A STUDY ON WEIBULL DISTRIBUTION FOR ESTIMATING THE PARAMETERS

Paritosh BHATTACHARYA
CEM, Kolaghat, Midnapore, India

E-mail: p_bhattacharya2001@yahoo.co.in
Rakhi BHATTACHARJEE
BITM, Shantiniketan, India

Abstract: The wind resource varies with of the day and the season of the year and even some extent from year to year. Wind energy has inherent variances and hence it has been expressed by distribution functions. In this paper, we present some methods for estimating Weibull parameters, namely, shape parameter (k) and scale parameter (c). The Weibul distribution is an important distribution especially for reliability and maintainability analysis. The suitable values for both shape parameter and scale parameters of Weibull distribution are important for selecting locations of installing wind turbine generators. The scale parameter of Weibull distribution also important to determine whether a wind farm is good or not. The presented method is the analytical methods and computational experiments on the presented methods are reported.

Key words: Wind Speed; Probability Distribution; Weibull Distribution; Linear Least Square Method

1. Introduction

Today, most electrical energy is generated by burning huge fossil fuels and special weather conditions such as acid rain and snow, climate change, urban smog, regional haze, several tornados, etc., have happened around the whole world. It is now clear that the installation of a number of wind turbine generators can effectively reduce environmental pollution, fossil fuel consumption, and the costs of overall electricity generation. Although wind is only an intermittent source of energy, it represents a reliable energy resource from a long-term energy policy viewpoint. Among various renewable energy resources, wind power energy is one of the most popular and promising energy resources in the whole world today. At a specific wind farm, the available electricity generated by a wind power generation system depends on mean wind speed (MWS), standard deviation of wind speed, and the location of installation. Since year-to-year variation on annual MWS is hard to predict, wind speed variations during a year can be well characterized in terms of a probability distribution function (pdf). This paper also addresses the relations among MWS, its standard deviation, and two important parameters of Weibull distribution.

2. Weibull distribution

The Weibull distribution is characterized by two parameters, one is the shape parameter k (dimensionless) and the other is the scale parameter c (m/s)
The cumulative distribution function is given by

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

And the probability function is given by

$$f(v) = \frac{dF(v)}{dv} = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} \exp\left[-\left(\frac{v}{c}\right)^k\right] \ldots\ldots(2)$$

The average wind speed can be expressed as

$$\bar{v} = \int_0^\infty v f(v) dv = \int_0^\infty \frac{vk}{c}\left[(\frac{v}{c})^{k-1}\right] \exp\left[-(\frac{v}{c})^k\right] dv \ldots\ldots(3)$$

Let $x = (\frac{v}{c})^k$, $x^{\frac{1}{k}} = \frac{v}{c}$ and $dx = \frac{k}{c}(\frac{v}{c})^{k-1} dv$

Equation (3) can be simplified as

$$\bar{v} = c\int_0^\infty x^{\frac{1}{k}} \exp(-x) dx \ldots\ldots\ldots\ldots(4)$$

By substituting a Gamma Function

$\Gamma(n) = \int_0^\infty e^{-x} x^{n-1} dx$ into (4) and let $y = 1 + \frac{1}{k}$ then we have

$$\bar{v} = c\Gamma\left(1 + \frac{1}{k}\right) \ldots\ldots\ldots\ldots(5)$$

The standard deviation of wind speed v is given by $\sigma = \sqrt{\int_0^\infty (v - \bar{v})^2 f(v) dv} \ldots\ldots(6)$

i.e.

$$\sigma = \sqrt{\int_0^\infty (v^2 - 2v\bar{v} + \bar{v}^2) f(v) dv}$$

$$= \sqrt{\int_0^\infty v^2 f(v) dv - 2\bar{v}\int_0^\infty v f(v) dv + \bar{v}^2}$$

$$= \sqrt{\int_0^\infty v^2 f(v) dv - 2\bar{v}.\bar{v} + \bar{v}^2} \ldots\ldots\ldots\ldots(7)$$

