

An Appraisal on Some Methods for Estimating the 2-Parameter Weibull Distribution with Application to Wind Speeds Sample

Patrick Osatohanmwun^{*}, Francis O Oyegue¹, Joseph E Osemwenkhae²
and Nosakhare Ekhosuehi³

^{*,1,2,3} Department of Mathematics, University of Benin, Benin City, Nigeria

*** Corresponding Author:** Profpato2014@gmail.com

 <https://orcid.org/0000-0002-9332-6020>

Received: 27th of May 2017 / Revised: 15th November 2017 / Published: 07th December 2018
©IAAppStat-SL2017

ABSTRACT

Six methods for estimating the Weibull shape and scale parameters are considered and compared in this paper. These methods are: the least squares method, weighted least squares method, method of moments, energy pattern factor method, method of L-moments and the maximum likelihood method. A simulation study as well as application to a real data set (wind speeds sample) was used to test the performance of different methods using the smallest mean square error criterion. Results from the simulation study indicated that the maximum likelihood method is the most efficient method when dealing with large sample sizes, while the weighted least squares method, method of moments and the method of L-moments were quite efficient for small and moderate sample sizes. The maximum likelihood method produced the best method when all six methods were applied to a wind speeds sample by possessing the smallest mean square error. A very useful result obtained from the study is that the weighted least squares method which performed considerably well in estimating the Weibull parameters. This is a rare incidence in many studies.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Keywords and phrase: L-moments; maximum likelihood; moments; least squares; weighted least squares and energy pattern factor

1. Introduction

Wind speed is a classic example of a stochastic variable. Due to this stochastic nature, its characteristics and behaviors in a given location can be captured by fitting specific probability distributions to a given wind speeds sample(s) collected over time. The 2-parameter Weibull distribution has become the most widely used probability distribution in wind speed analysis[1 – 9]. Results from many studies have suggested that it is very adequate in fitting many wind regimes whether in tropical or extra tropical latitudes due to its flexibility. Central to the use of the Weibull model in wind speed analysis is the estimation of the parameters of the Weibull distribution. Several methods for estimating the Weibull parameters have been proposed over the years. Generally, these methods can simply be classified into two broad categories, namely: graphical methods and statistical methods.

The graphical methods which has not only be found to be very simple to implement but also having the direct advantage of reporting visibly whether or not the Weibull model fits the observed sample well, as well as offer initial estimates that can be used as starting points for the statistical methods [10]. However, estimates obtained from the graphical methods have been found to be less accurate than those obtained using statistical methods. Some commonly used graphical methods include the empirical cumulative distribution function plot methods[11], Weibull probability plot [12] and hazard rate plot[11]. Statistical methods on the other hand, are based on mathematical and statistical theory which makes them more general and applicable to different data types. The Weibull estimators obtained using statistical methods are generally more accurate but complex to deal with than their graphical counterpart. While it may not be possible to obtain the asymptotic properties of the estimators based on the graphical methods, this is however possible with the statistical methods. Some

commonly used statistical methods include: method of moment [13], method of percentiles, method of maximum Likelihood estimation [11,13] , L-moments estimators [14] , interval estimation [15,16,17], Bayesian method [18] and the energy pattern factor method [19]. Many works have appeared in the literature over the years geared towards the determination of the best method for estimating the Weibull shape and scale parameters [20 – 25].

The rest of this paper is organized as follows. The Weibull distribution is introduced in section 2, some methods for estimating the Weibull parameters are considered in section 3. A simulation study and results from the simulation study is presented in sections 4 and 5 respectively. Application to wind speeds sample as well as results obtained from the application are contained in sections 6 and 7 respectively with conclusion in section 8.

2. Weibull Distribution

Weibull distribution is a continuous probability distribution named after Waloddi Weibull who described it in detail in 1951, although it was first identified by Maurice Frechet in 1927 and first applied by Rosin and Rammler in 1933 to describe the size distribution of particles [26]. The probability density function (PDF) and cumulative distribution function (CDF) of the Weibull distribution given by [27] are respectively of the form

$$f(x) = \frac{k}{c} \left(\frac{x}{c}\right)^{k-1} \exp\left(-\left(\frac{x}{c}\right)^k\right), \quad (1)$$

$$F(x) = 1 - \exp\left(-\left(\frac{x}{c}\right)^k\right), \quad (2)$$

$$x > 0, k, c, > 0,$$

where k and c are shape and scale parameters respectively. When applied to wind speed studies, the parameter k is a dimensionless shape parameter that shows the peakedness of the distribution of wind speeds at the measuring location, and for varying values of k the distribution of the wind speeds takes the shape of other distributions. For $k = 1$, the distribution is exponential, for $k = 2$ it is Rayleigh distribution, and for $k = 3.4$, the distribution becomes approximately normal. The

scale parameter c measured in the dimension of the wind speed (m/s) shows the dispersion of the wind speeds in a given location. The inverse CDF or quantile function, $x(F)$ of the Weibull distribution is obtained by finding the root of the equation $F(x(F)) = F$. This gives the quantile function as

$$x(F) = c(-\ln(1 - F))^{1/k}; \quad 0 < F < 1. \quad (3)$$

3. Methods for Estimating the Weibull Shape and Scale Parameters

In this section we consider some methods for estimating the Weibull parameters. These methods include the least squares method (LSM), weighted least squares method (WLSM), method of moments (MOM), energy pattern factor method (EPFM), method of L-moments (MOL) and maximum likelihood method (MLM).

