Performance Comparison of New Generation Parameter Estimation Methods for Weibull Distribution to Compute Wind Energy Density

Ahmet Emre Onay1, Emrah Dokur2,*, Mehmet Kurhan2
1Department of Energy Systems Engineering, Bilecik S. E. University, 11210 Bilecik, Turkey
2Department of Electrical Electronics Engineering, Bilecik S. E. University, 11210 Bilecik, Turkey
emrah.dokur@bilecik.edu.tr

Abstract—To install a wind energy conversion system to a region, the wind speed characteristics of that region must be identified. The two-parameter Weibull distribution is highly efficient in modeling wind speed characteristics. In this study, the wind speed data of 32 cities in three different regions of Turkey have been comparatively analysed to estimate Weibull distribution function parameters by the use of three well-known methods (Graphical Method (GM), Maximum Likelihood Method (MLM), Justus Moment Method (JMM)) and three new parameter estimation methods (Energy Pattern Factor Method (EPFM), Wind Energy Intensification Method (WEIM), Power Density Method (PD)) which have been proposed in recent years. Three years of hourly wind speed data of the specified regions have been used. The performance metrics of these analyses have been compared using Wind Energy Error (WEE), Root Mean Square Error (RMSE), and Coefficient of Determination ($R^2$). The results have shown that while the PD method has high model performance, the JMM is closely competitive with the MLM. Besides, the wind energy densities that were estimated by using actual data have been compared with the resulting Weibull distribution. It has been clear that the method that has the closest estimation to the actual values is the PD method.

Index Terms—Wind energy; Estimation methods; Weibull distribution; Wind energy intensification.

I. INTRODUCTION

The use of wind as a renewable and green energy is increasing every day. The demand for energy of the rapidly growing population is greater than ever. It is clear that not one energy source is going to be enough. Besides, the nonrenewable resources like fossil fuels are being depleted. For these reasons, coal, oil, and gas reserves are assumed to run out in 95, 23, and 25 years, respectively [1]. The climate change and the increased awareness of the people towards energy consumption make the work on renewable sources more urgent.

In this perspective, the use of the renewable energy resources like wind, solar, and hydroelectric power has increased significantly in recent decades. The use of wind power is especially important due to its advantages over the other resources, such as: i) consisting of fewer parts than the other sources, ii) being able to generate electricity in day and night times, iii) creating more environmentally friendly energy.

On the other hand, the power of the wind must be estimated to predict the amount of use we can get out of a wind plant. For this purpose, the main motivation of this study is the comparison of six different estimation methods for Weibull distribution in a case study in Turkey.

The Weibull distribution is first suggested by W. Weibull in 1951 [2] as a tool to use on statistical problems. Today, Weibull distribution is a widely used method not only in the characterization of the wind speed data, but also in predicting water level and rainfall, ceramic strength, solar irradiance assessment, and microbial survival data [3]–[5].

In this study, the two-parameter Weibull distribution is used to model the wind data. There are many studies on modeling wind speed characteristics that have been conducted using the two-parameter Weibull distribution.

A. Related Works

Some of the mentioned studies were conducted in Turkey, where this study took place. These studies use the Weibull distribution and other techniques like ANN (Artificial Neural Networks) [6], [7]. These studies show that the use of the tools, such as the Weibull distribution, is a trusted way to characterize the wind speed of a specific region. As a country that is in a good location in terms of wind energy, Turkey has an estimated 48 GW of wind energy potential [8].

When the literature is inspected on the use of Weibull parameter estimation methods, it can be seen that there is no one method that is agreed upon which works better than the others. In this perspective, there are many studies conducted on Weibull distribution and parameter estimation methods for lots of regions.

In his study, Ahmed [9] analysed the wind data in Halabja, Iraq region, to estimate the Weibull parameters. The four methods (Maximum Likelihood Method (MLM), Rank Regression Method (RRM), Mean-Standard Deviation (MSD), and Power Density Method (PD)) were analysed comparatively. The RMSE and $R^2$ were used to compare the methods.
Besides, in the study that was published by Sumair, Aized, Gardezi, and Aslam [10], they have compared some new and old parameter estimation methods in five different regions of Pakistan. These methods were Combined Linearized Moment Method (CLMM) and Method of Moment (MOM). The results are compared using $R^2$, Mean Square Error (MSE), RMSE, WEE, Mean Absolute Percentage Error (MAPE), and Chi-test.

