Comparison of Different Methods in Estimating Weibull Distribution Parameters for Wind Power Application

Yusuf Alper KAPLAN¹,*, Canan ALADAĞ²

Assistant Professor, Department of Energy Engineering, Osmaniye Korkut Ata University, Osmaniye Turkey¹
Research Assistant, Department of Energy Engineering, Osmaniye Korkut Ata University, Osmaniye, Turkey²

ABSTRACT: In recent years, there has been an increasing demand for energy due to the population growth and decline in fossil energy resources as well as the developing technology. Thus, there has been a recent tendency towards exploring new energy resources. In this study, the wind energy potential in Hatay was analyzed statistically based on the data regarding the wind speed measured hourly between 2009 and 2013. The wind speed distribution in the turbine location plays an important role besides wind turbine parameters in order to produce wind electricity. Weibull distribution was used in the research to reveal the local wind energy potential. Weibull parameters the c and k values were estimated analytically, and the mean wind speed and energy potential were determined based on these parameters. This study introduced the evaluation of different numerical methods which are used in calculating the Weibull parameters, namely Graphical Method (GM), Standard Deviation Method (SDM), Maximum Likelihood Method (MLM), Energy Pattern Method (EPM), Energy Trend Method (ETM) and Moment Method (MM).

KEYWORDS: Weibull Distribution, Numerical Methods, Wind Potential, Comparative Analysis.

I. INTRODUCTION

Wind energy is regarded as a clean energy resource for the environment with its renewable characteristic. It does not require any sort of transportation or high technology. Renewable wind energy has some other advantages as well. It is abundant and free in the atmosphere, and is a clean energy resource, which is easy to obtain and creates no environmental pollution [1]. Energy is also primary element of economic and social development in the world. Developed and developing countries tend to include the use of renewable energy resources into their future plans. These countries develop new technology in this field and increase their investment in order to utilize their potential more effectively [2]. The world will use renewable energy instead of using fossil fuels in order to meet the demands of the world’s energy [3,4]. Different numerical methods have been suggested in the literature in order to compute the Weibull Distribution parameters. Seguro and Lambert [5] estimated of Weibull distribution for sample wind data set using graphical, modified maximum likelihood and maximum likelihood methods. They declared that graphical method is the less accurate one, while maximum likelihood method is suggested for wind energy analysis. Safari and Gasore [6] studied wind energy potential for six different region in Rwanda. They calculated Weibull and Rayleigh distribution for all region by applying maximum likelihood method. They found maximum power density of 109,89 (W/m²) at Gisenyi. Bilgili and Sahin [7] investigated wind energy density in the southern and southwestern region of Turkey by using the Weibull and Rayleigh probability density functions. Rocha et al. [8] deal with the evaluation and comparison of different numerical methods for the assessment of efficiency in determining the parameters for the Weibull distribution function, using wind speed data collected in Camocim and Paracuru cities. Khahro et al. [9] used graphical, standard deviation, maximum likelihood, modified maximum likelihood and power density methods to specify the wind power potential of Babaurband in Pakistan. Weibull parameters were estimated at different heights for one year period. They also made energy generation prediction for a commercial wind turbine. Chang [10] reviewed the six kinds of numerical methods commonly used for estimating Weibull parameters. The Weibull distribution function has been
generally used to assessment of wind energy potential. The Weibull distribution has two parameters which are dimensionless shape ‘k’ and scale ‘c’ parameters.

II. WIND DATA ANALYSES OF HATAY REGION

A precise determination of probability distribution for wind speed data is the most important step in statistically evaluating wind speed and wind energy potential of a region. In this study, wind characteristic and wind energy potential that is located in the Mediterranean coastal city Hatay of Turkey where is shown in Figure 1, are analyzed using the annual average wind speed data of Turkish State Meteorological Service measured at 10 m height.

This study aims to determine the wind energy potential in Hatay. We have collected the five years data of the wind speed measured at 10-meter height from the General Directorate of State Meteorology. Based on the hourly-measured wind speed data, we have statistically analyzed the region’s wind energy potential. We have used actual data and different type of numerical method to determine the Weibull Distribution parameters. Figure 2 shows the density of wind directions according to wind frequency. It is clearly seen that the dominant wind direction is SSW direction.
II.1 Wind Power Calculation

Firstly it is necessary to determine wind power potential in order to benefit from wind energy. Recently, studies are ongoing worldwide on Weibull Distribution Function. The obtained results show that this method is most reliable method for wind energy estimation. Wind power is expressed as [5].

\[ P_A = \frac{1}{2} \rho A v^3 \]  

(1)

As it has been seen, wind speed is the most important input in order to determine the power potential [11].

