0, \infty) \] \hspace{1cm} (3)

2.1.2 The Weibull Distribution

This distribution is associated with the times to failure of nominally identical objects and is thought to describe the smallest time (from among a large sample size) before a failure arises.

The p.d.f of a Weibull distribution in two parameters, the shape (\( \beta \)) and scale parameters (\( x_0 \)), is given by,

\[ f(x; x_0, \beta) = \frac{\beta}{x_0} \left( \frac{x}{x_0} \right)^{\beta-1} e^{- \left( \frac{x}{x_0} \right)^\beta} , \quad \beta, x_0 > 0, \quad 0 < x \leq \infty \] \hspace{1cm} (4)

Upon substituting \( \beta = 2 \), the Weibull distribution transforms itself to the Rayleigh distribution. While \( \beta \) indicates the skewness, \( x_0 \) is called the nominal value, the characteristic value or the 63.2\textsuperscript{th} percentile or the scale parameter of the distribution. The c.d.f of the Weibull distribution is;

\[ F(x) = 1 - \exp \left( -\frac{x}{x_0} \right) ^\beta , \quad 0 < x < \infty, \quad \beta, x_0 > 0 \] \hspace{1cm} (5)
2.2 Probabilistic Modeling of Wind Power

The power contained in the wind is given by the kinetic energy of the flowing mass per unit time.

\[ P = \frac{1}{2} \rho A v^3 \]  \hspace{1cm} (6)

Where, \( \rho \) is the air density \( \equiv 1.225 \text{kg/m}^3 \), \( A \) is the rotor area \( \text{(m}^2) \), \( v \) is the wind velocity \( \text{(m/sec)} \) without rotor interference.

Since the stochastic power generated, \( P \) is a power law in velocity with an exponent 3 (Cubical law), \( P \) needs to be deduced by a transformation process. In a measure theoretic axiomatic approach to probability, if \( X \) is a random variable in a Borel Space, \( \mathcal{B} \Rightarrow \mathbb{R} \), then, a function of \( X \), say, \( Y \) and its distribution, is also a random variable in \( \mathcal{B} \Rightarrow \mathbb{R} \).

If \( X \) has a known c.d.f, the c.d.f of \( Y \) can be derived by admitting a real Borel measure; \( g: \mathcal{B} \Rightarrow \mathbb{R} \) to the probability of the random variable \( X \).

If the c.d.f of \( Y \) is \( Y = g^{-1}(y) \) is differentiable for all, then, its p.d.f can be generalized as under:

\[ f_y(y) = f_x[g^{-1}(y)] | \frac{dg^{-1}(y)}{dy} | \]

Now, in the present case, let the random variable \( X \) be represented as a power function in \( v \), be written as,

\[ Y = bX^2 \]

with \( b \) is a constant and \( Y \) is identified with Power \( P \), and \( X \) with wind velocity, \( v \)

\[ X = \left( \frac{y}{b} \right)^{\frac{1}{3}} \]; \[ G(y) = \left( \frac{y}{b} \right)^{\frac{1}{3}} \int_0^x \frac{x}{\sigma^2} \exp\left[ -\left( \frac{x^2}{2\sigma^2} \right) \right] dx \]; \hspace{1cm} \text{with} \ z = \frac{x^2}{2\sigma^2} \]

\[ G(y) = 1 - \exp\left[ -\left( \frac{y}{b} \right)^{\frac{1}{3}} \right] \]; \[ g(y) = \frac{1}{b^2} \left[ 1 - \exp\left[ -\left( \frac{y}{b} \right)^{\frac{1}{3}} \right] \right] \]

It is easy to show that \( g(y) \) has a form given by,

\[ g(y) = \left( \frac{y}{b} \right)^{\frac{2}{3}} \left( \frac{1}{3b} \right) \exp\left[ -\left( \frac{y}{b} \right)^{\frac{1}{3}} \right] \]  \hspace{1cm} (7)
On comparing the above expression in random variable $x$, with p.d.f of Weibull distribution,

$$f(x; x_0, \beta) = \frac{\beta}{x_0^\beta} \left(\frac{x}{x_0}\right)^{\beta-1} e^{-\left(\frac{x}{x_0}\right)^\beta}; \beta - 1 = -\frac{2}{3}; \beta = \frac{1}{3}; b = x_0 \text{ and } y = x$$

$$\therefore \left(\frac{x}{x_0}\right)^{\beta-1} = \left(\frac{x}{x_0}\right)^{-\frac{2}{3}} \quad (8)$$

This can be seen as the Weibull distribution in two – parameters with $\beta = \frac{1}{3}; x_0 = b$

2.3 Parameter Estimation

Several methods are indicated in processing random data as detailed below.

