Analysis of Wind Speed Variations and Estimation of Weibull Parameters for Wind Power Generation in Malaysia


School of Civil Engineering
University of Science Malaysia
Engineering Campus, Pulau Pinang, Malaysia

ABSTRACT
The use of wind as an energy source is becoming popular because it is non-polluting and renewable. There is a pressing need in Malaysia to develop site-based technology on wind engineering, which can be used for optimal design of wind turbines and wind farming. The wind speed time series data of east coast of Peninsular Malaysia were analyzed to study the seasonal and diurnal wind characteristics. The Weibull parameters were computed using maximum likelihood method, the modified maximum likelihood method and the graphical method. The three methods were compared for their performance. RMSE values show that the modified maximum likelihood method performed better than the graphical method in frequency distribution format. The Weibull parameters for the individual months have also been computed and the Weibull frequency distributions for all the four monsoon seasons are discussed.

Keywords: wind speed, seasonal and diurnal variations, Weibull parameters, maximum likelihood method, modified maximum likelihood method, graphical method.

INTRODUCTION
Due to the present day energy demand and growing environmental consciousness, it has become imperative to supplement our energy base with clean and renewable sources of energy. Wind is one of the potential renewable energy sources which can be harnessed in a commercial way. A detailed knowledge of wind characteristics is required for efficient planning and implementation of any wind engineering project. Wind blows due to the warming and cooling of the earth’s atmosphere and the changes in temperature. In extreme cases the wind energy can be destructive in nature. Fortunately, most regions of the world experience moderate range of wind speeds that can allow human to extract energy from the wind. There is a pressing need in Malaysia to develop site-based technology on wind engineering, which can be used for optimal design of wind turbines and wind farming.

The seasonal variation in Malaysia is conventionally expressed by four climatic seasons, namely the first inter-monsoon (the month of April), southwest monsoon (mid-May to September), the second inter-monsoon (the month of October) and the northeast monsoon (November to March). The months of April and October are the transition periods between the two monsoons. So far, studies on wind characteristics of Malaysia are limited.

* Corresponding author E-mail: shams_ci@yahoo.com
The natural wind, as an energy source, is extremely variable. In order to predict the energy output of a wind energy conversion system (WECS), there has been considerable interest in finding a suitable statistical model of wind speed frequency distribution in the last few years. Sopian et al. (1995) analyzed the wind speed data for wind potential in Malaysia and used the conventional graphical method for computing the Weibull parameters. Garcia et al. (1998) used Weibull distribution graphical method and lognormal models for wind speed data. Wind characteristics of Oman were studied by Sulaiman et al. (2002) using the graphical method to determine the Weibull parameters. Seguro and Lambert (2000) recommended the modified maximum likelihood method for the estimation of Weibull parameters using the time series wind data. This was based on a limited number of wind speed data of 3 days and he suggested that the true evaluation of the method requires many months or years of measured wind speed data.

In the present study, the hourly measurements of surface winds from 1995 to 2000 at four meteorological stations namely, Mersing, Kuantan, Kuala Terengganu and Kota Baharu located in the East Coast of Peninsular Malaysia at a height of about 14 m above the ground have been analyzed to study the seasonal and diurnal wind characteristics. The Weibull parameters have been computed using the maximum likelihood method, the modified maximum likelihood method and the graphical method. The root mean square errors (RMSE) for the fitted Weibull distribution and the actual frequency distribution have been used to determine the best wind speed distribution.

WIND DATA

The wind speed data are generally available in time series format, in which each data represents an instantaneous sample wind speed or an average of wind speed taken at short intervals of time. The wind data at the four stations (Table 1) belongs to the second category. The data used in this study is the hourly wind speed data measured from January 1995 to December 2000. All these stations are along the East Coast of Peninsular Malaysia. The wind speed data can be converted to the frequency distribution format whereby the frequency with which the wind speed falls within various ranges or bins. The wind data of Kuala Terengganu in the frequency distribution format has been shown in Table 2. The methods described in the next section can be used to estimate the Weibull parameters with the given wind speed.

