Analysis of wind speed and Estimation of Weibull Parameters by Three Numerical Methods in Al-Sulaimani Province

Awni A. Abdulahad\textsuperscript{1}, Mohammed A. Saleh\textsuperscript{2}, Basim A. Mohammed\textsuperscript{1}

\textsuperscript{1}Department Atmospheric Science, College of Science, Al-Mustansiriyah University
\textsuperscript{2}Ministry of Electricity, Planning Study Office

ABSTRACT

The Weibull distribution is an important distribution especially for reliability and maintainability analysis. The suitable values for both shape and scale parameters of Weibull distribution are important for selecting locations of installing wind turbine generators. In this study, three methods were used to estimate the Weibull parameters (Shape and Scale), namely, Maximum Likelihood Method (MLM), Standard Deviation Method (SDM) and Least Square Method (LSM). These methods were compared for their performance and analysis of actual wind speed according to the criterion such as Root Mean Square Error (RMSE). In this study, (RMSE) values show that the maximum likelihood method performed better than the standard deviation and least square methods for determining the value of (k) and (c) to fit Weibull distribution curves. Monthly mean wind speed and annual wind speed frequencies were calculated to note speed most available through year and high monthly mean wind speed.

INTRODUCTION

Currently Renewable energy has been in use for several years and this use can reduce the greenhouse effect that is the main cause of global warming. Wind energy is a source of clean energy that can be used to generate electricity without air pollution [1]. The assessment of wind resources at a given site is one of preliminary steps in setting of a wind farm project. The assessment of the wind resources involves analyzing in detail the wind speed, i.e., the shape and scale parameters. To determine the suitability of this site for wind energy generation; the mean wind speed, the shape and scale parameters of the site should be estimated. The estimated shape and scale parameters are used alongside with the various statistical functions to model the wind speed, and the wind distributions which are best to describe the variation of wind at the site are obtained [2].

S.A. Ahmeda and H.O. Mahammed [3] concluded that the Weibull distribution is fitting the measured monthly probability density distributions better than the Rayleigh distribution. Mohammadi and Mostafaeipour [4] used two methods (STDM and PDM) for wind data assessment in Zarrineh, Iran. Seguro and Lambert [5] calculated the value of the Weibull parameters by three methods. They recommended that the MLM is useful for time series wind speed data.

In the present study, three methods are presented for estimating Weibull parameters (Shape and Scale), namely, Maximum Likelihood Method (MLM), Least Square Method (LSM) and Standard Deviation Method (SDM). The aim of this work is to select a method that gives more accurate estimation for the Weibull parameters at this location in order to reduce uncertainties related to the wind energy output calculation from any Wind Energy Conversion Systems (WECS).

Wind Site Description

Hourly wind data between Jan 2012-Dec 2012 were collected in Sulaimani city North Iraq (35° 33' 24" North latitude, 45° 27' 11" East longitude and its elevation is 890m above sea level). On a height of (20m) above ground level.

Mean Wind Speed

The wind speed is one of the most important parameters in the wind profile of any given site. The mean wind speed indicates the suitability of a wind site for small-scale to large scale energy generation. The mean wind speed (m/s) of a given site is defined in Equation (1).
\[
\bar{v} = \frac{1}{n} \sum_{i=1}^{n} v_i
\]

Where \( \bar{v} \) is the mean of the values, \( v_i \) is the wind speed observation at \( i \)th time, and \( n \) is the number of wind speed data points [6].

**Modelling of the Wind Speed**

The wind speed variation at a given site is usually described using the wind distribution. Around the world, to identify the suitable statistical distribution for describing the wind speed variation, the following functions have been used and they include the Weibull [7]. The Weibull and Rayleigh functions are the widely accepted and extensively used statistical models for wind energy application [6].

