

On The Estimation of Parameters in a Weibull Wind Model and its Application to Wind Speed Data from Maiduguri, Borno State, Nigeria

Gongsin Isaac Esbond[1*], Fumilayo W. O. Saporu[2]
1 Department of Mathematics and Statistics, University of Maiduguri, Maiduguri, Borno State, Nigeria
2 National Mathematical Centre, Kwali, FCT-Abuja, Nigeria.

Abstract

Four methods of parameter estimation of the Weibull distribution are examined. These are maximum likelihood, method of moments, optimization and regression methods. It is shown how the parameters of the distribution can be obtained by each of these four methods using iterative techniques in numerical methods. These are illustrated by fitting Weibull and Rayleigh models to the wind speed data from Maiduguri. The model fits obtained by using each of these four methods of estimation are tested using four goodness–of –fit tests and compared using Root-Mean-Square-Error estimates. Results show that (i) only the Weibull model fits the data, (ii) differences in the corresponding estimates of parameters so obtained are thin, (iii) differences in the root-mean-square-error (RMSE) estimates are thin (iv) the RMSE estimates for the regression method is consistently the smallest; that is, the best fit and (v) the regression method is also the easiest to implement in obtaining estimates of the Weibull parameters and their standard errors. These indicate that the regression method could be a user's first choice in obtaining parameter estimates for a Weibull wind model.

Keywords: Parameter Estimation, Goodness-of-Fit tests, Root-Mean-Square-Error, Weibull distribution, Wind speed data, Wegstein's iterative method, optimization method, R programming language, Renewable Energy, New-Raphson iterative solution.

1 Introduction

Global attention is slowly drifting away from non-renewable energy like fossil fuels, oils and natural gases because of their by-product of environmental pollution that has adverse consequences on human health and climate. These hazardous consequences and the lessons of the oil crisis of 1973 have prompted the need to develop alternative and renewable energy sources. Hence today, there is a terrific global surge of research activities in this direction. Renewable energy sources are wind, solar, geothermal, hydro, biomass and ocean thermal energy. Their inexhaustible, environmentally friendly and economically viable characteristics make them more attractive for adoption in many countries (Proma et al, 2014) as an alternative energy source.

Wind energy is an inexpensive source of electric power generation. Hence research (Zhou et al, 2006) into wind power potential is of primary interest in many countries. Indices relevant to wind energy assessments are based on models for wind data. The primary interest of this paper therefore is to compare four methods of parameter estimation of the Weibull wind model, illustrating the same with wind speed data from Maiduguri, Nigeria. This is primarily done so that potential users could have an informed choice depending on their individual circumstance.

2 The Data

The data for illustration were obtained from the Nigerian Meteorological Agency (NIMET) office in Maiduguri. It is the wind speed data covering the period from September 1985 to December 2011, at hub height of 10 meters. The data is given in Table 2.1 below.

Table 2.1 Maiduguri Wind Speed Data (in m/s) from September 1985 to December 2011

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Year												
1985									3.06	3.98	3.66	4.02
1986	3.48	3.93	4.54	4.93	4.51	6.22	5.27	4.11	3.46	3.00	3.26	4.04
1987	3.50	3.85	3.95	5.53	5.44	4.87	5.39	3.71	3.62	2.23	3.57	2.43
1988	3.77	4.48	4.89	4.37	4.47	5.78	5.59	3.98	3.53	3.14	3.25	3.64
1989	4.52	4.94	4.32	4.15	4.94	5.55	5.22	3.89	3.30	3.33	3.39	3.29
1990	3.90	4.52	5.53	4.29	5.07	5.38	4.51	4.06	3.30	3.33	3.28	3.46
1991	3.74	3.60	4.46	4.64	4.29	4.79	4.18	3.17	3.24	2.97	3.42	3.37
1992	3.92	4.70	4.16	4.41	4.54	5.23	4.40	3.69	2.57	2.42	3.16	2.70
1993	3.77	1.74	4.02	3.86	4.77	4.53	4.00	3.25	2.85	2.75	2.91	3.35
1994	3.46	3.82	4.63	3.83	3.89	4.96	4.31	3.26	2.36	2.69	3.00	2.78
1995	2.85	3.24	4.08	4.22	4.15	5.01	4.05	2.55	2.45	1.64	2.49	2.78
1996	2.26	2.73	3.47	3.36	3.93	3.63	3.39	2.61	2.42	2.14	2.30	1.93
1997	2.82	4.18	3.72	3.45	3.49	3.49	3.48	2.27	2.12	2.09	2.16	2.25
1998	3.32	3.80	2.87	3.35	3.69	6.24	3.17	4.31	4.92	1.00	1.67	2.15
1999	2.32	2.69	2.94	3.71	3.53	3.58	3.28	2.46	2.11	1.91	2.08	2.19
2000	2.72	3.46	3.35	2.85	3.48	4.14	3.44	2.79	3.19	2.22	2.80	2.80
2001	2.86	3.53	3.68	3.84	4.30	4.21	3.91	2.96	2.57	2.16	5.10	1.81
2002	2.81	1.88	2.67	2.92	3.11	3.97	3.78	2.88	2.46	2.05	2.05	2.28
2003	1.56	1.44	2.16	2.65	2.75	2.26	1.51	0.90	1.59	2.18	2.40	2.38
2004	2.23	3.18	3.51	2.90	3.87	3.61	2.99	2.70	2.32	1.81	1.72	1.61
2005	2.42	2.51	2.97	2.94	2.92	3.09	2.63	1.82	1.64	1.51	1.59	1.08
2006	1.48	2.39	2.87	3.25	3.50	4.13	6.24	2.86	2.70	2.53	1.86	2.17
2007	3.12	2.91	3.29	3.59	3.61	3.91	3.24	2.19	1.91	2.13	1.86	2.17
2008	2.56	2.60	2.10	2.59	2.67	2.75	2.52	1.93	1.72	1.82	1.54	1.51
2009	1.43	1.85	2.12	2.19	1.80	2.55	2.45	1.75	1.26	1.28	1.05	1.31
2010	0.90	1.26	1.65	1.74	1.70	1.98	1.64	1.21	1.11	0.90	0.84	0.85
2011	1.44	1.96	1.68	2.48	1.96	1.68	1.30	3.40	1.06	1.06	1.07	1.13

3 Weibull Wind Model

Wind is the response of the atmosphere, arising from the uneven heating condition of the earth by the sun, which produces a pressure difference in the atmosphere, triggering wind to blow from high to low regions (Proma et al, 2014). The motion energy of wind can be harvested by modern wind turbines to generate electricity. Wind speed and its duration are the key factors used in designing and determining the use of wind energy. Wind power developers therefore measure actual wind resources, to determine the distribution of wind speeds because it is a key factor in the design of wind turbines.