**B. Structure of the Study**

This study aims to compare the efficiency of 6 different Weibull parameter estimation methods by utilizing three years of hourly wind data of 32 cities in three different regions of Turkey. The methods that have been used in this paper are three well-known (Graphical Method (GM), MLM, and Justus Moment Method (JMM)) and three newly proposed methods (Energy Pattern Factor Method (EPFM), Wind Energy Intensification Method (WEIM), and PD).

Section II focuses on the Weibull distribution and the wind data. Section III gives more information about the wind data and the method used to process the data. Section IV explains the parameter estimation methods for the Weibull distribution which were used in this study. Section V focuses on the calculation of the wind energy of the wind data. This calculation has been used in comparing the methods’ efficiencies at calculating the actual wind energy. In Section VI, the error performance comparison of these methods has been conducted using WEE, RMSE, and $R^2$. Lastly, in Section VII, all methods have been compared according to their efficiencies and error performances.

**II. WEIBULL DISTRIBUTION AND PARAMETER ESTIMATION**

**A. Weibull Probability Density Function**

The probability density function (PDF) is as follows

$$f(v) = \left(\frac{k}{c}\right)v^{k-1}e^{-\left(\frac{v}{c}\right)^k}.$$  \hspace{1cm} (1)

Equation (1) shows $f(v)$ as the Weibull function that is calculated for any wind speed $(v)$ (m/s). This value is the probability to observe the wind speed. Two parameters are estimated using Weibull distribution. These are the shape ($k$) and the scale ($c$) parameters. The calculation of these parameters is made by using different methods [11], some of which have been used in this study.

**B. Weibull Cumulative Distribution Function**

Equation (2) shows the $F(v)$ value as the cumulative distribution function (CDF)

$$F(v) = 1 - e^{-\left(\frac{v}{c}\right)^k}.$$  \hspace{1cm} (2)

**III. WIND DATA**

The wind data for this study have been acquired from SolarRadiationData (SoDa)’s website [12]. As previously mentioned, the data consists of three years of hourly wind speeds. However, the source has not been able to supply the data in a desired height. To resolve this issue, one of the tools to convert wind speeds at different heights has been used.

Figure 1 shows the studied areas of Turkey together with the cities in the areas. In Fig. 2, the calculated wind energy densities of the cities in these studied areas are given separately.

![Fig. 1. The view of the studied regions and cities on the map.](image)

![Fig. 2. The wind energy densities of the cities in (a) the Aegean, (b) Central Anatolia, and (c) Marmara regions, respectively.](image)

**Power Law**

Since the wind data were of 10 meters height, one of the methods in literature to convert wind speeds in different heights has been used to convert the data from 10 meters to 50 meters height. The applied method is called “Power Law” (PL) [13], and by using this method in (3), the data at the height of 50 meters have been acquired for every city.
In (3), \( \alpha \) is called “Hellman (or friction) exponent”, it is also known as “WSC”, which is related to atmospheric stability, wind speed, surface specifications, or height difference. WSC’s value differs from 0.40 in city centers to 0.10 in terrain with not many obstacles and lakes and oceans [14]. The widely used value of WSC in the world is 1/7 (0.143) [15]. In addition, \( v_i \) in (3) is the wind speed value, which was measured in \( h_1 \) height, and \( v_2 \) is the wind speed value at the desired \( h_2 \) height.

IV. METHODS USED TO ESTIMATE WEIBULL PARAMETERS

The GM, MLM, EPFM, WEIM, JMM, and PD methods have been used in this study. The MATLAB software has been used to calculate the parameters of the Weibull for each method.