**Weibull Probability Distribution Function**

In wind data analysis Weibull probability density functions are used to characterize the wind speed distribution. The Weibull distribution is often used in the statistical analysis of data. It is used to represent the wind speed distribution in wind energy analysis. The Weibull distribution function is given by:

\[
f(v) = \frac{k}{c} \left( \frac{v}{c} \right)^{k-1} \exp \left( \frac{-v}{c} \right)
\]

Where:

\[f(v) \geq 0, \quad v \geq 0; \quad k > 0, c > 0\]

Where \( f(v) \) is the frequency or probability of occurrence of wind speed \( v \), \( c \) is the Weibull scale parameter with unit equals to the wind speed unit and \( k \) is the dimensionless Weibull shape parameter. The higher value of \( c \) indicates higher wind speed, while the value of \( k \) shows the wind stability. Any probability equation in this paper including Equation (2) can be applied equally well whether the probability is in the form of relative (fractional or percent) or absolute (number of data points) [8, 9].

### Methods of Estimating the Parameters of Weibull Distribution

Weibull parameters regulate the wind speed for optimum performance of a wind conversion system as well as the speed range over which the device is likely to operate. It is therefore, very essential to accurately estimate the parameters of any candidate site for installation of wind energy conversion systems. Various methods have been proposed to estimate the parameters and the suitability of each method varies with sample data distribution, which nevertheless, varies from location to another. [8].

The proposed methods to determining \( k \) and \( c \) are:

3. Least Square Method (LSM).

**Maximum Likelihood Method (MLM)**

Maximum likelihood technique, with many required features is the most widely used technique among parameter estimation techniques. The (MLM) method has many large sample properties that make it attractive for use; it is asymptotically consistent, which means that as the sample size gets larger, the estimate converges to the true values [10].

Let \( v_1, v_2, v_3 \ldots \ldots \ldots v_n \) be a random sample size \( n \) drawn from a PDF \( f(v, \theta) \) where \( \theta \) is an unknown parameter. The likelihood function of this random sample is the joint density of \( n \) random variables and is a function of the unknown parameter. Thus,

\[
L = \prod_{i=1}^{n} f(v_i, \theta)
\]

Is the Likelihood function. The Maximum Likelihood Estimator (MLE) of \( \theta \), say \( \hat{\theta} \), is the value of \( \theta \) that maximizes \( L \) or, equivalently, the logarithm of \( L \).

Often, but not always, the (MLE) of Equation (4) is a solution of

\[
\frac{d \log L}{d \theta} = 0
\]

Now, we apply the (MLE) to estimate the Weibull parameters, namely the shape and scale parameters. Consider the Weibull probability density function (pdf) given in Equation (2), then likelihood function will be:

\[
L(v_1, v_2, \ldots, v_n, k, c) = \prod_{i=1}^{n} \frac{k}{c} \left( \frac{v_i}{c} \right)^{k-1} \exp \left( - \frac{v_i}{c} \right)
\]

On taking the logarithms of Equation (5), differentiating with respect to \( k \) and \( c \) in turn and equating to zero, we obtain the estimating equations:

\[
\frac{\partial \ln L}{\partial k} = \frac{n}{k} + \sum_{i=1}^{n} \ln v_i - \frac{\sum_{i=1}^{n} v_i^k}{\sum_{i=1}^{n} v_i} = 0
\]

\[
\frac{\partial \ln L}{\partial c} = \frac{-n}{c} + \frac{1}{c^2} \sum_{i=1}^{n} v_i^k = 0
\]

In eliminating \( c \) between Equations (6) and (7) and simplifying, one can get

\[
\frac{\sum_{i=1}^{n} v_i^k \ln v_i}{\sum_{i=1}^{n} v_i^k} - \frac{1}{k} = \frac{\sum_{i=1}^{n} \ln v_i}{\sum_{i=1}^{n} v_i}
\]

After rearranging Equation (8), it is possible to estimate the shape factor as follows,

\[
k = \left( \frac{\sum_{i=1}^{n} v_i^k \ln v_i}{\sum_{i=1}^{n} v_i^k} - \frac{\sum_{i=1}^{n} \ln v_i}{\sum_{i=1}^{n} v_i} \right)^{-1}
\]