Attempts (Ahmed and Mohammed, 2012, and Odo et al, 2012) have been made in modeling wind speed distribution. Of the various probability density functions for wind speed, the 2-parameter Weibull distribution is the most commonly used for wind energy studies. This is because it has flexible range of values within each of its parameters for which suitable choice can be found for most situations. Its density function is given by

$$f_X(x;\alpha,\beta) = \begin{cases} \beta\alpha^{-\beta}x^{\beta-1}e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, & x > 0,\ \alpha > 0 \ and\ \beta > 0 \\ 0 & otherwise \end{cases} \quad 3.1$$

The Rayleigh distribution is a special form of the Weibull distribution and its density function is given by

$$f_X(x;\alpha) = \begin{cases} 2\alpha^{-2}xe^{-\left(\frac{x}{\alpha}\right)^{2}}, & x > 0 \ and\ \alpha > 0 \\ 0 & otherwise \end{cases} \quad 3.2$$

4 Estimation

4.1 Review

Weibull distribution is a 2-parameter probability density function. Estimation of its parameters can be done largely by numerical methods (Johnson and Kotz, 1970). Of all the various methods of estimation of its parameters available, maximum likelihood is the most popular; this not surprising because by virtue of its properties, maximum likelihood estimators, in large samples, tend to be efficient, consistent and asymptotically normally distributed (Mood et al, 1963).

Fritz (2008) discussed how maximum likelihood estimates for these parameters can be obtained from the log-transform of the Weibull random variable and their confidence bounds from the idea of pivot and simulation. On the other hand, LDABOOK [8] shows how maximum likelihood estimates can be obtained directly without transformation and their standard errors obtained from the inversion of the Fisher's information matrix.

Regression methods for estimating parameters are also discussed in Fritz (2008) and LDABOOK [8]. Fritz's discussion again uses the log-transformation of the Weibull variate and constructs confidence bounds for the regression line based on simulation. LDABOOK [8] obtained least squares estimates for the parameters from the median rank values of failure times and their confidence bounds from the inversion of the Fisher information matrix.

There are several attempts in applying Weibull wind model to data. In all these, two methods of estimation are popular; maximum likelihood method (Shamshad et al, 2012; Ahmed and Mohammed, 2012) and regression method (Kostas and Despina, 2014 and Odo et al, 2012). Other methods used are in Nikolai et al (2014) and Dikko and Yahaya (2012). A list of some of the available methods can be found in Proma et al (2014). In all these, parameter estimates were produced without standard errors, and goodness-of-fit tests of the model to the data were not performed. This is not surprising. The complexity of the numerical methods involved in the computation of estimates derived by maximum likelihood method may perhaps have led to the choice of less efficient hand estimates.

4.2 The Problem

The use of four methods of estimation and the construction of their standard errors is the focus here. These are maximum likelihood, method of moments, optimization and regression methods. The inclusion of the method of moments is new. Maximum likelihood estimates are usually derived from the solution of the partial derivatives of the likelihood function with respect to the parameters set equal to zero. Here, the use of the optimization of the likelihood function to derive such solution is again new.

4.3 Maximum Likelihood Estimation

The log-likelihood function is given by

$$l(\alpha, \beta / x_1, x_2, \ldots, x_n) = -nlog\beta + n\beta log\alpha - (\beta - 1)\sum_{i=1}^{n} logx_i + \alpha^{-\beta}\sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta} \quad 4.1$$

and its first partial derivatives set equal to zero are

$$\frac{\partial l(\alpha, \beta / x_1, x_2, \ldots, x_n)}{\partial \alpha} = -\frac{n\beta}{\alpha} + \beta\alpha^{-(\beta+1)}\sum_{i=1}^{n} x_i^{\beta} = 0 \quad 4.2$$

$$\frac{\partial l(\alpha, \beta / x_1, x_2, \ldots, x_n)}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n} logx_i - nlog\alpha + \alpha^{-\beta}log\alpha\sum_{i=1}^{n} x_i^{\beta} - \alpha^{-\beta}\sum_{i=1}^{n} x_i^{\beta}\, logx_i = 0 \quad 4.3$$

Hence

$$\hat{\alpha} = \left[\frac{\sum_{i=1}^{n} x_i^{\beta}}{n}\right]^{1/\beta} \quad 4.4$$

$$\hat{\beta} = \left[\frac{\sum_{i=1}^{n} x_i^{\beta} logx_i}{\sum_{i=1}^{n} x_i^{\beta}} - \frac{\sum_{i=1}^{n} logx_i}{n}\right]^{-1}, \quad 4.5$$

Equations 4.4 and 4.5 are implicit equations. Consequently, their solutions which provide maximum likelihood estimates can only be obtained using, for example, Wegstein's iterative method (Salvadori and Baron, 1961and Contantinides, 1987) given by

$$\beta_{n+1} = \frac{\beta_{n-1} * g(\beta_n) - \beta_n * g(\beta_{n-1})}{\beta_{n-1} - g(\beta_{n-1}) - \beta_n + g(\beta_n)}, n \geq 2, \quad 4.6$$

where $g(\beta_n)$ is obtained by substituting the value of β_n in the right hand side of equation 4.5. Of course, β_1 is the initial value. The value of β so obtained from the iteration in 4.6 is substituted in 4.4 to estimate α.

The estimates of the standard errors of the parameters are obtained from the inversion of the Fisher's information matrix, given by

$$I(\alpha, \beta / x_1, x_2, \ldots, x_n) = -\begin{pmatrix} \frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \alpha \partial \beta} \\ \frac{\partial^2 l}{\partial \beta \partial \alpha} & \frac{\partial^2 l}{\partial \beta^2} \end{pmatrix} \quad 4.7$$

where l is the log-likelihood function given in equation 4.1.

4.4 Optimization Method

Here the objective function is minus the log-likelihood function of equation 4.1 and estimates of parameters can be obtained, for example, using the Newton-Raphson iterative solution method (Salvadori and Baron, 1961and Contantinides, 1987) given by

$$\hat{B}^{i+1} = \hat{B}^{i} + \rho J^{-1} \times -\boldsymbol{f} \quad i = 0, 1, 2, 3, \ldots \quad 4.8$$

where $B^i = \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix}$, α_o and β_o are the initial values of α and β,

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial \alpha} & \frac{\partial f_1}{\partial \beta} \\ \frac{\partial f_2}{\partial \alpha} & \frac{\partial f_2}{\partial \beta} \end{pmatrix}, \text{ and } \boldsymbol{f} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix},$$

where f_1 and f_2 are the first partial derivatives of minus log-likelihood with respect to the parameters α and β, respectively. Values for $\boldsymbol{f}$ and J are obtained using the i^{th} iteration values of the parameters α and β. The constant ρ is called the relaxation factor.

4.5 Method of Moments

Method of moments estimators are obtained by equating sample moments against population moments. Estimators (Mood et al, 1963) for Weibull parameters are given by

$$\hat{\alpha} = \frac{\hat{\beta}\bar{x}}{\Gamma\left(\frac{1}{\hat{\beta}}\right)} \quad 4.9$$

and

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i^2 \left[\Gamma\left(\frac{1}{\hat{\beta}}\right)\right]^2}{2n\bar{x}^2\Gamma\left(\frac{2}{\hat{\beta}}\right)} \quad 4.10$$

The standard errors of the estimates are estimated from the variances given by

$$var(\hat{\alpha}) = \frac{\hat{\alpha}^2}{n}\left[\Gamma\left(1+\frac{2}{\hat{\beta}}\right) - \left(\Gamma\left(1+\frac{1}{\hat{\beta}}\right)\right)^2\right]\left[\frac{\hat{\beta}}{\Gamma\left(\frac{1}{\hat{\beta}}\right)}\right]^2 \quad 4.11$$

and

$$var(\hat{\beta}) = \left\{\frac{\left[\Gamma\left(\frac{1}{\hat{\beta}}\right)\right]^2}{2\bar{x}^2\Gamma\left(\frac{2}{\hat{\beta}}\right)}\right\}^2 \frac{var(x_i^2)}{n} \quad 4.12$$

Again equations 4.9 and 4.10 are implicit equations hence method of moment estimates of parameters can also be obtained using Wegstein's iterative method as described above. The standard errors of the estimates can then be obtained by direct substitution of estimates of α and β into equations 4.11 and 4.12 before taking the square root.