A. Graphical Method (GM)

This method is created by using the cumulative distribution function. If the logarithm of the value in (2) is taken twice,

\[
\ln[-\ln[1 - F(v)]] = k \ln v - k \ln c, \tag{4}
\]

the value in (4) is acquired. As the slope of a straight line, we achieve the value of \( y = ax + b \) according to this formula. To find out the \( a \) and \( b \) values:

\[
k = a, \tag{5}
\]

\[
c = e^{(\frac{b}{a})}. \tag{6}
\]

are used. The steps to use this data are: i) by using wind speed data, the cumulative frequency distribution or frequency distribution is calculated, ii) \( (\ln v_i) \) and \( -\ln[1 - F(v_i)] \) values are calculated, and iii) by solving the problem of linear least squares and using (5) and (6), the shape and scale parameters are identified [16].

B. Maximum Likelihood Method (MLM)

This method is proposed by Stevens and Smulders [17]. The shape and scale parameters of the Weibull distribution are calculated by using (7) and (8):

\[
k = \left( \frac{\sum_{i=1}^{n} v_i^k \ln(v_i) - \sum_{i=1}^{n} \ln(v_i) \left( \frac{\sum_{i=1}^{n} v_i^k}{n} \right)}{\sum_{i=1}^{n} v_i^k} \right)^{-1}, \tag{7}
\]

\[
c = \left( \frac{\sum_{i=1}^{n} v_i^k}{n} \right)^{\frac{1}{k}}. \tag{8}
\]

where \( v_i \) is the wind speed and \( n \) is the number of non-zero wind speeds. Equation (7) needs to be solved by an iterative approach (\( k = 2 \) could be used as an initial value), then (8) could be solved as the last step of calculating the parameters.

C. Energy Pattern Factor Method (EPFM)

This method requires a value called “energy pattern factor” (\( Epf \)) to be calculated first

\[
Epf = \frac{\bar{v}}{\langle \bar{v} \rangle}. \tag{9}
\]

Then, the scale and shape parameters are calculated using (10) and (11), respectively [18]:

\[
k = 1 + \frac{3.69}{Epf^2}, \tag{10}
\]

\[
c = \frac{\bar{v}}{\Gamma \left( 1 + \frac{1}{k} \right)}. \tag{11}
\]

In addition, Akdag and Dinler [16] has stated (12) in their study about the \( Epf \) value

\[
\frac{\bar{v}}{\langle \bar{v} \rangle} = \frac{\Gamma(1+3/k)}{\Gamma(1+1/k)}. \tag{12}
\]

D. Wind Energy Intensification Method (WEIM)

This method has been proposed by Sumair, Aizedi, ur Rehman, and Rehman [19] and is similar to EPFM.

Equation (13) shows the calculation of a coefficient called “wind energy intensifier” (WEI), and after this calculation, the WEI value is used to estimate the scale \( c \) parameter in (14):

\[
A = WEI = \frac{1}{2} \rho \bar{v}^3 T = \left[ \int_0^\infty \rho v f(v) dv \right]^3, \tag{13}
\]

\[
c \propto \bar{v} \rightarrow c = A \bar{v}. \tag{14}
\]

Next, in (15), the calculation of the shape parameter \( (k) \) is done by using (c) value

\[
\ln c = - \frac{1}{k} \ln \bar{v} \rightarrow \ln c \cdot \frac{1}{k} \ln \bar{v}^2 = 0. \tag{15}
\]

E. Justus Moment Method (JMM)

The \( k \) and \( c \) values could be achieved as a function of mean wind speed and the standard deviation of wind speed [20], [21] in (16) and (17), respectively:

\[
k = \left( \frac{\sigma}{\bar{v}} \right)^{1.086}, \tag{16}
\]

\[
c = \frac{\bar{v}}{\Gamma \left( 1 + \frac{1}{k} \right)}. \tag{17}
\]

F. Power Density Method (PD)

This method uses the \( Epf \) variable that is calculated using EPFM. Equation (9) will be used in this method to calculate
The calculation of the $k$ value is as shown in (18)

$$\frac{\Gamma(1+3/k)}{\Gamma(1+1/k)} - Epf = 0. \quad (18)$$

Next, (19) will be used to calculate the $c$ value

$$c = \frac{\bar{v}}{\Gamma(1+1/k)}. \quad (19)$$

V. CALCULATION OF THE WIND ENERGY OF THE DATA

In this study, the Weibull parameter estimation methods have been compared for their accuracy of the results they provide for wind energy density as well. For this purpose, equation (20) has been used [19]

$$WED = \left[ \frac{1}{2} \int_0^\infty v^2 \times \left( \frac{k}{c} \right) \left( \frac{v}{c} \right)^{k-1} e^{(-\frac{v}{c})^k} d(v) \right] T = \left[ \frac{1}{2} \int_0^\infty \rho c^3 \left( 1 + \frac{3}{k} \right) \right] T, \quad (20)$$

where $T$ value represents time in hours.