Because \( k \) appears on both sides of the equation, the equation must be solved iteratively, and to fine a convergent value for \( k \), several iterations are required. Once \( k \) is determined, \( c \) can be estimated using Equation (10) as follows:

\[
c = \left( \frac{\sum_{i=1}^{n} v_i^k}{n} \right)^{1/k}
\]

Here, \( v_i \) is the wind speed in time step \( i \) and \( n \) the number of nonzero wind speed data points [11].

**Standard Deviation Method (SDM)**

This method is useful where only the mean wind speed and standard deviation are available. In addition, it has relatively simple expressions when compared with other methods. Moreover, it is unlike most of the other methods that may require more detailed wind data (which, in some cases, are not readily available) for the
determination of the Weibull distribution shape and scale parameters. The shape and scale factors are thus computed from the mean and standard deviation of wind data as an acceptable approximation in forms:

\[ k = \left(\frac{\sigma}{v_m}\right)^{-1.086} \]  
\[ c = \frac{v_m}{\Gamma\left(1 + \frac{1}{k}\right)} \]  

Where \( k \) and \( c \) can be estimated from Equations (11) and (12) [12].

Least Squares Method (LSM)

The third estimation technique we shall discuss is known as the Least Squares Method. It is so commonly applied in engineering and mathematics problems that are often not thought of as an estimation problem. We assume that a linear relation between two variables. For the estimation of Weibull parameters, we use the method of least squares and we apply it to the results. The cumulative density function of Weibull distribution with two parameters can be written as:

\[ F(v) = 1 - e^{-\left(\frac{v}{c}\right)^k} \]  

This function can be arranged as:

\[ (1 - F(v))^{-1} = e^{-\left(\frac{v}{c}\right)^k} \]  

If we take the natural logarithm of Equation (14), we get:

\[ -\ln(1 - F(v_i)) = k \ln v_i - k \ln c \]  

And then retake the natural logarithm of Equation (15), we get the following equation:

\[ \ln[-\ln(1 - F(v_i))] = k \ln v_i - k \ln c \]  

This is in the form of an equation of a straight line

\[ y_i = ax_i + b \]  

Where \( x_i \) and \( y_i \) are variables, \( a \) is the slope, and \( b \) is the intercept of the line on \( y \) axis, such that:

\[ y_i = \ln[-\ln(1 - F(v_i))] \]

\[ x_i = \ln v_i \]

\[ b = -k \ln c \]  

The idea is to determine the values of \( a \) and \( b \) in Equation (16) such that a straight line drawn through the \((x_i, y_i)\) points has the best possible fit. Parameters \( k \) and \( c \) are given by [13]:

\[ k = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \]

\[ c = \exp\left(\frac{\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} y_i}{nk}\right) \]

Statistical Error Analysis/Goodness of Fit

To find the best method for the analysis, some statistical parameters were used to analyze the efficiency of the above mentioned methods. The Root Mean Square Error test was used to achieve this goal [14].

Root Mean Square Error

The RMSE has been used to compare the actual deviation between the predicted and the actual (measured) values. The Root Mean Square Error value is defined by Equation (21)

\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)^2} \]  

Where, \( y_i \) is the \( (i^{th}) \) actual wind distribution (measured data), \( x_i \) is the predicted wind distribution from the Weibull distribution with the lowest (RMSE) value is chosen as the accurate function to be used for modeling the wind speed [6].

Comparisons and Accuracy of the Methods

Three methods to estimate the parameters of the Weibull wind speed distribution of wind energy analysis for Sulaimani region are presented. The application of each method is demonstrated using a sample wind speed data set, and a comparison of the accuracy of each method is also performed with the actual time series data for our case study, Sulaimani region. In order to compare the methods, monthly mean wind data of Sulaimani region is obtained from meteorological automatic station which 2012. To determine the accuracy of the three methods given in this paper, the Root Mean Square Error (RMSE) was given in equation (21).