4.6 Regression Method

The Weibull cumulative density function is given by

$$F(x) = 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}} \quad 4.13$$

Taking the natural logarithm of both sides of equation 4.13 twice, we have

$$log\left(-log\left(1 - F(x)\right)\right) = \beta log(x) - \beta log(\alpha) \quad 4.14$$

This is a straight line with gradient β and intercept $-\beta log(\alpha)$. Any statistical package can be used to produce β, the intercept $-\beta log(\alpha)$ and their respective standard errors. We derive the estimate for α from $-\beta log(\alpha) = intercept$ as

$$\hat{\alpha} = e^{-\frac{intercept}{\hat{\beta}}} \quad 4.15$$

and its standard error from $\hat{\beta}^2\left(log(\hat{\alpha})\right) = var(-\widehat{intercept})$ as

$$s.e(\hat{\alpha}) = e^{-\frac{se(\widehat{intercept})}{\hat{\beta}}} \quad 4.16$$

The standard error of $\hat{\alpha}$ computed this way is much easier than that suggested by Fritz (2008) and LDABOOK[8].

4.7 Computer Programs for Computation

Samples of the computer programs written in R statistical programming language that can be used in computing the parameter estimates and their standard errors from data are given below.

i. R Codes for Maximum Likelihood Estimation

```
# Program: WeibullMLE
# Estimating the Weibull Parameters using Maximum Likelihood Estimation
x <- c(data) # Wind speed data of various months
n <- length(x) # Sample wind data size
m <- mean(x)
```



```
d <- x – m # deviations from the mean
m2 <- sum(d^2)/n # second moment about the mean
m4 <- sum(d^4)/n # fourth moment about the mean
b <- m4/m2^2 # assumed initial value of the shape parameter β in terms of sample Kurtosis
y <- sum(log(x))/n
y0 <- sum(x^b)
y1 <- sum(x^b * log(x))
g1 <- (y1/y0 – y)^(-1)
b2 <- g1 # initial computed value of the shape parameter β
tolerance <- 0.00001
while (abs(b – b2) > tolerance) { # initializing the iterative process
y2 <- sum(x^b2)
y3 <- sum(x^b2 * log(x))
g2 <- (y3/y2 – y)^(-1)
d1 <- b * g2 – b2 * g1
d2 <- b – g1 – b2 + g2
b3 <- d1/d2 # Computing Wegstein variable of equation 3.18
b <- b3
y0 <- sum(x^b)
y1 <- sum(x^b * log(x))
g1 <- (y1/y0 – y)^(-1)
b2 <- g1
}
alpha.hat <- (sum(x^b)/n)^(1/b) # Computes the scale parameter α
beta.hat <- b
# Estimating the standard errors of the estimates
y2 <- sum(x^b * (log(x))^2)
# Computing the elements of the Hessian Matrix
a <- alpha.hat
d11 <- n * b/a^2 – b * (b+1) * a^(-b – 2) * y0
d12 <- -n/a + a^(-b – 1) * y0 – b * a^(-b – 1) * log(a) * y0 + b * a^(-b – 1) * y1
d21 <- d12
d22 <- -n/b^2 – a^(-b) * (log(a))^2  * y0 + 2 * a^(-b) * log(a) * y1 – a^(-b) * y2
H <- matrix(c(d11,d12,d21,d22), nrow = 2) # Hessian Matrix
I <- -H # Information matrix equals minus expectation of Hessian Matrix
VarCov <- solve(I) # Inverse of the information matrix gives the covariance matrix
valpha.hat <- VarCov[1,1] # variance of alpha.hat
vbeta.hat <- VarCov[2,2] # variance of beta.hat
SE <- sqrt(c(valpha.hat, vbeta.hat)) # standard errors of the parameter estimates
WeibullMLE <- c(alpha.hat, beta.hat) # ML Estimates of the Weibull parameters
WeibullMLE # Press Enter to return the estimated parameter values
SE # Press Enter to return the standard errors of the estimated parameters
```

ii. R Codes for Optimization Method

```
# Program: WeibullNM
# Estimating the Weibull Parameters using Numerical Method
x <- c(data) # Wind speed data of various months
n <- length(x) # Sample wind data size
m <- mean(x) # sample mean
d <- x – m # deviations from the mean
```

```
m2 <- sum(d^2)/n # second moment about the mean
m4 <- sum(d^4)/n # fourth moment about the mean
a <- m # initial value of the scale parameter α in terms of sample mean
b <- m4/m2^2 # initial value of the shape parameter β in terms of sample Kurtosis
B <- c(a, b)
y0 <- sum(x^b)
y1 <- sum(x^b * log(x))
f1 <- n * b/a - b * a^(-b – 1) * y0
f2 <- -n/b + n * log(a) – sum(log(x)) – a^(-b) * log(a) * y0 + a^(-b) * y1
F <- c(f1, f2)
for (i in 1:100000) {
y2 <- sum(x^b * (log(x))^2)
d11 <- -n * b/a^2 + b * (b+1) * a^(-b – 2) * y0
d12 <- n/a - a^(-b – 1) * y0 + b * a^(-b – 1) * log(a) * y0 – b * a^(-b – 1) * y1
d21 <- d12
d22 <- n/b^2 + a^(-b) * (log(a))^2 * y0 – 2 * a^(-b) * log(a) * y1 + a^(-b) *y2
H <- matrix(c(d11,d12,d21,d22), nrow = 2)
T <- solve(H, -F)
B <- B + 0.5 * T # 0.5 is the relaxation factor
a <- B[1]
b <- B[2]
y0 <- sum(x^b)
y1 <- sum(x^b * log(x))
f1 <- n * b/a - b * a^(-b – 1) * y0
f2 <- -n/b + n * log(a) – sum(log(x)) – a^(-b) * log(a) * y0 + a^(-b) * y1
F <- c(f1, f2)
}
WeibullNM <- B # Vector of parameter estimates
# Computing the variances and standard errors of the estimates
H <- -matrix(c(d11,d12,d21,d22), nrow = 2) # Hessian Matrix
I <- -H # Information matrix equals minus expectation of Hessian Matrix
VarCov <- solve(I) # Inverse of the information matrix gives the variance matrix
valpha.hat <- VarCov[1,1] # variance of alpha.hat
vbeta.hat <- VarCov[2,2] # variance of beta.hat
SE <- sqrt(c(valpha.hat, vbeta.hat)) # standard errors of the parameter estimates
WeibullNM <- c(alpha.hat, beta.hat) # Numerical Estimates of the Weibull parameters
WeibullNM # Press Enter to return the estimated parameter values
SE # Press Enter to return the standard errors of the estimated parameters
```