II.2 Statistical Analysis

There are many distribution functions to evaluate wind speed distribution and wind energy potential. Two-parameter Weibull and Rayleigh distribution functions are the most popular of all. Hourly wind speed and wind direction details are observed in a place and statistical results are calculated for modelling the frequency and probability of the obtained results [7, 13]. First, as is seen in Table 1, the periodical frequency (the blowing number) and probability of wind speed are determined. Wind speeds are grouped periodically in the second column of Table 1. The third column shows average wind speed for each speed ratio. The blowing number or frequency of each speed ratio is given in the fourth column [12].

\[ p(v_i) = \frac{f_i}{\sum_{i=1}^{N} f_i} \]  

(2)

Here, \( f_i \) is frequency of occurrence of each speed class and \( N \) is number of hours in the period of time considered. \( P(v_i) \) is the cumulated probability density which is illustrated in the sixth column of the Table 1.
Table 1. Periodical Frequency and Probability of Hourly Wind Speeds

<table>
<thead>
<tr>
<th>i</th>
<th>(v_i) (m/s)</th>
<th>(v_i^*) (m/s)</th>
<th>(f_i)</th>
<th>(p(v_i))</th>
<th>(P(v_i))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0-1</td>
<td>0,5</td>
<td>13396</td>
<td>0,31</td>
<td>0,31</td>
</tr>
<tr>
<td>2</td>
<td>1-2</td>
<td>1,5</td>
<td>15018</td>
<td>0,34</td>
<td>0,65</td>
</tr>
<tr>
<td>3</td>
<td>2-3</td>
<td>2,5</td>
<td>8212</td>
<td>0,19</td>
<td>0,84</td>
</tr>
<tr>
<td>4</td>
<td>3-4</td>
<td>3,5</td>
<td>3525</td>
<td>0,08</td>
<td>0,92</td>
</tr>
<tr>
<td>5</td>
<td>4-5</td>
<td>4,5</td>
<td>2396</td>
<td>0,05</td>
<td>0,97</td>
</tr>
<tr>
<td>6</td>
<td>5-6</td>
<td>5,5</td>
<td>1134</td>
<td>0,03</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>6-7</td>
<td>6,5</td>
<td>137</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>7-8</td>
<td>7,5</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The two-parameter Weibull distribution function is described for wind speed as follow Eq. (3) and as a cumulative distribution function given by Eq. (4) [14, 15]

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

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

Actual probability density function and cumulative probability distributions derived from the long-term wind speed data of Hatay during 2009 to 2013 period are depicted in Figure 3.

Fig. 3. The wind speed probability density and cumulative probability distributions for the year 2009 to 2013

III. DETERMINATION OF WEIBULL PARAMETERS

In practice, a variety of methods have been utilized to determine the wind energy of a region. A wind speed distribution measure or a frequency distribution is used to determine the wind energy potential of a specific region, if possible. If not, wind speed distribution can be demonstrated through other analytical distribution functions. One of
these functions is the Weibull distribution function. This distribution is not only quite flexible and simple, but it also complies with real data [16, 17]. In the present study, six of these methods:

- Graphical Method
- Standard Deviation Method
- Energy Trend Method
- Energy Pattern Method
- Maximum Likelihood Method
- Moment Method

are employed in order to characterize the wind speed data recorded for five-year period between 2009 and 2013. The Weibull distribution function, which is a three-parameter function, but for wind speed is known, it can be expressed mathematically in two parameter model.

### III.1 Graphic Method

The graphical method depends on the evaluation of cumulative distribution function of wind speed. Taking double logarithmic transformation of the cumulative distribution function yields Eq. 5 and Eq. 6 as follows,

\[
\left(\frac{v}{c}\right)^k = \ln[1 - P(v)]
\]  
\[ (5) \]

\[
k \ln(v) - k \ln(c) = \ln\left[-\ln[1 - P(v)]\right]
\]  
\[ (6) \]

Here, if we accept \( x = \ln(v) \), \( y = \ln[-\ln[1 - P(v)]] \), A = k and B = -k ln(c), we can obtain a linear equation as \( y = Ax + B \). Furthermore, we will obtain \( c = \exp(-B/A) \) from B = - ln (c) [7, 19]. Table 2 shows the obtained xi and yi parameters for graphical method.