VI. PERFORMANCE METRICS

The accuracy of the Weibull parameter estimation methods that have been studied in this paper has been compared by using some specific methods.

A. Wind Energy Error (WEE)

The wind energy calculated by using actual wind speed is different from the one calculated by using the Weibull distribution [19]. Equation (21) is used to check the accuracy of the calculated wind energy

$$WEE = \left| \frac{WED_w - WED_{act}}{WED_{act}} \right|, \quad (21)$$

where $WED_{act}$ is the actual wind energy density, $WED_w$ is the wind energy density calculated by using the Weibull distribution.

B. Root Mean Square Error (RMSE)

One of the methods to check the accuracy of the Weibull parameter estimation methods is RMSE

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{n} (v_i - w_i)^2}. \quad (22)$$

Equation (22) describes $v_i$ as the actual wind speed values, $w_i$ is the estimated values using probability density functions, and $n$ is the number of wind speed classes [23].

C. Coefficient of Determination ($R^2$)

Equation (23) describes this method [23]

$$R^2 = \frac{\sum_{i=1}^{n} (v_i - \bar{v})^2 - \sum_{i=1}^{n} (v_i - w_i)^2}{\sum_{i=1}^{n} (v_i - \bar{v})^2}. \quad (23)$$

VII. RESULTS AND DISCUSSION

All analyses have been realized for 32 cities using 6 different parameter estimation methods for two-parameter Weibull distribution. Figure 3 shows a comparison between the actual wind speed data and the Weibull parameter estimation methods, namely, GM, MLM, EPFM, WEIM, JMM, and PD for the regions of the studied cities. Figure 4 shows the comparison between the actual wind energy densities of each region and the calculated wind energy densities of the Weibull parameter estimation methods. Tables I–III show the estimated parameters by the Weibull parameter estimation methods and comparison of the accuracy of these methods by the use of some performance metrics (WEE, RMSE, and $R^2$) by using the wind data of the cities in three regions (Aegean, Central Anatolia, and Marmara) of Turkey.

The main comparison has been made by comparing the results that the methods have provided by using the performance metrics. The results for these metrics vary according to the model performance. The desired results for each metric are as follows.

For WEE, it is desired to have a number that is the closest to zero. For RMSE, results close to zero mean more accuracy [24]. For the Coefficient of Determination method, it is considered that the higher the value of ($R^2$), the more accurate the result is [25].

The results have been compared and the comparison has been given as a table after the comparison of all 6 methods with respect to WEE, RMSE, and $R^2$. Table IV shows the main evaluation of these methods according to Tables I–III and Fig. 3.

Firstly, the data has been compared per region. For the Aegean region, the PD method is better than the others with respect to WEE, followed by EPFM. The JMM is better with respect to RMSE, followed by MLM. Lastly, JMM is better with respect to $R^2$, followed by MLM.

For the Central Anatolia region, the PD method is better than the others with respect to WEE, followed by EPFM. The JMM is better with respect to RMSE, followed by MLM. Lastly, the JMM is better with respect to $R^2$, followed by MLM.

For the Marmara region, the PD is better than the others with respect to WEE, followed by EPFM. The JMM is better with respect to RMSE, followed by MLM. Lastly, the JMM is better with respect to $R^2$, followed by MLM.

The overall results show that for all regions the PD method is able to provide reliable data with respect to WEE, followed by EPFM. The JMM is better with respect to RMSE, followed by MLM. Lastly, the JMM is better with respect to $R^2$, followed by MLM for all regions.