Use

$$\int_0^\infty v^2 f(v) dv = \int_0^\infty v^2 \frac{k}{c}(\frac{v}{c})^{k-1} dv = \int_0^\infty c^2 x^{\frac{2}{k}} \frac{k}{c}(\frac{v}{c})^{k-1} dv = \int_0^\infty c^2 x^{\frac{2}{k}} \exp(-x) dx \ldots\ldots(8)$$

And put $y = 1 + \frac{2}{k}$, then the following equation can be obtained

$$\int_0^\infty v^2 f(v) dv = c^2\Gamma(1 + \frac{2}{k}) \ldots\ldots\ldots(9)$$

Hence we get

$$\sigma = \left[c^2\Gamma(1+\frac{2}{k}) - c^2\Gamma^2(1+\frac{1}{k}) \right]^{\frac{1}{2}}$$

$$= c\sqrt{\Gamma(1+\frac{2}{k}) - \Gamma^2(1+\frac{1}{k})} ....(10)$$

i.e. $\frac{1}{1-F(v)} = \exp\left[\left(\frac{v}{c}\right)^k\right]$........................(13)

i.e. $\ln\{\frac{1}{1-F(v)}\} = \left[\left(\frac{v}{c}\right)^k\right]$.......................(14)

But the cumulative Weibull distribution function is transformed to a linear function like below:

Again $\ln\ln\{\frac{1}{1-F(v)}\} = k\ln v - k\ln c$.....................(15)

Equation (14) can be written as $Y = bX + a$

where $Y = \ln\ln\{\frac{1}{1-F(v)}\}$, $X = \ln v$, $a = -k\ln c$, $b = k$

By Linear regression formula

Linear Least Square Method (LLSM)

Least square method is used to calculate the parameter(s) in a formula when modeling an experiment of a phenomenon and it can give an estimation of the parameters. When using least square method, the sum of the squares of the deviations S which is defined as below, should be minimized.

$$S = \sum_{i=1}^{n} w_i^2 [y_i - g(x_i)]^2 ....(11)$$

In the equation, xi is the wind speed, yi is the probability of the wind speed rank, so (xi, yi) mean the data plot, wi is a weight value of the plot and n is a number of the data plot.

The estimation technique we shall discuss is known as the Linear Least Square Method (LLSM). It is so commonly applied in engineering and mathematics problem that is often not thought of as an estimation problem. The linear least square method (LLSM) is a special case for the least square method with a formula which consists of some linear functions and it is easy to use. And in the more special case that the formula is line, the linear least square method is much easier. The Weibull distribution function is a non-linear function, which is

$$F(v) = 1 - \exp\left[-\left(\frac{v}{c}\right)^k\right]....................(12)$$

$$b = \frac{N\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{N\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2} \text{..........} \quad \text{.......(16)}$$

$$a = \frac{\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i Y_i}{N\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2} \text{.........(17)}$$

4. Maximum Likelihood Estimator(MLE)

The method of maximum likelihood (Harter and Moore (1965a), Harter and Moore (1965b), and Cohen (1965)) is a commonly used procedure because it has very desirable properties.

Let $x_1, x_2, \ldots\ldots\ldots\ldots x_n$ be a random sample of size *n* drawn from a probability density function $f(x,\theta)$ where θ is an unknown parameter. The likelihood function of this random sample is the joint density of the *n* random variables and is a function of the unknown parameter. Thus

$$L = \prod_{i=1}^{n} f_{X_i}(x_i, \theta) \text{..........} \quad (18)$$

is the Likelihood function. The Maximum Likelihood Estimator (MLE) of θ, say $\overline{\theta}$, is the value of *θ*, that maximizes *L* or, equivalently, the logarithm of *L* . Often, but not always, the MLE of q is a solution of

$$\frac{dLogL}{d\theta} = 0 \text{..............................(19)}$$

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

$$L(x_1, x_2, .., x_n, k, c) = \prod_{i=1}^{n} \left(\frac{k}{c}\right)\left(\frac{x_i}{c}\right)^{k-1} e^{-\left(\frac{x_i}{c}\right)^k} \text{......(20)}$$