Results and Discussion

Data of wind speed used in the present calculations were obtained during 2012 in Sulaimani city at a height of (20 m) above the ground. Wind speeds that taken every 10 s. Program was used to convert the data to an hourly values. The calculates obtained from this study can be summarized as follows:

1. Calculate mean, maximum and minimum wind speed at (20 m) height above the ground.
2. Calculate Weibull distribution parameters by three analytical methods (MLM, SDM and LSM) and select the best one which gives the minimum mean squared error.
3. Plot probability density distribution by the three analytical methods.

Monthly variation in wind speed

The wind speed is one of the most important parameters in the wind profile of at any given site. The mean wind speed refers to the suitability of a wind site from small scale to large scale energy generation. Figure (1) shows the monthly mean, minimum and maximum wind speed at 20 m height collected in (2012). The monthly mean wind speed is calculated using Equation (1). The highest monthly mean wind speed values are observed in Jun, July, August and September; therefore these months have the highest potential of wind energy generation at this site. The lowest monthly mean
wind speed values can be seen in winter months December and January. The maximum wind speed was observed in month Jun is 15.2 m/s and the minimum monthly wind speed was observed in January is 0.3 m/s. This shows the importance of studying wind speed seasonal variability as one of the important parameters in wind energy planning (which gives a meaning that the wind in our location is seasonable). Therefore, we can conclude that summer season is predominating and have a higher wind speed than the rest of seasons.

Methods of estimating Weibull parameters

The variation of wind speeds often described using the weibull two parameters \((c, k)\). The values of Weibull parameters introduce more explanation about the behavior of wind distribution in a given location. The Weibull scale parameter is parameter related with mean wind speed. The Weibull shape parameter is a dimensionless parameter. It refers to the shape of the distribution of the wind speeds curve. Some statistical methods which are widely accepted for estimating Weibull parameters \((k)\) and \((c)\) are:

1- The Maximum Likelihood Method which is used to calculate Weibull parameters by using Equation (9) and Equation (10), the results of using this method is shown in table (1).

<table>
<thead>
<tr>
<th>Months</th>
<th>(c) m/s</th>
<th>(k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>2.116</td>
<td>1.372</td>
</tr>
<tr>
<td>Feb</td>
<td>3.086</td>
<td>1.447</td>
</tr>
<tr>
<td>Mar</td>
<td>4.099</td>
<td>1.376</td>
</tr>
<tr>
<td>Apr</td>
<td>2.597</td>
<td>1.885</td>
</tr>
<tr>
<td>May</td>
<td>2.487</td>
<td>1.795</td>
</tr>
<tr>
<td>Jun</td>
<td>5.597</td>
<td>1.985</td>
</tr>
<tr>
<td>Jul</td>
<td>2.123</td>
<td>1.778</td>
</tr>
<tr>
<td>Aug</td>
<td>4.490</td>
<td>1.936</td>
</tr>
<tr>
<td>Sep</td>
<td>4.365</td>
<td>1.644</td>
</tr>
<tr>
<td>Oct</td>
<td>3.458</td>
<td>1.720</td>
</tr>
<tr>
<td>Nov</td>
<td>2.952</td>
<td>1.533</td>
</tr>
<tr>
<td>Dec</td>
<td>2.344</td>
<td>1.092</td>
</tr>
<tr>
<td>Mean</td>
<td>3.783</td>
<td>1.826</td>
</tr>
</tbody>
</table>

From this table, it is shown that the scale parameter values \((c\) m/s) varies between 2.132 - 5.611 m/sec with yearly mean of 3.813 m/sec, and the shape parameter values \((k\) varies between 1.286 - 1.941 with yearly mean of 1.656 for height 20m. The (RMSE) value for the maximum likelihood method is (0.0185).

The Standard Deviation Method: this method used to calculate Weibull parameters by using Equation (11) and Equation (12) the results are shown in table (2).

<table>
<thead>
<tr>
<th>Months</th>
<th>(c) m/s</th>
<th>(k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>2.103</td>
<td>1.576</td>
</tr>
<tr>
<td>Feb</td>
<td>3.766</td>
<td>1.554</td>
</tr>
<tr>
<td>Mar</td>
<td>3.840</td>
<td>1.683</td>
</tr>
<tr>
<td>Apr</td>
<td>3.272</td>
<td>1.932</td>
</tr>
<tr>
<td>May</td>
<td>3.587</td>
<td>2.010</td>
</tr>
<tr>
<td>Jun</td>
<td>3.981</td>
<td>2.111</td>
</tr>
<tr>
<td>Jul</td>
<td>4.517</td>
<td>1.981</td>
</tr>
<tr>
<td>Aug</td>
<td>4.385</td>
<td>2.009</td>
</tr>
<tr>
<td>Sep</td>
<td>4.277</td>
<td>1.938</td>
</tr>
<tr>
<td>Oct</td>
<td>3.385</td>
<td>2.100</td>
</tr>
<tr>
<td>Nov</td>
<td>2.848</td>
<td>1.685</td>
</tr>
<tr>
<td>Dec</td>
<td>2.471</td>
<td>1.238</td>
</tr>
<tr>
<td>Mean</td>
<td>3.093</td>
<td>1.881</td>
</tr>
</tbody>
</table>

From this table, it is noted that the \((c\) m/s) values varies between 2.10 - 5.497 m/s with yearly mean of 3.69 m/sec and the \((k\) varies between 1.25 - 2.2 with yearly mean of 1.86 for height (20m). The (RMSE) value for the least square method is (0.020).

By comparing the root mean square error values for the previous results it is obvious that the MLM has generally lower RMSE value than both the SDM and the LSM. Therefore, it can be conclude that, the maximum likelihood method is the best and more accurate to estimate the Weibull parameters at study location.

Probability density distribution

Weibull distribution can be used to describe the asymmetry properties of the wind frequency distribution. The Weibull distribution is an important distribution especially to the site reliability analysis, by using Equation (2). The suitable values for both shape and scale parameters of Weibull distribution are important
for selecting locations of installing wind turbine generators.

Figure (2) show the results for the calculated probability density function (pdf), for the three methods for estimating Weibull parameters. The results of these methods are depicted as a (red line). The blue bars show the relative frequency with each wind speed bin occurs at (1m/sec) (which is shown along x axis) for Sulaimani region. These bars are based on the results from the analysis wind speed.

From the figure (4.6) it could be concluded that:

1. The maximum likelihood and standard deviation methods are the most fitted methods to estimate the Weibull parameters from least square method.
2. The distribution shows that the most frequent winds are between 1–3 m/s.
3. There is a clear difference in the peaks of probability density function between maximum likelihood method and least square method, also standard deviation and least square method. This is the result of the difference in shape parameter values.
4. Higher value of the scale parameter implies the distribution is extended on a wider range and the probabilistic mean wind speed has a higher value.

Conclusions
In this work, statistical diagnosis of the best Weibull distribution methods for wind data analysis is presented. By using the available wind data, the values of shape factor \(k\) and scale factor \(c\) were determined using three methods and were then investigated as to how efficiently the methods can estimate the Weibull factors with minimum error. To satisfy the main objectives of this work, statistical tool (RMSE) was used to find the best method of Weibull distribution. The results show that:

1. The annual mean wind speeds in Sulaimani Province is 3.4 m/s for 2012, and the highest monthly mean wind speed values are observed in June, July, August and September. Therefore, conclude that summer season is predominating and have a higher wind speed than the rest of seasons.
2. The maximum likelihood method is the best and more accurate to estimate the Weibull parameters at study location.
3. The Weibull distribution shows that the most frequent winds are between 1–3 m/s.

References