3.1. Least squares method (LSM)

The LSM is extensively used in many practical problems in estimating the parameter(s) of a model. Consider the model of the form

$$Y_i = \beta_0 + \beta_1 X_i \quad i = 1, 2, \dots, n. \quad (4)$$

The least squares estimates of the parameters β_0 and β_1 are the values of the parameters which minimize the function

$$Z(\beta_0, \beta_1) = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2, \quad (5)$$

and thus, the estimators of β_0 and β_1 are given respectively by

$$\widehat{\beta}_0 = \frac{1}{n} \sum_{i=1}^n Y_i - \widehat{\beta}_1 \frac{1}{n} \sum_{i=1}^n X_i, \quad (6)$$

$$\widehat{\beta}_1 = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}. \quad (7)$$

Now, to estimate the Weibull parameters k and c using the LSM, we first admit the linearization of the Weibull CDF given by (2) which has the form

$$\ln(-\ln(1 - F(x))) = k \ln x - k \ln c. \tag{8}$$

Observe that (8) is of the form given by (4) with $Y_i = \ln(-\ln(1 - F(x_i)))$, $X_i = \ln x_i$, $\beta_0 = -k \ln c$ and $\beta_1 = k$. Suppose we have a random sample X_1, X_2, \dots, X_n of size n from the Weibull distribution with order statistics $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ and let $x_{(1)} < x_{(2)} < \dots < x_{(n)}$ be observed ordered observations, the CDF $F(x)$ can be estimated using the mean rank

$$\hat{F}(x_{(i)}) = \frac{i}{n + 1}, \tag{9}$$

Where i denotes the i^{th} smallest value of $X_{(1)} < X_{(2)} < \dots < X_{(n)}$, $i = 1, 2, \dots, n$. Taking (8) as (4), it follows that the estimates of β_0 and β_1 becomes

$$\hat{\beta}_0 = \frac{1}{n} \sum_{i=1}^n \ln(-\ln(1 - \hat{F}(x_{(i)}))) - \hat{\beta}_1 \frac{1}{n} \sum_{i=1}^n \ln x_{(i)}, \tag{10}$$

$$\hat{\beta}_1 = \frac{n \sum_{i=1}^n \ln x_{(i)} \ln(-\ln(1 - \hat{F}(x_{(i)}))) - \sum_{i=1}^n \ln x_{(i)} \sum_{i=1}^n \ln(-\ln(1 - \hat{F}(x_{(i)})))}{n \sum_{i=1}^n \ln^2 x_{(i)} - (\sum_{i=1}^n \ln x_{(i)})^2}. \tag{11}$$

Also, $\beta_1 = k$ means that

$$\hat{k} = \hat{\beta}_1, \tag{12}$$

and $\beta_0 = -k \ln c$ implies that $c = \exp\left(-\left(\frac{\beta_0}{k}\right)\right)$, hence

$$\hat{c} = \exp\left(-\left(\frac{\sum_{i=1}^n \ln(-\ln(1 - \hat{F}(x_{(i)}))) - \hat{k} \sum_{i=1}^n \ln x_{(i)}}{n \hat{k}}\right)\right). \tag{13}$$

The LSM estimates of the Weibull shape and scale parameters are obtained using (12) and (13) respectively.

3.2. Weighted Least Squares Method (WLSM)

For the model given by (4), the WLSM estimates of the parameters β_0 and β_1 are the values of the parameters which minimizes the function

$$Z(\beta_0, \beta_1) = \sum_{i=1}^n W_i (Y_i - \beta_0 - \beta_1 X_i)^2, \quad (14)$$

where $W_i, i = 1, 2, \dots, n$ is a weight factor. Hence the WLSM estimators of β_0 and β_1 are given by

$$\widehat{\beta}_0 = \frac{\sum_{i=1}^n W_i Y_i - \widehat{\beta}_1 \sum_{i=1}^n W_i X_i}{\sum_{i=1}^n W_i}, \quad (15)$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n W_i \sum_{i=1}^n W_i X_i Y_i - \sum_{i=1}^n W_i X_i \sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i \sum_{i=1}^n W_i X_i^2 - (\sum_{i=1}^n W_i X_i)^2}. \quad (16)$$

Taking (8) to be of the form (4) and using similar arguments as put forward for the LSM, one readily obtain the following results

$$\widehat{\beta}_0 = \frac{\sum_{i=1}^n w_{(i)} \ln \left(-\ln \left(1 - \widehat{F}(x_{(i)}) \right) \right) - \widehat{\beta}_1 \sum_{i=1}^n w_{(i)} \ln x_{(i)}}{\sum_{i=1}^n w_{(i)}}, \quad (17)$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n w_{(i)} \sum_{i=1}^n w_{(i)} \ln x_{(i)} \ln \left(-\ln \left(1 - \widehat{F}(x_{(i)}) \right) \right) - \sum_{i=1}^n w_{(i)} \ln x_{(i)} \sum_{i=1}^n w_{(i)} \ln \left(-\ln \left(1 - \widehat{F}(x_{(i)}) \right) \right)}{\sum_{i=1}^n w_{(i)} \sum_{i=1}^n w_{(i)} \ln^2 x_{(i)} - (\sum_{i=1}^n w_{(i)} \ln x_{(i)})^2}. \quad (18)$$