On taking the logarithms of (20), differentiating with respect to k and c in turn and equating to zero, we obtain the estimating equations

$$\frac{\partial \ln L}{\partial k} = \frac{n}{k} + \sum_{i=1}^{n} \ln x_i - \frac{1}{c}\sum_{i=1}^{n} x_i^k \ln x_i = 0 \text{...................(21)}$$

$$\frac{\partial \ln L}{\partial c} = \frac{-n}{c} + \frac{1}{c^2}\sum_{i=1}^{n} x_i^k = 0 \text{....................................(22)}$$

On eliminating c between these two above equations and simplifying, we get

$$\frac{\sum_{i=1}^{n} x_i^k \ln x_i}{\sum_{i=1}^{n} x_i^k} - \frac{1}{k} - \frac{1}{n}\sum_{i=1}^{n} \ln x_i = 0 \text{......................(23)}$$

which may be solved to get the estimate of k. This can be accomplished by Newton-Raphson method. Which can be written in the form

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \text{.......... ....}(24)$$

Where

$$f(k) = \frac{\sum_{i=1}^{n} x_i^k \ln x_i}{\sum_{i=1}^{n} x_i^k} - \frac{1}{k} - \frac{1}{n}\sum_{i=1}^{n} \ln x_i \text{..}(25)$$

And

$$f'(k) = \sum_{i=1}^{n} x_i^k (\ln x_i)^2 - \frac{1}{k^2}\sum_{i=1}^{n} x_i^k (k \ln x_i - 1) - (\frac{1}{n}\sum_{i=1}^{n} \ln x_i)(\sum_{i=1}^{n} x_i^k \ln x_i)\text{......}(26)$$

Once k is determined, c can be estimated using equation (22) as

$$c = \frac{\sum_{i=1}^{n} x_i^k}{n} \text{..........................}(27)$$

5. Results and Discussions

When a location has c=6 the pdf under various values of k are shown in Fig. 1. A higher value of k such as 2.5 or 4 indicates that the variation of Mean Wind speed is small. A lower value of k such as 1.5 or 2 indicates a greater deviation away from Mean Wind speed.

When a location has k=3 the pdf under various valus of c are shown in Fig.2. A higher value of c such as 12 indicates a greater deviation away from Mean Wind speed.

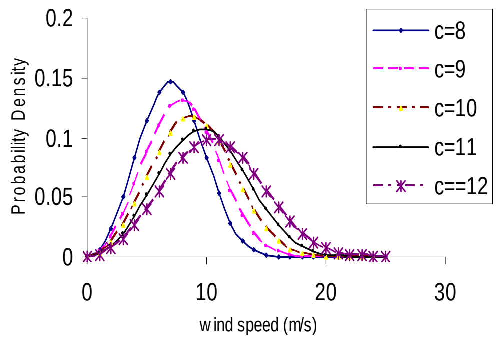

Fig.2: Weibull Distribution Density versus wind speed under a constant value of k=3 and different values of c

Fig. 3 represents the characteristic curve of $\Gamma\left(1+\frac{1}{k}\right)$. versus shape parameter k. The values of $\Gamma\left(1+\frac{1}{k}\right)$. varies around .889 when k is between 1.9 to 2.6.

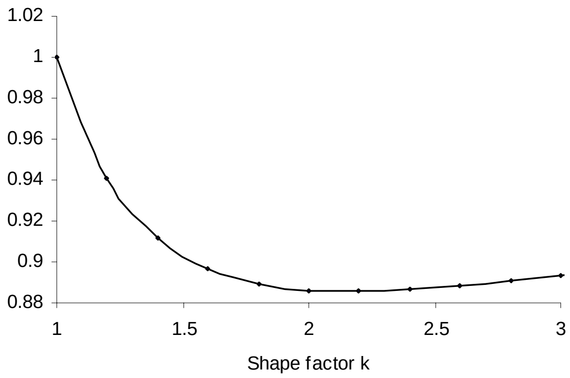

Fig. 3: Characteristic curve of Γ(1+1/k) versus Shape parameter k

Fig.4 represents the characteristic curve of $\frac{c}{\bar{v}}$ versus shape parameter k .Normally the wind speed data collected at a specified location are used to calculate Mean Wind speed. A good estimate for parameter c can be obtained from Fig.4 as $c = 1.128\bar{v}$ where k ranges from 1.6 to 4. If the parameter k is less than unity , the ratio $\frac{c}{\bar{v}}$ decrease rapidly. Hence c is directly proportional to Mean Wind speed for $1.6 \le k \le 4$ and Mean Wind speed is mainly affected by c. The most good wind farms have k in this specified range and estimation of c in terms of $\bar{v}$ may have wide applications.