**Table 1. Location of wind speed stations used in the study.**

<table>
<thead>
<tr>
<th>Station</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mersing</td>
<td>2° 27' N</td>
<td>103° 50' E</td>
<td>13.4</td>
</tr>
<tr>
<td>Kuantan</td>
<td>3° 47' N</td>
<td>103° 13' E</td>
<td>14.0</td>
</tr>
<tr>
<td>Kuala Terengganu</td>
<td>5° 23' N</td>
<td>103° 06' E</td>
<td>14.0</td>
</tr>
<tr>
<td>Kota Baharu</td>
<td>6° 10' N</td>
<td>102° 17' E</td>
<td>14.0</td>
</tr>
</tbody>
</table>

**Table 2. Wind speed data in frequency distribution and cumulative frequency distribution formats at Kuala Terengganu.**

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Wind Speed (m/s)</th>
<th>Frequency (Number)</th>
<th>Frequency (%)</th>
<th>Cum. Freq. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25</td>
<td>3626</td>
<td>8.771</td>
<td>8.771</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>4696</td>
<td>11.359</td>
<td>20.13</td>
</tr>
<tr>
<td>3</td>
<td>1.25</td>
<td>5741</td>
<td>13.887</td>
<td>34.017</td>
</tr>
<tr>
<td>4</td>
<td>1.75</td>
<td>5179</td>
<td>12.528</td>
<td>46.545</td>
</tr>
</tbody>
</table>
THE WEIBULL 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:

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

where \( p(V) \) is the frequency or probability of occurrence of wind speed \( V \), \( \gamma \) is the Weibull scale parameter with unit equals to the wind speed unit and \( k \) is the unitless Weibull shape parameter. The higher value of \( \gamma \) indicates that the wind speed is higher, while the value of \( k \) shows the wind stability. Any probability equation in this paper including Equation (1) can be applied equally well whether the probability is in the form of relative (fractional or percent) or absolute (number of data points). The cumulative Weibull distribution function \( P(V) \) gives the probability of the wind speed exceeding the value \( V \). It is expressed by:

\[ P(V) = \exp \left\{ - \left( \frac{V}{\gamma} \right)^k \right\} \]  

If the shape parameter is exactly 2, then the distribution is called the cumulative Rayleigh distribution. Manufacturers of wind turbines generally give the standard performance of their machines using the Rayleigh distribution.

ESTIMATION OF WEIBULL PARAMETERS

The graphical method is generally used for computing the Weibull parameters [1-3]. In graphical method, the cumulative wind speed distribution is required. It is plotted on a special Weibull graph paper and a line of best fit is drawn. This procedure is nowadays carried out using computers by performing the linear regression analysis. Seguro and Lambert [4]
suggested another more accurate and robust computer based approach given by the maximum likelihood method for estimating Weibull parameters. However, their suggestions for using the alternate method were based only on three days of hourly wind speed data. It was further suggested by them that the true evaluation would require many months or years of measured wind speed data.

Maximum Likelihood Method

The maximum likelihood method is used for the wind speed data in time series format. The maximum likelihood method was used by Stevens and Smulders [5] in their study for estimation of parameters of the Weibull wind speed distribution for wind energy utilization purposes. The shape factor $k$ and the scale factor $c$ can be estimated by the following equations:

$$ k = \left( \frac{\sum_{i=1}^{n} V_i^k \ln(V_i)}{\sum_{i=1}^{n} V_i} - \frac{n \sum_{i=1}^{n} \ln(V_i)}{n} \right)^{-1} $$  \hspace{1cm} (3)

$$ c = \left( \frac{1}{n} \sum_{i=1}^{n} V_i^k \right)^{\frac{1}{k}} $$  \hspace{1cm} (4)

where $n$ is the number of non zero data values.