The graphs in Fig. 3 show the comparison of the actual wind speed data to the probability distributions calculated using the different Weibull parameter estimation methods for the studied regions. Every two graphs have been created using the data of one region. Two graphs per region describe the probability distribution for the Aegean Region, and the absolute values for each method have been graphed to indicate the most efficient method. It can be concluded that while the GM and WEIM show better performance for some wind speed values, the overall best performance has been provided by the PD method.
Figure 4 shows the comparison between the actual wind energy densities and the calculated wind energy densities of the Weibull parameter estimation methods. For the Aegean region, while some of the estimation methods seem very close to the actual values, the method to provide the closest result to the actual value is PD, and EPFM closely follows. For the Central Anatolia region, some methods have competitive results. The main resolution is again that the PD is the closest, and the EPFM is the second closest for providing an accurate estimation of wind energy density. For the Marmara region, the comparison between the estimated and the actual wind energy densities has provided similar results. The main resolution is that while the PD has provided the closest estimations, the EPFM seem to be very close.

The Tables I–III include the comparison of the Weibull parameter estimation methods’ performances that have been studied in this paper, and Table IV includes the resulting evaluation according to these tables. In Table IV, all 6 methods have been ranked based on the performance metrics seen in Tables I–III. The values in this table are the mean for each method. The low rank means it is closer to the actual values, so it has preferable results for this study.

According to Table IV, the JMM has the lowest error values for RMSE and $R^2$ evaluations. On the other hand, the WEE evaluation has resulted with the PD method having the lowest error values.
Table IV concludes that for all three-performance metrics, the method with the lowest error value overall is the JMM. When all inspected, the one to provide the most efficient results has been the PD method. The GM method’s performance is low for all stations. Even so, it should be mentioned that since every region enjoys specific wind energy characteristics, the attained results in this paper regarding the performance of the parameters estimation methods of the Weibull distribution can only be extended to the regions with identical wind energy characteristics.

TABLE IV. THE RANKING OF THE WEIBULL PARAMETER ESTIMATION METHODS.

<table>
<thead>
<tr>
<th>Methods</th>
<th>WEE</th>
<th>RMSE</th>
<th>R²</th>
<th>Overall Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>JMM</td>
<td>1</td>
<td>0.04415</td>
<td>1</td>
<td>0.99566</td>
</tr>
<tr>
<td>MLM</td>
<td>4</td>
<td>0.04012</td>
<td>2</td>
<td>0.99648</td>
</tr>
<tr>
<td>EPFM</td>
<td>2</td>
<td>0.00003</td>
<td>3</td>
<td>0.99956</td>
</tr>
<tr>
<td>PD</td>
<td>1</td>
<td>2.8E-08</td>
<td>4</td>
<td>0.99859</td>
</tr>
<tr>
<td>WEIM</td>
<td>3</td>
<td>0.02009</td>
<td>6</td>
<td>0.94283</td>
</tr>
<tr>
<td>GM</td>
<td>6</td>
<td>0.08108</td>
<td>5</td>
<td>0.98580</td>
</tr>
</tbody>
</table>

The results show that when compared to the study conducted by Sumair, Aizad, Gardezi, ur Rehman, and Rehman [19], the WEIM has shown weaker results for the data of the regions that have been used in this paper.

VIII. CONCLUSIONS

In this paper, six different approaches, namely, GM, MLM, EPFM, WEIM, JMM, and PD have been analysed for their performances in estimating parameters of the two-parameter Weibull distribution. The comparison made to determine the most efficient method consisted of three steps:
1. Comparing the estimated wind speed probabilities;
2. Comparing the calculated wind energy densities;
3. Comparing the resulting performance criteria.

The results have been as follows:
– As it is clear from the Fig. 3, the PD method has been the closest at estimating the wind speed probability;
– As it is indicated in the Fig. 4, the PD method has made the closest predictions to the actual wind energy densities;
– While the JMM has showed a high performance, the PD method’s error values have also been low. Based on the results of the statistical analysis, it has been determined that the most accurate method for this study is the PD. The GM method showed weak performance for all stations.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES


This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 (CC BY 4.0) license (http://creativecommons.org/licenses/by/4.0/).