It follows that

$$\widehat{k} = \widehat{\beta}_1, \quad (19)$$

and

$$\widehat{c} = \exp \left(- \left(\frac{\sum_{i=1}^n w_{(i)} \ln \left(-\ln \left(1 - \widehat{F}(x_{(i)}) \right) \right) - \widehat{k} \sum_{i=1}^n w_{(i)} \ln x_{(i)}}{\widehat{k} \sum_{i=1}^n w_{(i)}} \right) \right). \quad (20)$$

In this paper, the weight function $w_{(i)} \quad i = 1, 2, 3, \dots, n$ proposed by [21] is used. This is defined as

$$w_{(i)} = \left(\left(1 - \widehat{F}(x_{(i)}) \right) \ln \left(1 - \widehat{F}(x_{(i)}) \right) \right)^2. \quad (21)$$

3.3. Method of moments (MOM)

It is a standard practice in statistics when fitting a parametric family of distribution to a data set to estimate the parameters by equating the sample moments to those of the fitted distribution. For a continuous random variable X with PDF $f(x)$, the r^{th} moment about the origin is defined as

$$\mu'_r = \int_{-\infty}^{\infty} x^r f(x) dx \quad r = 1, 2, \dots \quad (22)$$

If the random variable X follows the Weibull distribution, the r^{th} moment about the origin is given by

$$\mu'_r = \int_0^{\infty} x^r \frac{k}{c} \left(\frac{x}{c}\right)^{k-1} \exp\left(-\left(\frac{x}{c}\right)^k\right) dx \quad r = 1, 2, \dots \quad (23)$$

Evaluating (23) gives

$$\mu'_r = c^r \Gamma\left(1 + \frac{r}{k}\right), \quad (24)$$

where $\Gamma(\cdot)$ is the complete gamma function defined as

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx. \quad (25)$$

The mean of the Weibull random variable X is obtained from (24) by setting $r = 1$ to obtain

$$\begin{aligned} \mu'_1 &= c \Gamma\left(1 + \frac{1}{k}\right). \end{aligned} \quad (26)$$

The variance defined as $\sigma^2 = \mu'_2 - (\mu'_1)^2$ for the Weibull distribution is thus given as

$$\sigma^2 = c^2 \left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right) \right]. \quad (27)$$

For a random sample X_1, X_2, \dots, X_n of size n , the sample r^{th} moment is defined as

$$m_r = \frac{1}{n} \sum_{i=1}^n X_i^r \quad r = 1, 2, \dots \quad (28)$$

The sample mean is obtained from (28) by setting $r = 1$ to obtain

$$\begin{aligned} m_1 &= \bar{X} \\ &= \frac{1}{n} \sum_{i=1}^n X_i, \end{aligned} \quad (29)$$

and consequently, the sample variance is given by

$$S^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}. \quad (30)$$

The MOM estimators of the Weibull shape and scale parameters k and c is obtained by equating the r^{th} population moment in (24) to the r^{th} sample moment in (28). In particular, from (26) we obtain

$$c = \frac{\mu'_1}{\Gamma\left(1 + \frac{1}{k}\right)}. \quad (31)$$

Substituting (31) into (27) gives

$$\begin{aligned} \sigma^2 &= \frac{(\mu'_1)^2}{\Gamma^2\left(1 + \frac{1}{k}\right)} \left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right) \right], \quad (32) \end{aligned}$$

hence,

$$\frac{\Gamma\left(1 + \frac{2}{k}\right)}{\Gamma^2\left(1 + \frac{1}{k}\right)} - \left(\frac{\sigma}{\mu'_1}\right)^2 - 1 = 0. \quad (33)$$

For a sample of observations, (33) becomes

$$\frac{\Gamma\left(1 + \frac{2}{k}\right)}{\Gamma^2\left(1 + \frac{1}{k}\right)} - \left(\frac{S}{\bar{X}}\right)^2 - 1 = 0. \quad (34)$$

Observe that (34) is a function of k . The estimate \hat{k} of the shape parameter k is obtained by numerically finding the root of the equation given in (34). Once \hat{k} is obtained, the scale parameter c is then obtained from (31) as

$$\hat{c} = \frac{\bar{X}}{\Gamma\left(1 + \frac{1}{\hat{k}}\right)}. \quad (35)$$

3.4. Energy Pattern Factor Method (EPFM)

The EPFM for estimating the Weibull parameters as developed by [19] within the context of wind speed analysis is based on the energy pattern factor (EPF). The EPF is the ratio of the average wind power

density (WPD) to the cube of the mean wind speed. The WPD is obtained as follows.