iii. R Codes for Method of Moments

```
# Program: WeibullMME
# Estimating the Weibull Parameters using Method of Moments Estimation
x <- c(data)
n <- length(x) # Sample Size
m <- mean(x)
d <- x – m # deviations from the mean
m2 <- sum(d^2)/n # second moment about the mean
m4 <- sum(d^4)/n # fourth moment about the mean
b <- m4/m2^2 # assumed initial value of the shape parameter β in terms of sample Kurtosis
y1 <- m
```

```
y2 <- sum(x^2)/n # Computing the second moment about the origin
f1 <- 2 * (y1^2) * gamma(2/b)
f2 <- y2 * (gamma(1/b))^2
g1 <- f2/f1
b2 <- g1 # first estimate of the shape parameter β
tolerance <- 0.00001
while (abs(b – b2) > tolerance) { # Initializing the iterative process
f3 <- 2 * (y1^2) * gamma(2/b2)
f4 <- y2 * (gamma(1/b2))^2
g2 <- f4/f3
d1 <- b * g2 – b2 * g1
d2 <- b – g1 – b2 + g2
b3 <- d1/d2
b <- b3
f1 <- 2 * (y1^2) * gamma(2/b)
f2 <- y2 * (gamma(1/b))^2
g1 <- f2/f1
b2 <- g1
}
alpha.hat <- b * y1/gamma(1/b) # Computing the scale parameter α
beta.hat <- b
WeibullMME <- c(alpha.hat, beta.hat) # Parameter estimates
v1 <- gamma(1 + 1/b)
v2 <- gamma(1 + 2/b)
v3 <- b/gamma(1/b)
valpha.hat <- (v3^2) * (alpha.hat^2) * (v2 – v1^2)/n
sealpha.hat <- sqrt(valpha.hat) # standard error of alpha.hat
v4 <- gamma(1/b)
v5 <- gamma(2/b)
z <- x^2
vbeta.hat <- (v4^2/(2 * v5 * y1^2))^2 * var(z)/n
sebeta.hat <- sqrt(vbeta.hat) # standard error of beta.hat
SE <- c(sealpha.hat, beta.hat)
WeibullMME # Press Enter to return the estimates of the parameters
SE # Press Enter to return the standard errors of the estimates
```

iv. R Codes for Regression Method

```
# Program: WeibullRegM
# Estimating the Weibull Parameters using Regression Method
x <- c(data)
n <- length(x) # Sample Size
z <- sort(x)
k <- log(z)
i <- 1:n
F <- i/(n+1)
y <- log(-log(1-F))
syk <- sum(y*k) – sum(y)*sum(k)/n
ssy <- sum(y^2) – ((sum(y))^2)/n
ssk <- sum(k^2) – ((sum(k))^2)/n
beta.hat <- syk/ssk # Estimate of the Weibull shape parameter
```

```
b0 <- mean(y) – beta.hat*mean(k)
alpha.hat <- exp(-b0/beta.hat) # Estimate of the Weibull scale parameter
WeibullReg <- c(alpha.hat, beta.hat) # Parameter estimates based on Regression method
sse <- ssy – beta.hat*syk # Error sum of squares
mse <- sse/(n-2) # Mean square error
vbeta.hat <- mse/ssk
sebeta.hat <- sqrt(vbeta.hat)
vb0 <- mse*(1/n + ((mean(k))^2)/ssk)
seb0 <- sqrt(vb0)
sealpha.hat <- exp(-seb0/beta.hat)
seWeibullReg <- c(sealpha.hat, sebeta.hat) # Standard errors of the parameter estimates
WeibullReg # Press enter to return the Weibull parameter estimates
seWeibullReg # Press enter to return the standard errors of the parameter estimates
```

5 Application to Data

5.1 Estimation of Parameters

The monthly meteorological data for illustration is from Maiduguri and are given in Table 2.1. The Weibull and Raleigh distributions were fitted to the data. The computer programs given above were used in obtaining estimates of the parameters and their standard errors for the four methods of estimation considered in this study. The initial values for the parameters α and β were computed from the wind speed values as the sample mean and sample kurtosis, respectively. These provided reasonable guess estimates because α (in units of wind speeds) and β (dimensionless) are the scale and shape parameters, respectively of the weibull distribution. The relaxation factor, ρ, which helps to stabilize the Newton-Raphson iterative process on the path of convergence in the optimization method is assumed to be 0.5.The program for the regression method is straight forward using the mean ranks. The results for the Weibull and Raleigh distributions are given in Tables 5.1 and 5.2, respectively for the four methods of estimation considered, while the graphs of the observed and fitted distributions are given in Figures 5.1a - 5.1c.

Table 5.1 Parameter Estimates for the Weibull Distribution

	Maximum Likelihood Method				Optimization Method				Method of Moments				Regression Method			
Month	α	Se(α)	β	Se(β)	α	Se(α)	β	Se(β)	α	Se(α)	β	Se(β)	α	Se(α)	β	Se(β)
Jan	3.12	.191	3.42	.566	3.12	.191	3.42	.566	3.12	.211	3.26	.386	3.19	.961	2.62	.101
Feb	3.47	.208	3.34	.559	3.47	.209	3.34	.559	3.46	.232	3.22	.391	3.51	.965	2.80	.088
Mar	3.82	.197	3.99	.659	3.82	.197	3.99	.615	3.82	.212	3.96	.420	3.85	.968	3.54	.095
Apr	3.86	.188	4.26	.698	3.86	.188	4.26	.648	3.86	.199	4.29	.426	3.89	.967	3.76	.101
May	4.10	.183	4.62	.729	4.10	.183	4.62	.729	4.10	.204	4.46	.411	4.15	.958	3.74	.122
Jun	4.60	.253	3.83	.616	4.60	.253	3.83	.616	4.59	.275	3.72	.415	4.64	.961	3.25	.092
Jul	4.13	.252	3.39	.521	4.13	.252	3.39	.521	4.12	.265	3.37	.411	4.17	.960	2.95	.094
Aug	3.11	.187	3.43	.549	3.11	.187	3.43	.549	3.11	.203	3.31	.389	3.16	.972	2.75	.076
Sep	2.91	.180	3.34	.486	2.91	.180	3.34	.486	2.90	.182	3.46	.420	2.95	.972	3.24	.092
Oct	2.49	.160	3.22	.490	2.49	.160	3.22	.490	2.49	.166	3.23	.416	2.51	.977	2.90	.083
Nov	2.81	.203	2.81	.413	2.81	.203	2.81	.413	2.81	.208	2.82	.411	2.83	.973	2.58	.076
Dec	2.72	.185	3.04	.473	2.72	.185	3.04	.473	2.72	.194	2.99	.400	2.74	.979	2.67	.062

S E = Standard Error of the estimated parameter

Table 5.2 Parameter Estimates for the Rayleigh Distribution

	Maximum Likelihood Method		Optimization Method		Method of Moments		Regression Method	
Month	α	S.E(α)	α	S.E(α)	α	S.E(α)	α	S.E(α)
Jan	2.95	.2953	2.95	.2953	3.16	.3301	3.93	.9540
Feb	3.28	.3215	3.28	.3215	3.50	.3589	3.78	.9519
Mar	3.59	.3524	3.59	.3524	3.90	.4000	4.32	.9337
Apr	3.63	.3559	3.63	.3559	3.96	.4059	4.40	.9307
May	3.86	.3786	3.86	.3786	4.22	.4330	4.70	.9296
Jun	4.33	.4329	4.33	.4329	4.68	.4892	5.14	.9410
Jul	3.90	.3821	3.90	.3821	4.18	.4283	4.54	.9482
Aug	2.94	.2886	2.94	.2886	3.15	.3230	3.40	.9538
Sep	2.74	.2688	2.74	.2688	2.95	.3020	3.26	.9385
Oct	2.36	.2312	2.36	.2312	2.52	.2582	2.73	.9460
Nov	2.68	.2576	2.68	.2576	2.82	.2837	3.01	.9592
Dec	2.58	.2531	2.58	.2531	2.74	.2805	2.93	.9571

S E = Standard Error of the estimated parameter

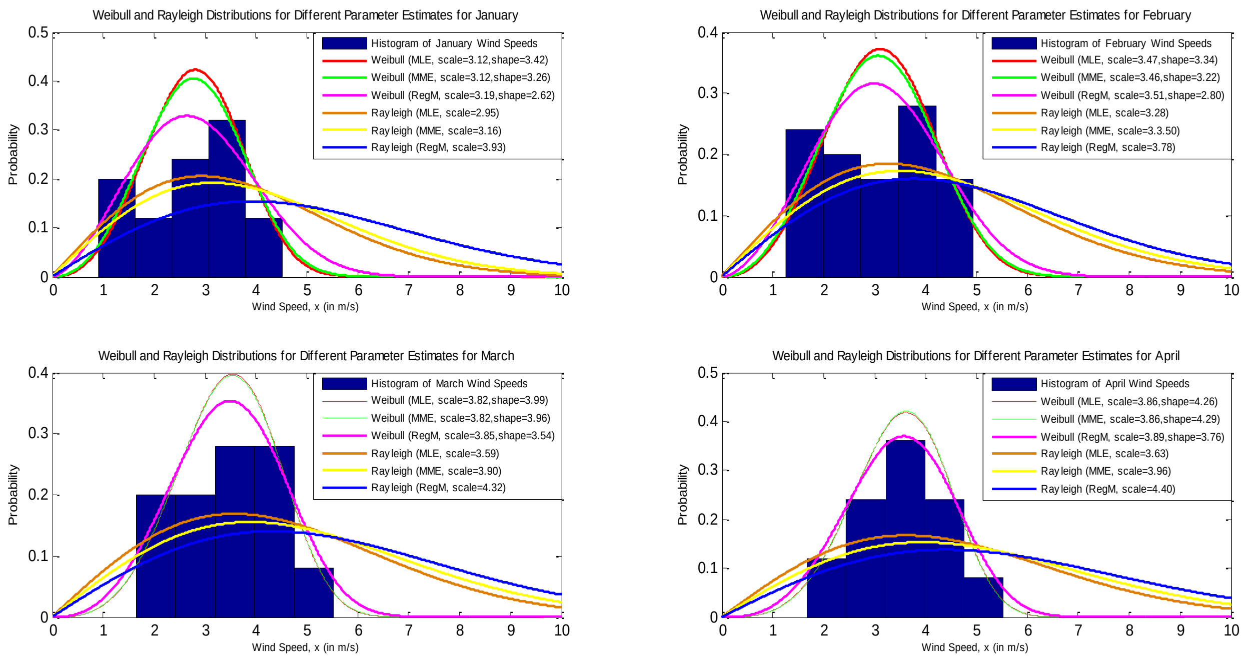

Figure 5.1a Density Histograms, Weibull and Rayleigh Curves for the various Methods of Parameter Estimates for Months of January, February, March and April

ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.6, No.7, 2016

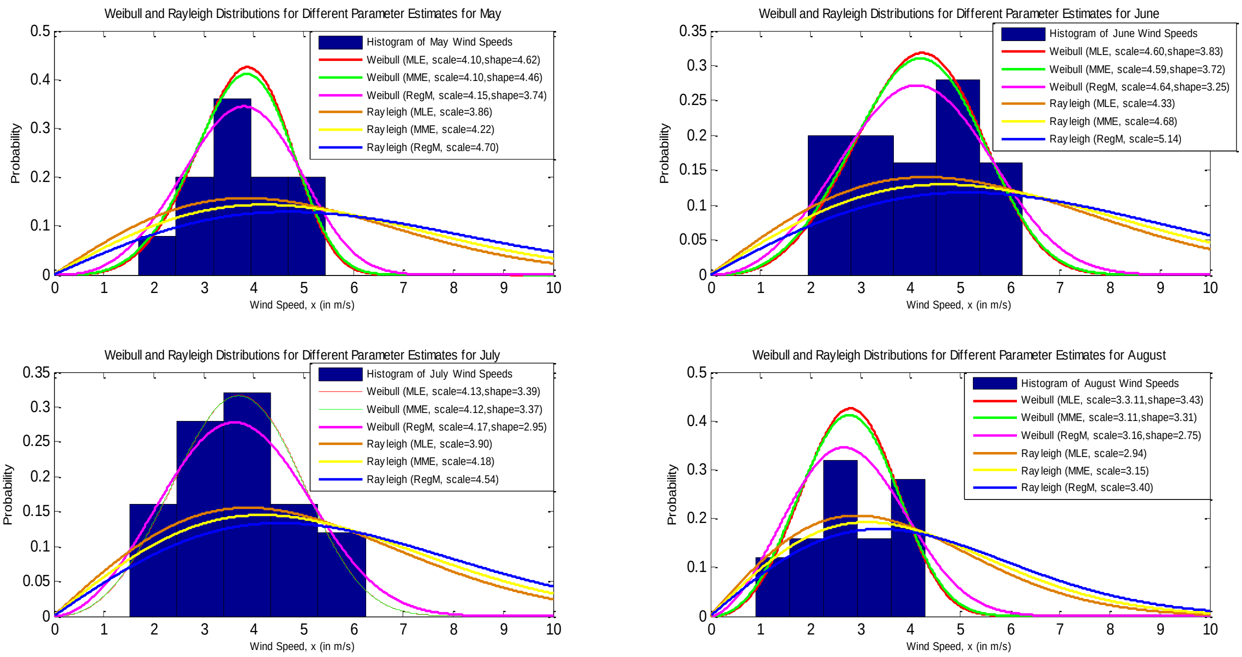

Figure 5.1b Density Histograms, Weibull and Rayleigh Curves for the various Methods of Parameter Estimates for Months of May, June, July and August.

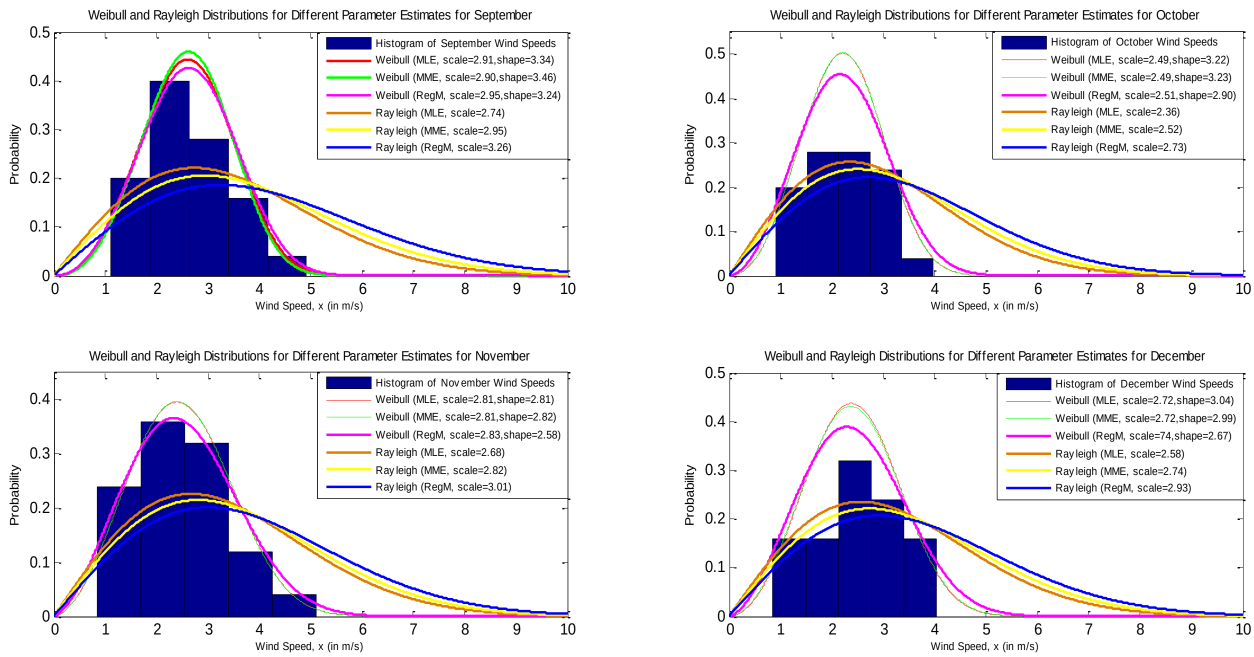

Figure 5.1c Density Histograms, Weibull and Rayleigh Curves for the various Methods of Parameter Estimates for Months of September, October, November and December.

5.2 Goodness-of-Fit Test

Each model fitted to the wind data was assessed using four goodness-of-fit tests listed below

A. Chi-squared goodness-of-fit test is given by

$$\chi^2 = \frac{\sum_{j=1}^{k}(o_j - e_j)^2}{e_j} \qquad 5.1$$

B. Likelihood Ratio test (Bayo, 1984) given by

$$Y^2 = -2\sum_{j=1}^{k} o_j log\left(\frac{e_j}{o_j}\right) \quad 5.2$$

where o_j and e_j are the observed and expected values for the j^{th} interval. It should be noted that the probability for the j^{th} interval p_j involved in the computation of e_j was computed by integrating the density function of interest over the range of the interval for which p_j is appropriate.

C. Kolmogorov-Smirnov test is given by

$$D = \max_{x_{(i)}}\left|F_o\left(x_{(i)}\right) - F_n\left(x_{(i)}\right)\right| \quad 5.3$$

D. Anderson-Darling test (Atif and Elif, 2012) is given by

$$A^2 = -\frac{1}{n}\sum_{j=1}^{n}(2j-1)\left\{log\left(F_o\left(x_{(i)}\right)\right) + log\left(1 - F_o\left(x_{(n-i+1)}\right)\right)\right\} - n \quad 5.4$$

The results of the goodness-of-fit tests for Weibull and Rayleigh distributions are provided in Tables 5.4 and 5.4, respectively.

Table 5.3 Test of Goodness-of-Fit of the Weibull Distribution

		Chi-Square			**Likelihood Ratio**			**Kolmogorov-Smirnov**			**Anderson-Darling**		
Month/	Method	χ^2	C-V	D	Y^2	C-V	D	K-S	C-V	D	A^2	C-V	D
Jan	MLE	0.696	3.84	DNR	0.985	3.84	DNR	0.106	0.281	DNR	0.458	0.740	DNR
	MME	0.676	3.84	DNR	1.132	3.84	DNR	0.100	0.281	DNR	0.422	0.740	DNR
	RgME	1.400	3.84	DNR	3.348	3.84	DNR	0.107	0.281	DNR	0.551	0.740	DNR
Feb	MLE	0.719	3.84	DNR	2.441	3.84	DNR	0.101	0.275	DNR	0.312	0.745	DNR
	MME	0.683	3.84	DNR	2.608	3.84	DNR	0.091	0.275	DNR	0.290	0.745	DNR
	RgME	0.649	3.84	DNR	4.141	3.84	DNR	0.068	0.275	DNR	0.382	0.745	DNR
Mar	MLE	2.483	3.84	DNR	2.455	3.84	DNR	0.088	0.275	DNR	0.178	0.745	DNR
	MME	2.370	3.84	DNR	2.366	3.84	DNR	0.086	0.275	DNR	0.173	0.745	DNR
	RgME	1.068	3.84	DNR	1.463	3.84	DNR	0.064	0.275	DNR	0.194	0.745	DNR
Apr	MLE	0.233	3.84	DNR	0.305	3.84	DNR	0.064	0.275	DNR	0.133	0.745	DNR
	MME	0.230	3.84	DNR	0.297	3.84	DNR	0.066	0.275	DNR	0.134	0.745	DNR
	RgME	0.567	3.84	DNR	0.878	3.84	DNR	0.053	0.275	DNR	0.217	0.745	DNR
May	MLE	0.990	3.84	DNR	1.097	3.84	DNR	0.078	0.275	DNR	0.181	0.745	DNR
	MME	0.731	3.84	DNR	0.915	3.84	DNR	0.086	0.275	DNR	0.169	0.745	DNR
	RgME	0.238	3.84	DNR	1.216	3.84	DNR	0.108	0.275	DNR	0.324	0.745	DNR
Jun	MLE	0.686	3.84	DNR	2.619	3.84	DNR	0.074	0.281	DNR	0.216	0.740	DNR
	MME	0.725	3.84	DNR	2.816	3.84	DNR	0.075	0.281	DNR	0.194	0.740	DNR
	RgME	0.953	3.84	DNR	4.469	3.84	R	0.058	0.281	DNR	0.211	0.740	DNR
Jul	MLE	0.312	3.84	DNR	1.184	3.84	DNR	0.193	0.275	DNR	0.209	0.745	DNR
	MME	0.318	3.84	DNR	1.188	3.84	DNR	0.195	0.275	DNR	0.206	0.745	DNR
	RgME	0.817	3.84	DNR	2.661	3.84	DNR	0.177	0.275	DNR	0.268	0.745	DNR
Aug	MLE	1.209	3.84	DNR	2.633	3.84	DNR	0.504	0.275	R	0.255	0.745	DNR
	MME	1.060	3.84	DNR	2.731	3.84	DNR	0.500	0.275	R	0.228	0.745	DNR
	RgME	0.892	3.84	DNR	4.508	3.84	R	0.462	0.275	R	0.330	0.745	DNR
Sep	MLE	0.804	3.84	DNR	0.908	3.84	DNR	0.065	0.269	DNR	0.291	0.745	DNR
	MME	0.846	3.84	DNR	0.916	3.84	DNR	0.073	0.269	DNR	0.278	0.745	DNR
	RgME	0.801	3.84	DNR	0.975	3.84	DNR	0.072	0.269	DNR	0.329	0.745	DNR
Oct	MLE	4.280	3.84	R	4.300	3.84	R	0.094	0.275	DNR	0.215	0.745	DNR
	MME	4.262	3.84	R	4.278	3.84	R	0.094	0.275	DNR	0.216	0.745	DNR
	RgME	5.143	3.84	R	5.420	3.84	R	0.089	0.275	DNR	0.252	0.745	DNR
Nov	MLE	0.953	3.84	DNR	1.179	3.84	DNR	0.073	0.269	DNR	0.301	0.745	DNR
	MME	0.974	3.84	DNR	1.196	3.84	DNR	0.073	0.269	DNR	0.302	0.745	DNR
	RgME	0.576	3.84	DNR	0.997	3.84	DNR	0.072	0.269	DNR	0.329	0.745	DNR
Dec	MLE	3.859	3.84	R	4.372	3.84	R	0.091	0.275	DNR	0.260	0.745	DNR
	MME	3.734	3.84	DNR	4.293	3.84	R	0.088	0.275	DNR	0.244	0.745	DNR
	RgME	3.138	3.84	DNR	4.234	3.84	R	0.074	0.275	DNR	0.240	0.745	DNR

C-V = Critical Value, D = Decision, DNR = Do not reject H_o, R = Reject H_o

Table 5.4 Test of Goodness-of-Fit of the Rayleigh Distribution

		Chi-Square			Likelihood Ratio			Kolmogorov-Smirnov			Anderson-Darling		
Month/ Method		χ^2	C-V	D	Y^2	C-V	D	K-S	C-V	D	A^2	C-V	D
Jan	MLE	5.65	5.99	DNR	8.47	5.99	R	0.205	0.275	DNR	1.645	0.745	R
	MME	4.68	5.99	DNR	8.58	5.99	R	0.162	0.275	DNR	1.272	0.745	R
	RgME	5.11	5.99	DNR	14.01	5.99	R	0.297	0.275	R	2.780	0.745	R
Feb	MLE	5.14	5.99	DNR	9.85	5.99	R	0.152	0.269	DNR	1.502	0.745	R
	MME	3.99	5.99	DNR	10.42	5.99	R	0.115	0.269	DNR	1.251	0.745	R
	RgME	3.34	5.99	DNR	12.03	5.99	R	0.165	0.269	DNR	1.463	0.745	R
Mar	MLE	5.76	5.99	DNR	9.05	5.99	R	0.213	0.269	DNR	2.457	0.745	R
	MME	4.61	5.99	DNR	9.35	5.99	R	0.159	0.269	DNR	1.957	0.745	R
	RgME	4.34	5.99	DNR	11.35	5.99	R	0.206	0.269	DNR	2.248	0.745	R
Apr	MLE	9.83	5.99	R	12.91	5.99	R	0.251	0.269	DNR	2.934	0.745	R
	MME	9.04	5.99	R	13.44	5.99	R	0.200	0.269	DNR	2.372	0.745	R
	RgME	9.35	5.99	R	15.82	5.99	R	0.218	0.269	DNR	2.674	0.745	R
May	MLE	8.61	5.99	R	12.84	5.99	R	0.260	0.269	DNR	3.194	0.745	R
	MME	7.40	5.99	R	13.57	5.99	R	0.197	0.269	DNR	2.567	0.745	R
	RgME	7.30	5.99	R	16.25	5.99	R	0.238	0.269	DNR	2.849	0.745	R
Jun	MLE	8.99	5.99	R	14.09	5.99	R	0.209	0.275	DNR	2.025	0.745	R
	MME	7.57	5.99	R	14.87	5.99	R	0.157	0.275	DNR	1.598	0.745	R
	RgME	6.97	5.99	R	17.12	5.99	R	0.191	0.275	DNR	1.822	0.745	R
Jul	MLE	6.05	5.99	R	9.44	5.99	R	0.282	0.269	R	1.695	0.745	R
	MME	5.34	5.99	DNR	10.16	5.99	R	0.157	0.269	DNR	1.358	0.745	R
	RgME	5.21	5.99	DNR	12.07	5.99	R	0.177	0.269	DNR	1.547	0.745	R
Aug	MLE	5.43	5.99	DNR	9.91	5.99	R	0.486	0.269	R	1.615	0.745	R
	MME	4.09	5.99	DNR	10.44	5.99	R	0.438	0.269	R	1.243	0.745	R
	RgME	3.37	5.99	DNR	11.99	5.99	R	0.382	0.269	R	1.362	0.745	R
Sep	MLE	5.43	5.99	DNR	7.39	5.99	R	0.197	0.264	DNR	2.217	0.745	R
	MME	4.77	5.99	DNR	7.55	5.99	R	0.151	0.264	DNR	1.765	0.745	R
	RgME	4.94	5.99	DNR	9.307	5.99	R	0.220	0.264	DNR	2.010	0.745	R
Oct	MLE	10.86	5.99	R	11.99	5.99	R	0.185	0.269	DNR	1.507	0.745	R
	MME	10.72	5.99	R	12.54	5.99	R	0.144	0.269	DNR	1.218	0.745	R
	RgME	11.31	5.99	R	14.20	5.99	R	0.152	0.269	DNR	1.386	0.745	R
Nov	MLE	1.39	5.99	DNR	2.55	5.99	DNR	0.138	0.264	DNR	1.043	0.745	R
	MME	1.00	5.99	DNR	2.65	5.99	DNR	0.115	0.264	DNR	0.863	0.745	R
	RgME	0.99	5.99	DNR	3.47	5.99	DNR	0.157	0.264	DNR	0.991	0.745	R
Dec	MLE	5.30	5.99	DNR	7.33	5.99	R	0.167	0.269	DNR	1.110	0.745	R
	MME	4.35	5.99	DNR	7.40	5.99	R	0.126	0.269	DNR	0.867	0.745	R
	RgME	3.93	5.99	DNR	8.25	5.99	R	0.112	0.269	DNR	0.990	0.745	R

C-V = Critical Value, D = Decision, DNR = Do not reject H_o, R = Reject H_o

5.3 Root-Mean-Square-Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{j=1}^{k}\left(f_o(x_j)-f(x_j)\right)^2}{k-s-1}} \qquad 5.5$$

where x_j is the mid-value for the j^{th} interval, $f_o(x)$ is the observed frequency distribution, $f(x)$ is the fitted distribution and s is the number of parameters estimated. RMSE was computed in each case for the four methods of parameter estimation considered. The method of estimation with the smallest RMSE is adjudged the best. The results are tabulated in Table 5.5

Table 5.6 Root-Mean-Square-Error

	Weibull			Rayleigh		
Month	MLE	MME	RegM	MLE	MME	RegM
January	.0557	.0536	.0529	.1110	.0869	.1439
February	.0443	.0421	.0390	.0971	.0784	.0844
March	.0337	.0329	.0255	.1307	.1047	.1122
April	.0261	.0266	.0276	.1469	.1207	.1273
May	.0343	.0324	.0359	.1549	.1274	.1328
June	.0398	.0371	.0268	.1169	.0918	.0993
July	.1216	.1237	.1125	.1727	.1341	.0993
August	.3516	.3491	.3252	.3346	.2990	.2223
September	.0343	.0347	.0358	.1198	.0972	.1052
October	.0385	.0386	.0374	.1011	.0812	.0857
November	.0407	.0409	.0375	.0766	.0619	.0673
December	.0446	.0428	.0355	.0817	.0622	.0659

6 Discussions of Results

The results for the Weibull model parameters estimates shown in Table 5.1 indicate clearly that there are virtually no differences between the estimates obtained for maximum likelihood and optimization methods. Same is true for the Rayleigh distribution results shown in Table 5.2. This is not surprising because the objective function in both cases is the log-likelihood function. The difference in the two methods being only in the way the objective function is maximized; one is through differentiation and the other by optimization. Hence, the result for the optimization method is subsequently omitted from the table of results. The estimates obtained by maximum likelihood and method of moments are quite close. Those for the regression method are within twice the standard error (approximate 95% confidence limits) of the corresponding maximum likelihood estimates. That is, the regression method estimates are possible values for those of the maximum likelihood method. Consequently, any of these four methods can be used in estimating the parameters of the Weibull model.

The Rayleigh distribution parameter estimates shown in Table 5.2 indicate that the three methods of estimation have similar estimates; differences are observed only in the first decimal place. However, estimates obtained for the maximum likelihood method are consistently lower than those of the others over the months of the year.

The graphs of the fitted models (Weibull and Rayleigh) to the data shown in Figures 5.1a – c shows that the Rayleigh distribution is consistently a poor fit for each method of estimation examined. This conclusion is buttressed by the results of the goodness-of-fit tests for the Rayleigh distribution shown in Table 5.4. Rayleigh model is rejected in almost all the cases considered by the four types of goodness-of-fit tests examined. On the other hand, the graphs for the Weibull model in Figure 5.1a – c appear to fit the data for the three methods of parameter estimation examined. The results of the goodness-of-fit tests shown in Table 5.3 also confirm this; the Weibull model was not rejected for most cases considered by the three methods of estimation used in computing the parameter values.

The results for the root-mean-square-error shown in Table 5.5 show that estimation of parameters by regression method has the smallest RMSE estimates for almost all the months of the year. However the difference in the RMSE estimates for the three methods of estimation considered are not pronounced. Corresponding RMSE estimates for the Rayleigh model are quite higher than those of Weibull. This again indicates that the Weibull model is a better fit to the data considered.

7 Conclusion

The use of four methods of estimation of parameters in a Weibull distribution was discussed. These are maximum likelihood, method of moments, optimization and regression methods. They were implemented on the monthly meteorological wind data from Maiduguri for illustrative purpose. Weibull and Rayleigh models were both applied to the data. The results show that the Weibull model is a good fit to the data while the Rayleigh model was rejected in most cases considered. The differences in the corresponding parameter and RMSE estimates obtained are thin. It is noted that parameter estimation by regression method is the easiest to implement and the Weibull model parameter estimates obtained by regression method consistently gave the smallest RMSE. These two reasons suggest that the regression method can be a potential user’s first choice in obtaining parameter estimates for a Weibull wind model.

References

Ahmed S. A. & Mohammed H. O. (2012). A Statistical Analysis of Wind Power Density Based on the Weibull and Rayleigh Models of “Penjwen Region” Sulaimani/Iraq. *Jordan Journal of Mechanical and Industrial Engineering,* Volume 6, No. 2, pp 135-140.

Atif E. & Elif T. (2012). On Some Properties of Goodness-of-fit Measures Based on Statistical Entropy. [online] Available @ www.arpapress.com/Volumes/Vol13Issue1/IJRRAS_13_1_22.pdf, October 2012

Bayo L. H. (1984). Comparisons of χ^2, Y^2, Freeman-Tukey and Williams’s Improved G^2 Test Statistics in Small Samples of One-way Multinomial. *Biometrika*, **71(2)**, pp. 415-8.

Constantinides A. (1987). *Applied Numerical Methods and Applied Computers*. McGraw-Hill, New York.

Dikko I. & Yahaya D. B. (2012). Evaluation of Wind Power Density in Gombe, Yola and Maiduguri, North Eastern Nigeria. *Journal of Research in Peace, Gender and Development,* Vol. 2(5) pp.115-122.

Fritz S. (2008). Inference for the Weibull Distribution. *Stat 498B Industrial Statistics* [online] Available @ https://www.stat.washington.edu/people/fritz/DATAFILES498B2008/WeibullBounds.pdf, May 22, 2008

Johnson N. L. & Kotz S. (1970). *Continuous Univariate Distributions*. Vols 1 and 2. Houghton Mifflin Company, Boston.

Kostas P. & Despina D. (2009). Statistical Simulation of Wind Speed in Athens, Greece based on Weibull and ARMA Models. *International Journal of Energy and Environment,* Vol.3 Issue 4, pp. 151-158.

Lawal S. & Isa G. (2013). Electricity Generation Using Wind in Katsina State, Nigeria. *International Journal of Engineering Research and Technology (IJERT),* Vol.2(2),pp 1-13

LDABOOK[8]. The Weibull Distribution. *Weibull.com, ReliaWiki.org : ReliaSoft Corporation*

Mood A. M., Graybill F. A., & Boes D. C. (1974). *Introduction to the Theory of Statistics*. International Student Edition, McGraw-Hill Edition.

Nikolai N., Gudrun N. P., Halldor B., Andrea N. H., Kristjan J., Charlotte B. H & Niels-Erik C. (2014). The Wind Energy Potential of Iceland. *Renewable Energy* V0l.69 pp 290-299.

Odo F. C., Offiah S.U. & Ugwuoke P. E. (2012). Weibull Distribution-Based Model for Prediction of Wind Potential in Enugu, Nigeria. *Advances in Applied Science Research,* Pelagia Research Library, 3(2), pp 1202-1208.

Proma A. K., Pobitra K. H., & Sabbir R. (2014). Wind Energy Potential Estimation for Different Regions of Bangladesh. *International Journal of Renewable and Sustainable Energy,* Vol. 3(3), pp 47-52.

Salvadori, M. G. & Baron M. L. (1961). *Numerical Methods in Engineering*. 2^{nd} ed. Prentice Hall, Inc., Englewood Cliffs, N.J. Iteration and Relaxation Methods are discussed in Chapter 1.

Shamshad A., Wan H. W. M. A., Bawadi M. A. & Mohd. S. S. A. (2012). Analysis of Wind Speed Variation and Estimation of Weibull Parameters for Wind Power Generation in Malaysia, [online] Available @ https://core.ac.uk/download/files/423/11960230.pdf

Zhou W., Hongxing Y. & Zhaohong F. (2006). Wind Power Potential and Characteristic Analysis of Pearl River Delta Region, China. *Renewable Energy,* 31, 739 – 753.