1. Graphical Method
2. Maximum Likelihood Estimation (MLE)
3. Monte Carlo Algorithm
4. Best Linear Unbiased Estimation (BLUE)
5. Best Linear Invariant Estimation (BLIE)

However, in this paper, the Graphical method, Maximum Likelihood Estimation and Monte Carlo Algorithms are used for estimating and forecasting the wind power and are explained in the following sub sections.

2.3.1 Graphical Method

It is often convenient to make a rough estimation of the parameters using Graphical methods before analytical techniques are used in their precise estimation, such as the method of MLE. The graphical estimates are used as seed values in MLE \(^{(15)}\).

The c.d.f of the two-parameter is used as the governing equation for this purpose. Starting from the expression for c.d.f in the random variable $X = x$;

$$F(X|x_0, \beta) = 1 - \exp\left(-\left(\frac{x}{x_0}\right)^\beta\right); \beta > 0; x_0 > 0 \quad (9)$$

By a natural logarithmic transformation, this expression can be re-written as;

$$\ln\ln\left(\frac{1}{1 - F(x)}\right) = \beta\ln x - \beta\ln x_0 \quad (10)$$

This is a linear equation in $\ln x$. The transformed c.d.f $F(x)$ can be expressed as

$$Y = mx - C \quad \text{in which} \quad Y = \ln\ln\left(\frac{1}{1 - F(x)}\right); \quad (11)$$

$$X = \ln x; \quad \beta = m; \quad C = \ln x_0;$$
\[ Y_i \] are order statistics and can be computed using \( F(x_i) = \frac{i}{n} \) for \( x_i \)

Where, \( x_i \) is the \( i^{th} \) ordered observation and \( n \) is the sample size. \( x_i \) is got by the inverse transformation as

\[ x_i = \exp \left( \frac{Y_i}{\beta} \right) + \ln x_0 \]  \hspace{1cm} (12)

### 2.3.2 Likelihood Estimation Technique

Maximum Likelihood Estimation (MLE) is among the more powerful methods of estimating the parameters of any distribution. The p.d.f of the distribution representing the Pr\( [X = x]\) shall be known to make further calculations. The MLE has many optimal properties: sufficiency, consistency and efficiency. Many of the inference methods in statistics are developed based on MLE. The MLE’s are essentially asymptotically unbiased estimators (large sample estimators) based on a preconditioned uniform prior. It can be shown that the estimator is the mode of the posterior distribution \([16, 17]\).

Let \( x_1, x_2, x_3, \ldots x_n \) be the samples of independent and identically distributed observations coming from a distribution with an unknown probability density function \( f_0 (\cdot) \). It is desirable to find an estimator \( \hat{\theta} \) which would be as close to the true value \( \theta_0 \) as possible.

Both the observed variables \( x_i \) and the parameter \( \theta \) can be vectors, thus;

\[ L(\theta | x_1, x_2, x_3, \ldots x_n) = f(x_1, x_2, x_3, \ldots x_n | \theta) = \prod_{i=1}^{n} f(x_i | \theta_j) \]

The likelihood function, \( L \), can be maximized by equating the partial derivatives of \( L \) w.r.t \( \theta_j \).

\[ \frac{\partial L}{\partial \theta_j} = 0 \]  \hspace{1cm} (13)

This technique can be applied to the case of Weibull whose p.d.f is given by Eqn. (4).

The partial derivative, (Eqn.12) results in a pair of non-linear equations, the Log Likelihood equations, thus

\[ x_0^\beta = \frac{1}{n} \sum_{i=1}^{n} x_i^\beta \]  \hspace{1cm} (14)

\[ \frac{1}{\beta} = \frac{1}{n} \sum_{i=1}^{n} x_i^\beta \ln x_i - \frac{1}{n} \sum_{i=1}^{n} \ln x_i \]  \hspace{1cm} (15)

These two nonlinear simultaneous equations can be solved by iterative numerical methods such as predictor corrector method. The parameters so obtained are used in making forecasts beyond the time window over which the data is available.
2.3.3 Application of Monte Carlo Algorithm in Forecasting

One of the possible methods of forecasting randomly changing wind power is to recognize as a time series and apply Monte Carlo Algorithm, as described below.

Consider a standard exponential c.d.f given by, \( F(x) = 1 - e^{-x} \) and can be transformed into a Weibull c.d.f using the transformation,

\[
X = \left( \frac{W}{x_0} \right)^{\beta}
\]

in which \( W \) is the random variable characterizing Weibull distribution.

Suppose a random variable \( U \), is uniformly distributed in the interval \([0,1]\) then the random variable is distributed as Weibull with parameters \( x_0 \) and \( \beta \) given by

\[
W = x_0 [\ln(-\ln(U))]^{\beta}
\]

It may be seen that \( \ln(U) \) represents \( X \) and this expression can be employed to generate random Weibull variates with parameters \( x_0 \) and \( \beta \). The variates can be computed by generating uniform random numbers between 0 and 1. The forecast wind power calculated by the simulation procedure using \( x_0 \) and \( \beta \) of the observed wind power can be thought of as a subset of the original population comprising the actual wind power data and the forecast wind power. It is to be noted, that, this simulation procedure gives a time series similar in form and content to the actual data and hence the forecast can be taken as an extension of time series beyond the observation time window. It is possible to re-construct the time series in the observed time window also. However, this series cannot be equated point-by-point to the actual time series. The reason being, the reconstruction is a Monte Carlo procedure that shows a global but not a local similarity. In fact, any reconstruction/forecast procedure entails a certain measure of uncertainty insofar as the point-to-point comparison is concerned. It is also important to note that different sets of random numbers are to be used in every new simulation procedure.

3. Results and Analysis

The data used in this paper is, obtained from a site in Karnataka, India. About 54,000 wind velocity readings are taken over a full year at intervals of 10 minutes. The data is collected at a hub height of 40m and is truncated below cut-in and above furling speeds. The spectrum of the wind speed is chosen from 3m/s to 14m/s for which the variation of power is from 14.2MW to 503.8MW

3.1 Distribution of Wind Speeds

The wind speeds are fitted to three commonly used statistical distributions, the Rayleigh, the Weibull and the Normal. The Normal pdf plot of wind velocity is shown in Fig. 1, indicating the mean value of 5.1
It is observed, that the closest fit is the Rayleigh. Also, the wind power generated is proportional to the cube of the wind speed. This follows, as has already been shown, a Weibull distribution in two parameters. The Rayleigh p.d.f plot of wind velocity is, as expected, skewed towards left from the mean value as shown in Fig. 2. The Rayleigh distribution is a special case of Weibull distribution and is often used in fitting non-extremal data.
The Fig. 3 shows the Weibull cdf plot of wind velocity. The Shape ($\beta$) and the Scale ($x_0$) parameter of the model as determined from the plot are found to be $\beta = 3.1$ and $x_0 = 6.31$

From Fig. 4 it is observed that the Weibull probability density function of wind velocity graph is skewed towards right with a small skewness from the mean value and found to be close to the Gaussian distribution.

3.2 Prediction of future time series in Wind Power forecasting

The observed wind power generated over a period of 81 days, sampled daily and the possible power producible (forecast) has also been carried out. It can be seen that the forecast seems to be in the expected order. The simplest way to forecast the realizable wind power is to reconstruct the power, treating it as a random variable. Expressions

$$X = \left(\frac{W}{x_0}\right)^\beta$$ and $$W = x_0[-\ln(U)]^{1/\beta}$$ are used to obtain the stochastic nature of reconstructed (estimated) and forecast wind power.

In Fig. 5, the plots of observed, forecast and re-constructed (estimated) wind power are shown. While the forecast and original data, Fig. 4(a) and 4(b) show similarity in many
respects, the reconstructed power in the observed time window, Fig.4(c), seems somewhat different. Possible explanation for such happenings is given in the discussion section.

Given a set of random wind power data, assuming that the wind power data comes from Weibull population, the parameters are estimated by both Graphical and MLE methods. Also, the data has been reconstructed as ordered observation in the data spectrum; forecast has been made for 81 days beyond existing data spectrum.

Fig. 6 shows the Weibull plot of observed wind power over a period of 81 days, sampled daily. A visual examination of the fit indicates that the Weibull plot is nearly linear. The shape parameter (β) and the scale parameter (x₀) of the model are determined using the plot and the values are found to be β = 1.4 and x₀ = 12.3 respectively.

Fig. 7 shows the Weibull plot of estimated wind power generated over a period of 81 days in the observed time window. The graph shows a pronounced curvature and hence a possible existence of a third parameter. The shape parameter (β) and the scale parameter (x₀) of the model are 1.4 and 12.9 respectively.

In Fig. 8 the Weibull plot of forecast wind power over a period of 5 days beyond the actual observation period using the expression (Eqn.12) derived from Weibull c.d.f is shown. The shape parameter (β) and the scale parameter (x₀) of the model are determined using the plot and the values are found be 1.4 and 13.

\[
x₀(\ln(\text{windpower})) = 12.3; \beta = 1.4
\]

\[
x₀(\ln(\text{windpower})) = 12.9; \beta = 1.4
\]
Long time forecasting of time series is a highly sought after paradigm. However, it is well known that the uncertainties involved in forecasting the time series wind power data beyond the observed time window is only possible if the series is stationary in time. Stationarity implies that the variance of the time series remains invariant with respect to the inspection time window, or the lag as it is often called. Even when the stationarity is guaranteed, the estimation involves large errors in prediction besides associated poor repeatability as well as reproducibility. Hence, the authors have restricted the prediction to 5 days beyond the observed period.

\[ x_0 (\ln(\text{windpower})) = 13; \beta = 1.4 \]

\[ x_0 = \ln(\text{windpower}) = 12.1; \beta = 1.3 \]

Use of expression (Eqn.12) gives order statistics, that is, ordered observations only and there seems to be no way of reproducing the stochastic nature of the forecasts. It was therefore found appropriate to use simulation methods, such as MCS for this purpose as has been described earlier. Fig. 9 shows the Weibull plot of wind power forecast over a period of 81 days beyond the period of actual data using Monte Carlo estimation (Eqns.16,17). The shape parameter (\( \beta \)) and the scale parameter (\( x_0 \)) of the model determined using the plot are \( \beta = 1.3 \) and \( x_0 = 12.1 \) which is closer to the parameters of observed data. Also, forecasts can be made for any number of days using the Weibull parameters derived from actual data.
As has already been said, the estimation accuracy can be improved using analytical methods such as MLE. The Weibull parameters computed using the graphical and ML estimates (Eqns.12, 14, 15) are listed in Table I.

<table>
<thead>
<tr>
<th>Wind Power</th>
<th>Graphical</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>ln ($x_0$)</td>
<td>1.4</td>
</tr>
<tr>
<td>Estimated/Reconstructed</td>
<td>12.9</td>
<td>1.4</td>
</tr>
<tr>
<td>Forecast</td>
<td>13.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

It may be observed from the graphical method that the value of $x_0$ differs from the observed data due to the curve fitting for the given data. This can be seen as an approximate method. Whereas, the MLE method gives the Weibull parameters, that are very close to the parameters of the observed data.

### 3.3 Estimation of uncertainties in the forecast

Any time series forecast shall, always be validated against observed value, if it exists. The common measures of errors in prediction or the uncertainties are MAE, MAPE and RMSE \[^{18, 19, 20}\]. While these methods have their limitations, they are simple tools for evaluating the accuracy of forecast. The efficiency of forecasting process can be expressed in terms of the closeness of the estimated power to the actual or observed power. The degree of deviation by which the predicted value differs from the actual value can be computed using the following expressions:

\[
\text{MAE} = \frac{\sum_{t=1}^{N} |E_t|}{N} 
\]

\[
\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{N} E_t^2}{N}}
\]

\[
\text{MAPE} = \frac{\sum_{t=1}^{N} |E_t|}{\sum_{t=1}^{N} Y_t} \times 100
\]

where, $E_t = Y_t - F_t$; where, $E_t$ is the forecast error at period $t$, $Y_t$ is the actual data at period $t$, $F_t$ is the forecast for period $t$ and $N$ is the sample size. The values of Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) have been computed for the ordered wind power data and are tabulated in Table 2.
4. Discussion

A perusal of the literature on this topic suggests that, more often, extreme distributions are invoked in the analysis of wind speed data. However, the degree of variability in the acquired data does not seem to warrant consideration of statistics of extremes. Also, an examination of data set at any location indicates that, by and large, the test data fits a near Gaussian distribution. However, in view of inconsistencies in the tail distributions, it was thought appropriate to use Rayleigh distribution. In this paper, other distribution functions are also considered for the sake of comparison and to show that a Gaussian or a Weibull-type distribution is untenable in so far as the analysis of raw-data on wind speeds is concerned.

Now, the power generated in a wind turbine has been shown to be a function of the cube of velocity. It therefore stands to be assessed based on random nature of wind speeds. This is not a straightforward exercise.

Distribution of a function of a random variable is in general, very different from the original distribution. For instance,

\[ P = \alpha v^3 \text{ or } P = b v^3 \]

It has been shown that, in such an event, the transformed variable P is seen to be Weibull distributed with a constant shape parameter \( \beta \), the characteristic parameter in the two-parameter Weibull is a function of \( b \). This is considered as an important result.

The Borel mapping method seems quite effective in transforming \( v \) on to \( P \). The difficulties faced in obtaining the c.d.f from p.d.f by a process of integration often preclude the analysis of data using c.d.f. In such cases, an approximate estimate of the parameters can be made using p.d.f. However, MLE offers considerable advantages over the graphical methods and have been found to be more accurate.

The Weibull parameters derived from the actual data are used to forecast the wind power for 81 days, beyond the observed data spectrum using, Monte Carlo techniques. In discussing the results of forecasting using Weibull parameters, two important points are to be noted;

- Efficacy of the model used in forecasting
- The acceptability of predicted values

The wind power forecast is done by considering the quarterly data from April-June 2011. On the question of the sufficiency of the model over the prediction interval, the

<table>
<thead>
<tr>
<th>Measures of Error</th>
<th>Observed Vs. Forecast</th>
<th>Estimated (Re-constructed) Vs. Forecast</th>
<th>Observed Vs. Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>21.4 MW</td>
<td>24.3 MW</td>
<td>25.2 MW</td>
</tr>
<tr>
<td>RMSE</td>
<td>27.1 MW</td>
<td>31.2 MW</td>
<td>32.1 MW</td>
</tr>
<tr>
<td>% MAPE</td>
<td>14.16 MW</td>
<td>18.47 MW</td>
<td>18.75 MW</td>
</tr>
</tbody>
</table>
Monte Carlo algorithm applied to estimate the Weibull variates is seen to offer a simple and effective method. It is to be noted, however, that the predicted / estimated time series cannot be compared on a point – to - point basis. The reason for this is that the very process of generating the variates is by using several sets of Uniform random numbers, each set being different from the other. However, the predicted variates preserve the global properties of the Weibull distribution, at least asymptotically (large sample approximation).

In any statistical exercise, the validation of the simulation and other estimated results are mandatory. The estimation of possible uncertainties is made based on several measures of aggregate errors. Table 2 includes these quantities. The authors opine that the estimated / predicted wind power generation, are within acceptable limits, given that the power is a randomly changing quantity. It has been found that the results obtained by Weibull estimation (Eqn.12) are applicable for short term forecast, whereas, the Monte Carlo simulation method presented here, can be applied for forecasting over a considerably longer duration.

5. Conclusions

The terms of reference of the current research paper was to try and analyze wind speed data in a specified location. In so doing, different distribution functions were examined. Based on the analysis presented in this paper, the following broad conclusions can be drawn:

- Among the probability function considered here, the Rayleigh distribution is seen to be of relevance. As such, this distribution is given the preference.
- In this work, the distribution of the generated power is deduced by appropriate mathematical transforms (Borel transforms). An interesting result ensues; while the wind velocity is Rayleigh distribution, the power is Weibull distribution.
- The forecasting method described under the relevant section, was a sequel to commercial requirements of power stations. Often, such information is sought in power system planning and connected reliability studies.
- The Weibull parameters derived for the different cases (as shown in Table 1) are found to be near accurate as seen from the Weibull plots.
- The result indicates that the Mean Absolute Percentage Error (Table 2), being much smaller than one mean value, appears to justify the forecasting procedure.
- The method of forecasting proposed here indicates that prediction intervals can be as long as is desired. Higher prediction accuracies (lower uncertainties) can be achieved by modeling the time series using more accommodative modeling techniques. For example, smoothing algorithms such as Autoregressive, Moving Average and combinations thereof could be tried to make comparisons that are more effective. This is an objective for future work.
References


[19]. Jaesung Jung and Robert P. Current Status and Future Advances for Wind Speed and Power Forecasting, Broadwatera Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg,VA, USA. (Article).