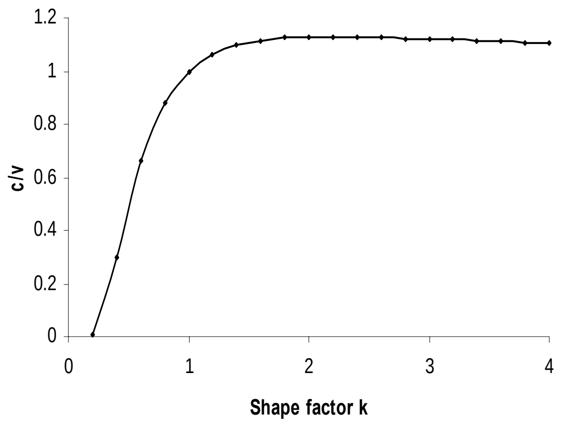

Fig.4: Characteristic curve of c/ṽ versus shape parameter k

Example: Consider the following example where x_i represents the Average Monthly Wind Speed (m/s) at kolkata (from 1st March, 2009 to 31st March, 2009)

March2009	Wind Speed (m/s)	March, 2009	Wind Speed (m/s)
1	0.56	17	0.28
2	0.28	18	0.83
3	0.56	19	1.39
4	0.56	20	1.11
5	1.11	21	1.11
6	0.83	22	0.83
7	1.11	23	0.56
8	1.94	24	0.83
9	1.11	25	1.67
10	0.83	26	1.94
11	1.11	27	1.39
12	1.39	28	0.83
13	0.28	29	2.22
14	0.56	30	1.67
15	0.28	31	2.22
16	0.28		

Also let $F(x_i) = \dfrac{i}{n+1}$ and using equations (16) and (17) we get k= 1.013658 and c=29.9931

But if we apply maximum Likelihood Method we get k = 1.912128 and c=1.335916. There is a huge difference in value of c by the above two methods. This is due to the mean rank of $F(x_i)$ and k value is tends to unity.

6. Conclusions

In this paper, we have presented two analytical methods for estimating the Weibull distribution parameters. The above results will help the scientists and the technocrats to select the location for Wind Turbine Generators.

References

1. Mann, N. R., Schafer, R. E., and Singpurwalla, N. D., **Methods for statistical analysis of reliability and life data**, John Wiley and Sons, New York, 1974
2. Engelhardt, M., "**On simple estimation of the parameters of the Weibull or extreme-value distribution**", Technometrics, Vol. 17, No. 3, August 1975.
3. Mann, N. R. and K. W. Fertig , "**Simplified efficient point and interval estimators of the Weibull parameters**", Technometrics, Vol. 17, No. 3, August 1975
4. Cohen, A. C., "**Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples",** Technometrics, Vol. 7, No. 4, November 1965.
5. Harter, H. L. and A. H. Moore, "**Point and interval estimators based on order statistics, for the scale parameter of a Weibull population with known shape parameter**", Technometrics, Vol. 7, No. 3, August 1965a
6. Harter, H. L. and A. H. Moore, "**Maximum likelihood estimation of the parameters of Gamma and Weibull populations from complete and from censored samples**", Technometrics, Vol. 7, No. 4, November 1965b
7. Stone, G. C. and G. Van Heeswijk, "**Parameter estimation for the Weibull distribution**", IEEE Trans. On Elect Insul. VolEI-12, No-4, August, 1977.
8. P. Gray and L. Johnson, **Wind Energy System** Upper Saddle River, NJ: Prentice-Hall, 1985