Wind speed as air in motion possesses kinetic energy. For a given air mass m the kinetic energy present in air movement with velocity x is given as

$$K.E = \frac{1}{2} mx^2. \quad (36)$$

To obtain the mass m of flowing air passing through an area A perpendicular to its velocity, its volume after time t has elapsed is multiplied by the air density ρ (whose value in this paper is taken to be $1.225\text{kg}/\text{m}^3$) this gives the value of m as

$$m = A\rho xt. \quad (37)$$

Putting (37) into (36) gives the total wind energy as

$$E_x = \frac{1}{2} A\rho x^3 t. \quad (38)$$

To obtain the total wind power, we consider that power is the rate of change of energy with time. Thus we differentiate (38) with respect to t to obtain the wind power

$$P_x = \frac{1}{2} A\rho x^3. \quad (39)$$

Observe that the wind power is proportional to the cube of the wind velocity x . The actual average wind power based on a sample of wind speeds observations is obtained from (39) by defining x^3 as the mean value of the third power of the wind speeds observations. It follows that the average WPD associated with the PDF $f(x)$ of the wind speed random variable X is given as

$$WPD = \int_{-\infty}^{\infty} P_x f(x) dx. \quad (40)$$

Observe also that the average WPD given by (40) is obtained by computing the third moment about zero of the wind speed random variable X . If X follows the Weibull distribution, one readily obtain the WPD using (24) as

$$WPD = 0.6125c^3\Gamma\left(1 + \frac{3}{k}\right). \quad (41)$$

The cube of the mean wind speed for a wind speed random variable following the Weibull distribution is given by

$$(\mu'_1)^3 = c^3 \Gamma^3\left(1 + \frac{1}{k}\right). \quad (42)$$

It follows that the EPF for a wind speed random variable X following the Weibull is given by

$$EPF = \frac{0.6125c^3\Gamma\left(1 + \frac{3}{k}\right)}{c^3 \Gamma^3\left(1 + \frac{1}{k}\right)} = \frac{0.6125\Gamma\left(1 + \frac{3}{k}\right)}{\Gamma^3\left(1 + \frac{1}{k}\right)}. \quad (43)$$

Equating the population moments which make up (43) to their corresponding sample moments we obtain

$$\frac{\Gamma\left(1 + \frac{3}{k}\right)}{\Gamma^3\left(1 + \frac{1}{k}\right)} = \frac{\frac{1}{n}\sum_{i=1}^n X_i^3}{\bar{X}^3},$$

which is further expressed as

$$\frac{\Gamma\left(1 + \frac{3}{k}\right)}{\Gamma^3\left(1 + \frac{1}{k}\right)} - \frac{\frac{1}{n}\sum_{i=1}^n X_i^3}{\bar{X}^3} = 0. \quad (44)$$

The estimate of k is numerically obtained by finding the root of the equation given by (44). Once the estimate of k is obtained, the estimate of c can be obtained using (35).

3.5.Method of L-moments (MOL)

The moments based methods for estimating the Weibull parameters discussed in sections 3.3 and 3.4 are not always satisfactory. The sample moments, especially when the sample size is small can significantly be different from those of the probability distribution from which the sample was drawn; and the estimated parameters of distributions fitted by the method of moments are often markedly less accurate than those obtainable by other estimation procedures such as the maximum likelihood method. An alternative approach to the conventional moments based methods is the method based on L-moments as proposed by [28]. L-moments though analogous to the conventional moments are estimated by linear combinations of order statistics, i.e. by L-statistics. L-moments have the theoretical advantages over conventional moments of being able to

characterize a wider range of distributions and, when estimated from a sample, of being more robust to the presence of outliers in the data.

Let X be a real-valued random variable with CDF $F(x)$ and quantile function $x(F)$, and let $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ be the order statistics of a random sample of size n drawn from the distribution of X . The L-moments of X as defined by [28] are the quantities

$$\lambda_r \equiv r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} EX_{r-k:r} \quad r = 1, 2, \dots \quad (45)$$

The natural estimator of λ_r based on an observed sample of data is a linear combination of the ordered data values, i.e. an L-statistic. The expectation of an ordered statistic as defined in [29] may be written as

$$EX_{j:r} = \frac{r!}{(j-1)!(r-1)!} \int x\{F(x)\}^{j-1}\{1-F(x)\}^{r-j} dF(x). \quad (46)$$

Substituting (46) into (45), expanding the binomials in $F(x)$ and summing the coefficients of each power of $F(x)$ gives

$$\lambda_r = \int_0^1 x(F)P_{r-1}^*(F)dF, \quad r = 1, 2, \dots, \quad (47)$$

where

$$P_{r-1}^*(F) = \sum_{k=0}^r p_{r,k}^* F^k,$$

and

$$p_{r,k}^* = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k}.$$

$P_{r-1}^*(F)$ is the r^{th} shifted Legendre polynomial, related to the usual Legendre polynomial $P_r(u)$ by $P_r^*(u) = P_r(2u-1)$. Shifted Legendre polynomials are orthogonal on the interval (0,1) with constant weight function [30]. The first four L-moments are

$$\begin{aligned} \lambda_1 &= EX = \int_0^1 x(F)dF, \\ \lambda_2 &= \frac{1}{2}E(X_{2:2} - X_{1:2}) = \int_0^1 x(F)(2F-1)dF, \end{aligned} \quad (48)$$

$$\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3}) = \int_0^1 x(F)(6F^2 - 6F + 1)dF,$$

$$\lambda_4 = \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) = \int_0^1 x(F)(20F^3 - 30F^2 + 12F - 1)dF.$$

If X follows the Weibull distribution, then the first two L-moments of X using the system in (48) are given by

$$\lambda_1 = c\Gamma\left(1 + \frac{1}{k}\right), \tag{49}$$

$$\lambda_2 = c\Gamma\left(1 + \frac{1}{k}\right)\left(1 - \frac{1}{2^{1/k}}\right). \tag{50}$$

The L-moments describe in (48) above are for a probability distribution. In practice, they are often estimated from a finite sample. Suppose $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ be an ordered sample, the sample L-moments can be characterized better by the estimator of the probability weighted moments β_r . Following [31], an unbiased estimator of β_r based on the ordered sample is defined as

$$b_r = n^{-1} \binom{n-1}{r} \sum_{j=r+1}^n \binom{j-1}{r} X_{j:n}. \tag{51}$$

The sample L-moments ℓ_r are related to b_r by

$$\ell_{r+1} = \sum_{k=0}^r p_{r,k} b_k, \quad r = 0, 1, 2, \dots, n-1, \tag{52}$$

where

$$p_{r,k} = (-1)^{r-k} \binom{r}{k} \binom{r+1}{k}.$$

The first two sample L-moments are

$$\ell_1 = b_0, \tag{53}$$

$$\ell_2 = 2b_1 - b_0. \tag{54}$$

To estimate the Weibull parameters, we equate the population L-moments given in (49) and (50) to their corresponding sample estimates given in (53) and (54). The following results follow.

$$c\Gamma\left(1 + \frac{1}{k}\right) = \ell_1,$$

implies that

$$c = \frac{\ell_1}{\Gamma\left(1 + \frac{1}{k}\right)}. \quad (55)$$

$$c\Gamma\left(1 + \frac{1}{k}\right)\left(1 - \frac{1}{2^{1/k}}\right) = \ell_2. \quad (56)$$

Substituting (55) into (56) and evaluating gives the estimate \hat{k} of k as

$$\hat{k} = \frac{-\log 2}{\log\left(\frac{\ell_1 - \ell_2}{\ell_1}\right)}. \quad (57)$$

Once \hat{k} is determined from (57), the estimate \hat{c} of c is then estimated from (55) by

$$\hat{c} = \frac{\ell_1}{\Gamma\left(1 + \frac{1}{\hat{k}}\right)}. \quad (58)$$

3.6. Maximum Likelihood Method (MLM)

Given the PDF $f(x)$ of the Weibull distribution with parameters k and c , and a random independent sample of observations x_1, x_2, \dots, x_n of size n , the MLM estimates of the parameters k and c are obtained by maximizing the log-likelihood function

$$L = \sum_{i=1}^n \ln(f(x_i)). \quad (59)$$

For the Weibull distribution, the log-likelihood function is given as

$$L = n \ln k + (k - 1) \sum_{i=1}^n \ln x_i - nk \ln c - \sum_{i=1}^n \left(\frac{x_i}{c}\right)^k. \quad (60)$$

To obtain the estimate of k and c we take the partial derivative of (60) w.r.t. k and c and equate the resulting partial derivatives to zero. This gives

$$\frac{\partial L}{\partial k} = \frac{n}{k} + \sum_{i=1}^n \ln x_i - n \ln c - \sum_{i=1}^n \left(\frac{x_i}{c}\right)^k \ln\left(\frac{x_i}{c}\right) = 0, \quad (61)$$

$$\frac{\partial L}{\partial c} = \frac{\beta}{c} \sum_{i=1}^n \left(\frac{x_i}{c}\right)^\beta - \frac{nk}{c} = 0. \quad (62)$$

Observe that (61) and (62) are systems of non-linear equations in k and c . From (62) we obtain

$$c = \left(\frac{1}{n} \sum_{i=1}^n x_i^k \right)^{\frac{1}{k}}. \quad (63)$$

Substituting c in (63) into (61) and evaluating gives

$$\frac{1}{k} + \frac{1}{n} \sum_{i=1}^n \ln x_i - \frac{\sum_{i=1}^n x_i^k \ln x_i}{\sum_{i=1}^n x_i^k} = 0. \quad (64)$$

The system of equation in (64) is non-linear in k , the estimate \hat{k} of k can be numerically determined by finding the root of the equation. Once \hat{k} is obtained the estimate \hat{c} of c is then calculated from (63) as

$$\hat{c} = \left(\frac{1}{n} \sum_{i=1}^n x_i^{\hat{k}} \right)^{\frac{1}{\hat{k}}}. \quad (65)$$

4. Simulation Study

For the simulation study a random sample is generated from the Weibull distribution with $k = 2.5$ and $c = 1$ for sample sizes $n = 5, 20, 40, 80, 100, 150$. For each sample size $B = 5000$ bootstrap samples are generated. For each methods discussed in section 3, 5000 estimates $\hat{k}_1, \hat{k}_2, \dots, \hat{k}_{5000}$ of the parameter k and 5000 estimates $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_{5000}$ of the parameter c are obtained. The means (\bar{k} and \bar{c}) of these estimates for k and c are computed. To compare the performance of the methods of estimation, the mean square errors is computed from the relation

$$MSE = \frac{1}{n} \sum_{i=1}^n [\hat{F}(x_i) - F(x_i)]^2, \quad (66)$$

where $\hat{F}(x_i)$ is obtained by substituting \bar{k} and \bar{c} (for each method) in (2) while $F(x_i) = i/n$ is the empirical CDF (ECDF). The method with the smallest mean square error value is taken as best among the others. Results from the simulation study are contained in Table 1.

5. Discussion of Results from Simulation Study

Results from the simulation study as presented in Table 1 and Figure 1, clearly reveal that the MSE decreases as the sample size increases for the 6 methods considered. For $n = 5$, the MSE for the MOM is smallest followed by that of the MOL. The EPFM reported the highest MSE followed by that of WLSM, LSM and MLM respectively. For sample size $n = 20$, the WLSM reported the lowest MSE follow by that of MLM, MOL, MOM, LSM and EPFM. For sample size $n = 40$ the MLM reported the smallest MSE. Sample sizes $n = 80$ and $n = 100$ report the MOL as the method with the smallest MSE. The MLM reported the smallest MSE for sample size $n = 150$. It is evident from the study that the EPFM reported the highest MSE for all sample sizes. Again, for small sample sizes the WLSM appears to be very efficient as well as in large sample sizes. The MOL and MLM can also be viewed as good methods for small, moderate and large sample sizes.

6. Application to Wind Speed Sample

For the application, the Weibull distribution, using the 6 methods discussed in section 3 was used to fit the daily mean wind speeds sample covering 3 years (2012-2015) obtained from the National Center for Energy and Environment (NCEE) University of Benin, Benin City. A total sample size of 988 observations was used for the analysis out of a possible sample size of 1096. The missing observations were due to temporary shut down and maintenance of recording systems at the Center. Results from the analysis are contained in Table 2 which also contains the mean square errors obtained using the 6 methods and computed from the relation in (66).

7. Discussion of Results from Application

Results obtained from using the 6 methods to fit the Weibull distribution to the wind speeds data as contained in Table 2 and Figure 2 suggests that the Weibull distribution provides a very good model for the wind speeds data. This is evident from the fact that the MSEs obtained using the 6 Weibull methods are small and the fitted density

in Figure 2 also testify to that. The MLM provided the best method for estimating the Weibull parameters for the wind speeds sample due to it having the smallest MSE. The WLSM and MOL also proved to be good methods as well. This result further confirmed the one obtained from the simulation study on the efficacy of the MLM, WLSM and MOL in estimating the Weibull parameters.

Table 1: Simulation results of the parameters estimation $k = 2.5, c = 1$

Sample size	Method	\bar{k}	\bar{c}	MSE
$n = 5$	LSM	2.9192	1.0936	0.0151
	WLSM	3.1554	1.0947	0.0152
	MOM	4.1418	1.0640	0.0105
	EPFM	1.9962	1.0904	0.0194
	MOL	3.7510	1.0700	0.0110
	MLM	4.9579	1.0570	0.0113
$n = 20$	LSM	2.9511	0.9412	0.0038
	WLSM	2.7657	0.9128	0.0018
	MOM	2.8881	0.9406	0.0037
	EPFM	1.6526	0.9380	0.0120
	MOL	2.9679	0.9395	0.0037
	MLM	2.8428	0.9413	0.0036
$n = 40$	LSM	2.0360	0.9736	0.0014
	WLSM	2.2449	0.9737	9.1×10^{-4}
	MOM	2.3783	0.9555	6.7×10^{-4}
	EPFM	1.5444	0.9413	0.0060
	MOL	2.3233	0.9558	6.5×10^{-4}
	MLM	2.5099	1.0196	6.4×10^{-4}
$n = 80$	LSM	2.6186	1.0130	2.7×10^{-4}
	WLSM	2.6412	1.0178	2.8×10^{-4}
	MOM	2.7517	1.0077	2.4×10^{-4}
	EPFM	1.6427	1.0025	0.0055
	MOL	2.7209	1.0082	2.3×10^{-4}
	MLM	2.7977	1.0081	2.6×10^{-4}
$n = 100$	LSM	2.5911	1.0393	2.3×10^{-4}
	WLSM	2.6357	1.0298	2.4×10^{-4}
	MOM	2.6484	1.0362	2.2×10^{-4}
	EPFM	1.6043	1.0274	0.0053
	MOL	2.6465	1.0363	2.1×10^{-4}
	MLM	2.6651	1.0367	2.5×10^{-4}

$n = 150$	<i>LSM</i>	2.2777	1.0735	1.9×10^{-4}
	<i>WLSM</i>	2.6216	1.0480	1.7×10^{-4}
	<i>MOM</i>	2.5298	1.0575	1.4×10^{-4}
	<i>EPFM</i>	1.5690	1.0449	0.0041
	<i>MOL</i>	2.5492	1.0573	1.2×10^{-4}
	<i>MLM</i>	2.5156	1.0558	1.1×10^{-4}

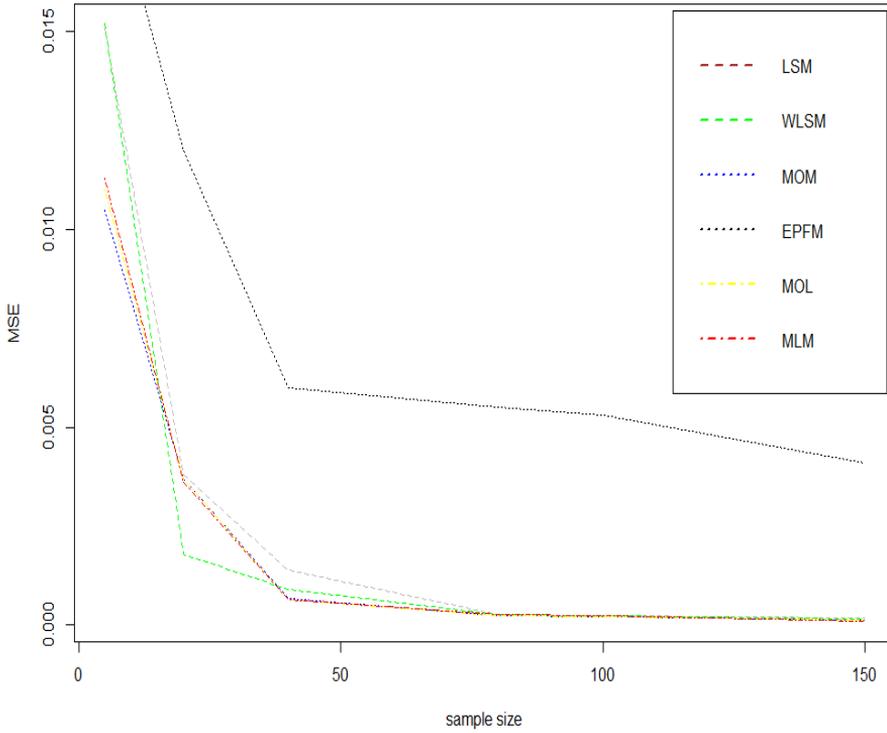


Figure 1: Mean square errors from Weibull methods

Table 2: Results from fitting the Weibull distribution to wind speed data

Sample size	Meth od	\hat{k}	\hat{c}	<i>MSE</i>
-------------	---------	-----------	-----------	------------

$n = 988$	<i>LSM</i>	2.3498	1.3773	1.1×10^{-4}
	<i>WLSM</i>	2.2583	1.3857	7.7×10^{-5}
	<i>MOM</i>	2.2101	1.3820	1.0×10^{-4}
	<i>EPFM</i>	1.4071	1.3439	0.0079
	<i>MOL</i>	2.2980	1.3816	8.1×10^{-5}
	<i>MLM</i>	2.1793	1.3803	1.3×10^{-5}

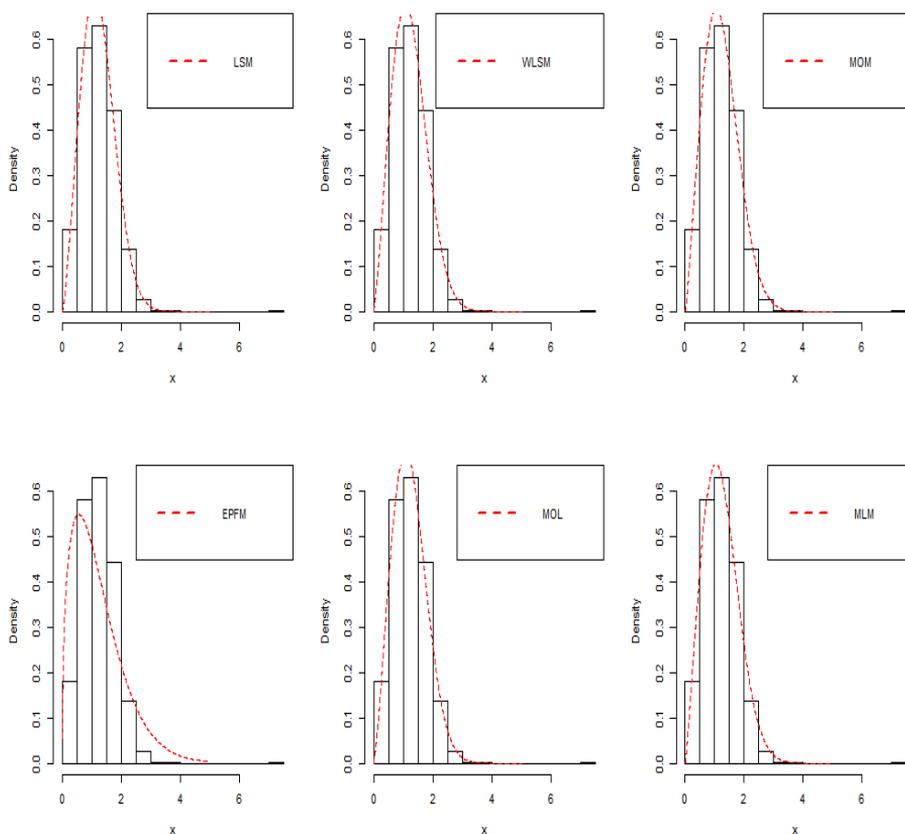


Figure 2: Fitted Weibull density of wind speeds sample

8. Conclusion

In this paper, we have discussed six methods for estimating the Weibull shape and scale parameters using simulation study and application to a real data set (wind speeds sample). We recommend based on the results obtained in this paper that the MLM be used when dealing with large sample sizes. The WLSM, MOM and MOL should also be considered when fitting small and moderate sizes of samples.

Acknowledgement

The authors are very grateful to the anonymous referee(s) and the editor-in-chief for their useful comments which largely helped to improve this paper.

References

1. Akdag SA, Ali D (2009). A new method to estimate Weibull parameters for wind energy applications. *Energy Conversion and Management*.**50**, 1761-1766.
2. Azad, A.K., Rasul, M.G., and Yusaf, Talal. (2014). Statistical diagnosis of the best Weibull methods for wind power assessment for agricultural applications, *Energies*, **7**, 3056-3085; doi: 10.3390/en7053056.
3. Bergman, B. (1986). Estimation of Weibull Parameters Using a Weight Function, *Journal of Materials Science Letters*, **5**, 611-614.
4. Celik, A.N. (2004). A statistical analysis of wind power density based on the Weibull and Rayleigh models at the Southern region of Turkey. *Renewable Energy*, **29**, 593-604.
5. Chu, Y. K., and Ke, Ch. J. (2012). Computation approaches for parameter estimation of Weibull distribution, *Mathematical and Computational Applications*, **17**(1); 39-47.
6. David, H.A. (1981). *Order Statistics*, 2nd edition. New York: Wiley.
7. Goda, Y., Kudaka, M., and Kawai, H. (2010). Incorporation of Weibull distribution in L-moments method for regional frequency analysis of peaks-over-threshold wave heights. In "Proceedings of the International Conference of Coastal Engineering.". Available at [http://journals.tdl.org/ICCE/article/view file/1154/pdf_42](http://journals.tdl.org/ICCE/article/view/file/1154/pdf_42).

8. Gupta, R., and Biswa, A. (2010). Wind data analysis of Silchar (Assam, India) by Rayleigh and Weibull methods. *Journal of Mechanical Engineering Research*, **2**, 10-24.
9. Hosking, J.R.M. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics, *Journal of Statistical Royal Society*, B, **52**(1); 105-124.
10. Hosking, J.R.M., and Wallis, J. R. (1997), *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press, Cambridge, UK.
11. Johnson, N.L., Kotz, S., and Balakrishnan, N. (1995). *Continuous univariate distributions*, vol. 2, second edition, New York, John Wiley and sons, Inc.
12. Kececioglu, D. (1991) *Reliability Engineering Handbook*. Prentice-hall, Englewood Cliffs, New Jersey.
13. Lanczos, C. (1957). *Applied Analysis*, p.286. London: Pitman.
14. Lawless, J.F. (1978) Confidence interval estimation for the Weibull and extreme value distributions. With a discussion by Easterling, R.G. *Technometrics*, **20**, 355-365.
15. Lawless, J.F. (1982) *Statistical Models and Methods for Lifetime Data*. 2nd edition, John Wiley & Sons, New York.
16. Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*. 3rd edition, John Wiley & Sons, New York.
17. Mann, N.R. (1968) Point and interval estimation procedures for the two-parameter Weibull and extreme-value distributions. *Technometrics*, **10**, 231-256.
18. Nelson, W. (1982) *Applied Life Data Analysis*. John Wiley & Sons, New York.
19. O'Connor, P.D.T. (2002) *Practical Reliability Engineering*. 4th edition, Chichester, Wiley.
20. Odo, F.C., Offiah, S.U., and Ugwuoke, P.E. (2012). Weibull distribution-based model for prediction of wind Potential in Enugu, Nigeria. *Advances in Applied Science Research*, **3**(2); 1202-1208.
21. Osatohanmwen, P., Oyegue, F.O., and Ogbonmwan, S.M. (2016). Statistical analysis of wind energy potential in Benin City using the 2-parameter Weibull distribution. *International Journal for Renewable Energy and Environment*, **2**, 22-31.
22. Petersen, T.L., Troen, I., Frandsen, S., and Hedegaard, K. (1981). Wind Atlas for Denmark, *RISO*, Denmark.
23. Pobočiková, I., and Sedliačková, Z. (2014). Comparison of four methods for estimating the Weibull distribution parameters, *Applied Mathematical Sciences*, **8**, 41-4149.

24. Sarkar A., Singh, S. and Mitra, D. (2011). Wind Climate modeling using Weibull and Extreme value distribution. *International Journal of Engineering Science and Technology*,**3**(5); 100-106.
25. Sarkar, A.,and Kasperki, M. (2009). Weibull parameters for Wind Speed distribution in India. *Proceedings of 5th National conference on Wind Engineering*, 134-158.
26. Soland, R.M. (1969) Bayesian analysis of the Weibull process with unknown scale and shape parameters. *IEEE Transactions on Reliability*, **18**, 181-184.
27. Trustrum, K., and Jayatilaka, A. S. (1979). On estimating the Weibull modulus for a brittle material, *Journal of Material Science*, **14**, 1080–1084.
28. Ulgen, K., and Hepbasli, A. (2002). Determination of Weibull parameters for wind Energy analysis of Izmir, Turkey. *International Journal of Energy Research*,**26**, 495-506.
29. Wentink, T. (1976). Study of Alaskan Wind power potential and its possible application. *Final Rep., Rep No. NSF/RANN/SE/AER 74-0023/FR – 76/1*, Geophysical Institute, University of Alaska.
30. Wu, D., Zhou, J., and Li, Y. (2006). Methods for estimating Weibull parameters for brittle materials, *Journal of Material Science*, **41**, 5630-5638.
31. Zaharim, A., Razali, A.M., Abidin, R.Z., and Sopian, K. (2009). Fitting of statistical distributions to wind speed data in Malaysia. *European Journal of Scientific Research*, **26**(1); pp.6-12.