<table>
<thead>
<tr>
<th>i</th>
<th>vi*(m/s)</th>
<th>f</th>
<th>( x_i = \ln(v_i) )</th>
<th>( y_i = \ln[-\ln(1 - P(v_i))] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0,5</td>
<td>13396</td>
<td>-0,693</td>
<td>-1,00836</td>
</tr>
<tr>
<td>2</td>
<td>1,5</td>
<td>15018</td>
<td>0,4055</td>
<td>0,044175</td>
</tr>
<tr>
<td>3</td>
<td>2,5</td>
<td>8212</td>
<td>0,9163</td>
<td>0,591324</td>
</tr>
<tr>
<td>4</td>
<td>3,5</td>
<td>3525</td>
<td>1,2528</td>
<td>0,907925</td>
</tr>
<tr>
<td>5</td>
<td>4,5</td>
<td>2396</td>
<td>1,5041</td>
<td>1,262902</td>
</tr>
<tr>
<td>6</td>
<td>5,5</td>
<td>1134</td>
<td>1,7047</td>
<td>1,744859</td>
</tr>
<tr>
<td>7</td>
<td>6,5</td>
<td>137</td>
<td>1,8718</td>
<td>2,185622</td>
</tr>
</tbody>
</table>
After determining the numerical values of A and B in \( y = Ax + B \) which is shown in Figure 4, \( k \) and \( c \) parameters can be calculated as follow:

\[
k = A , \quad c = \exp \left( -\frac{B}{A} \right)
\]  

(7)

III.II Standard Deviation Method

Standard Deviation Method is a method that, data distribution allows to solve the relationship between average and standard deviation values between the shape parameters using numerical methods for determining the scale of 1 to 10 and the shape parameter. Shape and scale parameters can be expressed as in Eq. 8 and Eq. 9 [10, 14].

\[
k = \left( \frac{n}{\bar{V}_{in}} \right)^{1.6667}
\]

(8)

Here \( \Gamma \) is Gamma function;

\[
C = \frac{\gamma_m}{\phi \left( 1 + \frac{1}{\lambda} \right)}
\]

(9)

Here, \( \phi \), shows Standard deviation \( \gamma_m \), average speed;

\[
\bar{V}_m = \frac{1}{n} \sum_{i=1}^{n} v_i
\]

(10)

\[
\sigma' = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (v_i - \bar{V}_m)^2}
\]

(11)

III.III Energy Pattern Method

The energy pattern method is related to the mean of wind speed. This method uses the mean of wind speed cubes and the cube of mean wind speed to calculate the energy pattern factor Epf and then the Weibull parameters [9].

\[
Epf = \frac{1}{\left( \frac{n}{\bar{V}^3} \right)}
\]

(12)
III. IV Energy Trend Method

Energy tendency factor is defined as wind energy obtained by the ratio of total wind energy to the average wind speed. After calculating energy trend factor (EPF) the shape parameter \(k\) can be obtained by following equation \([10, 22]\).

\[
R = 1 + \frac{E_{EPF}}{EPF} 
\]

\[
c = \sqrt{\frac{\text{mean}}{\text{standard deviation}}} 
\]

III. V Maximum Likelihood Method

The maximum likelihood estimation method is a mathematical expression recognized as a likelihood function of the wind speed data in time series format. In this method, many numerical iteration can be required to determine the \(k\) and \(c\) parameters of the Weibull function. In this method Weibull parameters \(k\) and \(c\) values are calculated by the following equation\([8, 9, 20]\).

\[
E_{EPF} = \frac{1}{2} \sum\frac{V_i^2}{V_m^2} 
\]

\[
R = \frac{1}{2} \left( \frac{E_{EPF}}{EPF} \right)^{0.622} 
\]

The scale parameter calculated as follows;

\[
c = \left( \frac{1}{n} \sum V_i^c \right)^{1/k} 
\]

III. VI Moment Method

Moment method is one of the oldest method is used to determine Weibull distribution parameters. The first two moments of the Weibull density function are utilized to calculate the parameters ‘\(k\)’ and ‘\(c\)’. The calculations are based on standard deviation, average wind velocity and gamma function for parameter \([10, 13]\).

\[
k = \left( \frac{1.0984}{V_m} \right)^{1.0989} 
\]
Here \( \Gamma \) is Gamma function;
\[
c = \frac{v_m}{\Gamma\left[\frac{1}{\beta}\right]}
\]  
(21)

The Weibull shape factor \( k \) is very important factor to know about the characteristics of the wind wave of a particular wind site. The Weibull scale factor \( c \) indicates the potentiality of the wind power of that site. Bigger the value of \( c \) parameter means the more potential of wind sites. The performance of five numerical methods in estimating Weibull Distribution parameters between 2009 and 2013 has been graphically compared.

**IV. OBTAINED RESULTS AND DISCUSSION**

**IV.I Statistical Error Analysis**

To determine the efficiency of the used methods and to find the best method of aforementioned methods, the following three statistical indicators were used: The root mean square error (RMSE), Analysis of variance \( (R^2) \) and Mean percentage error (MPE) [20].

\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - x_i)^2}
\]  
(22)

\[
R^2 = \frac{\sum_{i=1}^{N} (y_i - x_i)^2 - \sum_{i=1}^{N} (y_i - \bar{y})^2}{\sum_{i=1}^{N} (y_i - x_i)^2}
\]  
(23)

\[
\text{MPE} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{|y_i - x_i|}{y_i} \right) \times 100\%
\]  
(24)

By applying the above equations the performance of the Weibull distribution function can be evaluated where the model with the lowest RMSE, MPE and \( R^2 \) which gives the closest outcome to a value of “1” is a best method which gives the closest result to the actual data [21].

<table>
<thead>
<tr>
<th>Method</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>0.57378</td>
<td>0.61033</td>
<td>0.50643</td>
<td>0.50222</td>
<td>0.49510</td>
</tr>
<tr>
<td>SDM</td>
<td>0.00272</td>
<td>0.00253</td>
<td>0.00275</td>
<td>0.00306</td>
<td>0.00283</td>
</tr>
<tr>
<td>EPM</td>
<td>0.00216</td>
<td>0.00248</td>
<td>0.09624</td>
<td>0.03410</td>
<td>0.00146</td>
</tr>
<tr>
<td>ETM</td>
<td>0.08680</td>
<td>0.20615</td>
<td>0.04245</td>
<td>0.02979</td>
<td>0.05564</td>
</tr>
<tr>
<td>MLM</td>
<td>0.05749</td>
<td>0.01817</td>
<td>0.00806</td>
<td>0.00658</td>
<td>0.00689</td>
</tr>
<tr>
<td>MM</td>
<td>0.00060</td>
<td>0.00003</td>
<td>0.03811</td>
<td>0.00003</td>
<td>0.00055</td>
</tr>
</tbody>
</table>

The most accurate annually mean RMSE analysis are observed in the methods of MM because of close value of “1” for all years and the worst values GM are occurred in all years as shown in Table. 3.

It is clearly seem that RMSE analysis result; MM presents better performance than other methods in 2009, 2010, 2012 and 2013, SDM presents better performance than other methods in 2011.
Table 4. The Annually $R^2$ analysis

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>0.691780</td>
<td>0.713030</td>
<td>0.103631</td>
<td>0.319250</td>
<td>0.038266</td>
</tr>
<tr>
<td>SDM</td>
<td>0.999993</td>
<td>0.999995</td>
<td>0.999973</td>
<td>0.999950</td>
<td>0.999960</td>
</tr>
<tr>
<td>EPM</td>
<td>0.999960</td>
<td>0.999950</td>
<td>0.967623</td>
<td>0.993916</td>
<td>0.999990</td>
</tr>
<tr>
<td>ETM</td>
<td>0.992946</td>
<td>0.967260</td>
<td>0.993700</td>
<td>0.995350</td>
<td>0.987853</td>
</tr>
<tr>
<td>MLM</td>
<td>0.996905</td>
<td>0.999745</td>
<td>0.999773</td>
<td>0.999770</td>
<td>0.99981</td>
</tr>
<tr>
<td>MM</td>
<td>1.00000</td>
<td>1.00000</td>
<td>0.99490</td>
<td>1.00000</td>
<td>0.99999</td>
</tr>
</tbody>
</table>

The obtained $R^2$ results are shown in Table 4, the GM’s performance in estimating Weibull parameters is worst one and the other methods give very close results for all years. Based upon the results of $R^2$ analysis; MM presents better performance than other methods in 2009, 2010, 2012 and 2013, SDM presents better performance than other methods in 2011.

Table 5. The Annually MPE analysis

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>-27,268</td>
<td>-27,204</td>
<td>-38,649</td>
<td>-37,707</td>
<td>-35,640</td>
</tr>
<tr>
<td>SDM</td>
<td>-0,040</td>
<td>0,0353</td>
<td>-0,127</td>
<td>-0,180</td>
<td>-0,155</td>
</tr>
<tr>
<td>EPM</td>
<td>-0,038</td>
<td>-0,008</td>
<td>2,517</td>
<td>0,662</td>
<td>-0,003</td>
</tr>
<tr>
<td>ETM</td>
<td>0,712</td>
<td>-1,209</td>
<td>0,811</td>
<td>1,291</td>
<td>1,413</td>
</tr>
<tr>
<td>MLM</td>
<td>-0,954</td>
<td>-0,052</td>
<td>0,301</td>
<td>-0,101</td>
<td>-0,203</td>
</tr>
<tr>
<td>MM</td>
<td>-0,0133</td>
<td>0,0006</td>
<td>0,7376</td>
<td>-0,0015</td>
<td>0,1892</td>
</tr>
</tbody>
</table>

The obtained MPE results are shown in Table 5, according to obtained results, the GM’s performance in estimating Weibull parameters is worst one and the other methods give very close results for all years. It is estimated that MPE analysis; MM presents better performance than other methods in 2009, 2010 and 2012, SDM presents better performance than other methods in 2011, EPM presents better performance than other methods in 2013.

IV.II Power Density Assessments

As known, the different between the wind energy calculated by actual time series data and by the Weibull Distribution Function data can be used as an indicator to evaluate the accuracy of used method. The average wind speed ‘$V_m$’ and wind power density ‘$P_m$’ can be computed by the following equations for real time series datas [22]:

$$V_m = \frac{1}{n} \sum_{i=1}^{n} v_i$$  \hspace{1cm} (25)

$$P_m = \frac{1}{2} \rho V_m^3$$  \hspace{1cm} (26)

$V_m = 1.8515$ and $P_m = 11,04864 \text{ w/m}^2$. The results are obtained from 43824 datas which are measured hourly wind speed for the year 2009 to 2013.
The mean wind velocities and wind power densities have also been calculated with the help of k and c parameters. The average wind speed ‘\( V_w \)’ and wind power density ‘\( P_w \)’ of Weibull Distribution Function can be estimated by the following equations:

\[
V_w = c \Gamma \left( 1 + \frac{1}{k} \right)
\]

\[
R_w = \frac{1}{2} \rho c^2 \Gamma \left( 1 + \frac{2}{k} \right)
\]

Table 6. The mean wind speed and power densities of all methods

<table>
<thead>
<tr>
<th></th>
<th>K</th>
<th>c</th>
<th>( V_w )</th>
<th>( P_w )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>1.179</td>
<td>1.351</td>
<td>1.277</td>
<td>5,274</td>
</tr>
<tr>
<td>SDM</td>
<td>1.466</td>
<td>2.045</td>
<td>1.851</td>
<td>10,938</td>
</tr>
<tr>
<td>EPM</td>
<td>1.461</td>
<td>2.046</td>
<td>1.853</td>
<td>11,028</td>
</tr>
<tr>
<td>ETM</td>
<td>1.554</td>
<td>2.087</td>
<td>1.876</td>
<td>10,454</td>
</tr>
<tr>
<td>MLM</td>
<td>1.457</td>
<td>2.047</td>
<td>1.855</td>
<td>11,103</td>
</tr>
<tr>
<td>MM</td>
<td>1.452</td>
<td>2.042</td>
<td>1.852</td>
<td>11,099</td>
</tr>
</tbody>
</table>

According to the calculated power densities based on five-year data, it was found that MM, SDM, EPM, ETM and MLM provided very desirable performance while the GM gives the weak performance for all years.

V. CONCLUSIONS

For the wind tribunes to be installed in the province of Hatay, the compatibility of the wind speed with Weibull distribution function has been studied by using the measured wind speed rates that had been measured for five years between 2009 and 2013, and numerical results were presented. Results obtained through graphical method were shown for each month and year separately. In this study, wind speed data were measured at 10 m of ground clearance by the Automatic Meteorology Observation Station, which belongs to the General directorate of the Meteorology Service and is located in the centre of Hatay. From the above investigations, the all methods except GM give close results to each other. It cannot be expected that the same method has the best result for all period and for all used statistical performance tests.

Considering the fact that the data used herein are about the measurements made at 10 m of ground clearance, it is understood that there is much more quality wind potential with sufficient power density at approximately 100 m of ground clearance.

REFERENCES