Modified Maximum Likelihood Method

The modified maximum likelihood method is applied by converting the wind speed time series in frequency distribution format. The shape parameter $k$ and the scale parameter $c$ are estimated using the following relationship:

$$ k = \left( \frac{\sum_{i=1}^{n} V_i^k \ln(V_i) P(V_i)}{\sum_{i=1}^{n} V_i^k P(V_i)} - \frac{n \sum_{i=1}^{n} \ln(V_i) P(V_i)}{P(V \geq 0)} \right)^{-1} $$  \hspace{1cm} (5)

$$ c = \left( \frac{1}{P(V \geq 0)} \sum_{i=1}^{n} V_i^k \right)^{\frac{1}{k}} $$  \hspace{1cm} (6)

where $V_i$ is the wind speed central to bin $i$, $n$ is the number of bins and $P(V_i)$ is the frequency with which the wind speed falls within bin $i$.

Equation 5 is solved using iterative procedure by guessing a suitable initial value of $k$. Equation 6 is then solved explicitly to estimate $c$.

Graphical Method
For two parameter Weibull distribution, several papers have appeared dealing with determination of shape parameter $k$ and the scale parameter $c$. The wind speed data is required in cumulative frequency distribution format for analysis using graphical method. Therefore, the time series data should first be sorted into bins.

The following relationship can be obtained from Equation 2:

$$\ln\{-\ln(P(V))\} = k \ln V_i - k \ln c$$  \hspace{1cm} (7)

Hence, a plot of $\ln\{-\ln(P(V))\}$ versus $\ln V_i$ represents a straight line with a slope of $K$ and a $y$ intercept $-k \ln c$. The logarithmic transformation becomes the basis of this method. In order to generate a straight line, observations of calms should be omitted from the data. The line of best fit is then determined using a least square regression.

**WIND SPEED CHARACTERISTICS**

The wind speed data at four stations are statistically analysed. Figure 1 shows the mean daily wind speed values at Kuala Terengganu for the year 1995. It is clear from the diagram that the wind data has the seasonal effects. The maximum annual mean wind speed is at Mersing which is about 3.0 m/s. The mean monthly wind speed is higher in January and December and it varies from 4.5 to 5.0 m/s. The diurnal variations in mean and maximum wind speeds for Mersing in the months of April, July, October and January, representing all the four monsoon seasons, are shown in Figure 2. The figure highlights several interesting characteristics of the wind speed at Mersing. The average maximum speed is 7-8 m/s in the month of January and it varies from 4 to 6 m/s in the month of July. The months of January and October show the highest and the lowest average wind speed respectively.

![Figure 1. Mean daily wind speed values at Kuala Terengganu.](image)

The diurnal variation of the hourly mean wind speed for all the years shows only one peculiar feature as depicted in Figure 3 for Kuala Terengganu. There is a clear bell shaped trend which is evident throughout the year. This similar pattern has been observed for all the years for the other three stations. This is due to the effect of solar heating balance. The wind speeds are usually reduced during the night but increased during the day [6]. From the point of view of wind energy generation, it is advantageous that most of the wind energy is produced during the day since electricity consumption is higher than at night.
Figure 2. Monthly average diurnal wind speed at Mersing for April, July, October and January, representing the four seasons.

Figure 3. Diurnal wind speed variation for all the individual years for Kuala Terengganu.

WEIBULL DISTRIBUTION PARAMETERS

The Weibull distribution parameters for the wind speed data at all the four stations were estimated and presented in Table 3 by the three methods as mentioned above. Proper care has been taken wherever required to avoid the log of zero values. The results of the parameters are being presented in Table 3.
Table 3. Calculated Weibull parameters by the three methods.

<table>
<thead>
<tr>
<th>Station</th>
<th>Bin Size</th>
<th>Max. Likelihood</th>
<th>Mod. Max. Likelihood</th>
<th>Graphical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>K</td>
<td>C</td>
<td>RMSE</td>
</tr>
<tr>
<td>Mersing</td>
<td>0.25</td>
<td>1.800</td>
<td>3.050</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>1.850</td>
<td>3.120</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.850</td>
<td>3.120</td>
<td>2.05</td>
</tr>
<tr>
<td>Kuantan</td>
<td>0.25</td>
<td>1.050</td>
<td>2.045</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>1.200</td>
<td>2.200</td>
<td>2.65</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.389</td>
<td>2.330</td>
<td>2.47</td>
</tr>
<tr>
<td>Terengganu</td>
<td>0.25</td>
<td>1.160</td>
<td>2.110</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>1.260</td>
<td>2.230</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.395</td>
<td>2.309</td>
<td>2.13</td>
</tr>
<tr>
<td>Kota Baharu</td>
<td>0.25</td>
<td>0.991</td>
<td>1.980</td>
<td>2.55</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>1.133</td>
<td>2.150</td>
<td>3.18</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.329</td>
<td>2.308</td>
<td>3.41</td>
</tr>
</tbody>
</table>

Avg.    | Avg.    |
2.23    | 2.83    |

For the modified maximum likelihood and the graphical methods the Weibull parameters have been determined for bin sizes 0.25 m/s, 0.50 m/s and 1.00 m/s. For testing the accuracy of the parameters, the root mean square errors (RMSE) were calculated at different bin sizes for the frequency distribution and the estimated distribution from the Weibull parameters obtained. Table 3 reveals some important findings. The comparison of RMSE shows that the results obtained, using the maximum likelihood method, are better than the modified maximum likelihood method and graphical method without considering the size of the bin. The table also revealed one important finding i.e. the accuracy of both the modified maximum likelihood and the graphical methods improved with the reduced bin size. This is because large bin size results in less detailed information that is available to the algorithm. The main reason for the poor results from the graphical method is that the least squares regression is carried out in its cumulative frequency distribution and not on its actual wind speed data. This way, all the points are given equal weightage although some of the bins may have larger number of data points than others. This problem can be solved to a certain extent by reducing the bin size, but the bin size cannot be reduced below the resolution of the wind speed data. The wind speed probability distribution and the fitted Weibull distribution by maximum likelihood method have been shown in Figure 4 for all the wind speed data at Mersing.

The Weibull parameters, computed for the data at Mersing for individual months are shown in Table 4. The average of RMSE values for the maximum likelihood method is lowest here. This also strengthens the statement that maximum likelihood method is better for the computation of the Weibull parameters.

Based on the parameters computed by the maximum likelihood method, the Weibull distributions for Mersing for all the four monsoon categories have been shown in Figure 5. The southwest monsoon period (May to September) exhibits almost the similar distribution pattern for all the five months. However, the pattern varies for each month during the...
northeast monsoon. During the transition period (the months of April and October), the Weibull distribution is similar to that of the southwest monsoon.

![Figure 4](image1.png)

**Figure 4.** (a) Frequency distribution and Weibull distribution. (b) Cumulative frequency distribution.

![Figure 5](image2.png)

**Figure 5.** Weibull Probability distribution functions for all the four seasons.

### Table 4. The Weibull Parameters for individual month for wind speed data at Mersings

<table>
<thead>
<tr>
<th>Months</th>
<th>Max. Likelihood</th>
<th>Mod. Max. Likelihood</th>
<th>Graphical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CONCLUSIONS
At all the stations under study, the wind energy generation can be higher in the northeast monsoon period as the winds are strongest during this period. The average wind speed at Mersing is about 3 m/s which is the place of maximum wind speed in Peninsular Malaysia. Medium size or small wind turbines can be used to generate electricity in the region. The true estimation of Weibull parameters is important for the manufacturer of the wind turbines who need to know the performance of their turbines. The maximum likelihood method, the modified maximum likelihood method and the graphical method have been evaluated using real life wind speed data. The root mean square error was always found to be lower for the maximum likelihood method. The graphical method is less accurate and its accuracy is affected by the bin size in the cumulative distribution format. Therefore, the maximum likelihood method is recommended for the estimation of the Weibull parameters.

REFERENCES: