International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 17
July 2018

Comparative Study of Modeling Wind Speed Data; a Case Study of Maiduguri, Nigeria

Obanla O. James[1*], Awariefe Christopher[2], Afolabi Nasimot O.[3]
1 & 3 Department of Statistics, Ogun State Institute of Technology, Igbesa, Ogun State, Nigeria
2. Department of Statistics, Delta State Polytechnics, Ozoro, Delta State, Nigeria.

ABSTRACT

Wind speed is the most important factor of the wind energy, because of the random nature of wind speed statistical methods are useful in modeling it. The aim of this study is to model wind speed by considering three different probability distributions which are well suited in fitting distributions with varieties of shapes which are widely used in scientific field and often used to model random effect, monitoring environment, wind and rainfall size: are two parameter Lognormal, the two-parameter Weibull and the two-parameter Gamma. These distributions are applied on the wind speed data recorded at the Nigerian Meteorological Agency (NiMET) Office in Maiduguri, at hub height of 10 meters. The period covered by the data is September, 1985 to December, 2011. In order to determine the distribution that best fit the wind data we used the MLE as method of parameter estimation and Kolmogorov-Smirnov test for decision guide. Based on the results the two-parameter Weibull performs the best fit on the wind speed data histogram.

Keywords: Wind speed, Weibull distribution, Gamma distribution, Lognormal distribution, Maximum Likelihood Estimator, Kolmogorov Smirnov test

1.0 INTRODUCTION

Energy is a quantitative property that can be transferred to an object in order to perform work on it. It is a conserved quantity that can be converted in form but cannot be created or destroyed. Wind energy is an environment friendly and renewable energy like solar and is theoretically inexhaustible. Wind speed is a clean energy which is an alternative energy compared to fossils fuel that causes pollution to lower part of the atmosphere. Since wind energy is a clean and renewable energy, systems transforming wind power to electrical energy has been on the rise of recent Aras, et al. (2003). Currently in Nigeria, the sources of electricity generation is mainly through water dams and thermal energy via gas and there is still serious shortfall of energy requirement for the teeming population of Nigeria, wind energy can be a very good source of energy that will add to its inadequate megawatt that is less than 5,000. Since wind speed is the most important variable of the wind energy, therefore, an accurate estimation of probability distribution of wind speed values is very important in evaluating wind speed energy potential of a region. Yilmaz and Çelik (2008).The gamma, Weibull and Lognormal distributions are useful in modeling life time data. The Gamma distribution is useful in lifetime Meeker and Escobar (2008) and signal processing Vasegi (2008). Bhaumik (2009) used gamma distribution to modeled physical and biological data. The Weibull distribution is commonly used in the practical studies related to the wind energy modeling Auwera et al. (1980).

Probability distribution are used to model random phenomenon, like wind speed. Several probability distributions have been used in literature to describe the wind speed. Celik (2003), Stevens and Smulders (1979) and Carta and Ramírezes (2007) used Weibull distribution describe wind speed. Other authors such as Morgan et al (2011), Yilmaz and Celik (2008) and Zhou et al (2010) used and compared the other probability distributions that provide better fit to the wind speed . The objective of this study is to model wind speed by considering three different probability distributions which are well suited in fitting distributions with varieties of shapes and are widely used in scientific field and often used to model random effect, monitoring environment, wind and rainfall size: the two parameter Lognormal, the two-parameter Weibull and the two-parameter Gamma will be considered.

2.0 METHODS

2.1. Probability Density Function (PDF) of Weibull Distribution

Weibull cumulative distribution function is defined as

$$F(x) = 1 - \exp\left[-\left(\frac{t}{\beta}\right)^{\alpha}\right] \qquad (2.1)$$

While the two-parameter Weibull probability density function is given as

$$f(x) = \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{t}{\beta}\right)\alpha\right], t \geq 0, \alpha, \beta > 0 \qquad (2.2)$$

The two parameter Weibull distribution can be written as a special case of the three parameter Weibull probability distribution where $\delta = 0$. When $\beta = 1$, in this research work we will be dealing with the two-parameter Weibull distribution.

2.2 Maximum Likelihood Estimation of Weibull Distribution

Let $x_1, x_{2,\dots,}, x_n$ be a random sample of size n drawn from a probability density function f(x;θ), where θ is an unknown parameter, then the likelihood function of this random sample is the joint density of the n random variables and is a function of the unknown parameter. Consider the Weibull probability density function;

$$L(x_1,\dots,x_n, \alpha, \beta) = \prod_{i}^{n}\left(\frac{\alpha}{\beta}\right)\left(\frac{xi}{\beta}\right)^{\alpha-1} e^{\left(\frac{xi}{\beta}\right)\alpha} \qquad (2.3)$$

Taking the logarithm and differentiating with respect to α and β and equating to zero. We obtain the following equation

$$\frac{\partial \ln L}{\partial \alpha} = \frac{n}{\beta} + \sum_{i=1}^{n} \ln xi - \frac{1}{\beta}\sum_{i=1}^{n} x^{\alpha} \ln xi = 0 \qquad (2.4)$$

$$\frac{\partial \ln L}{\partial \beta} = -\frac{n}{\beta} + \frac{1}{\beta^2}\sum xi^{\alpha} \ln xi = 0 \qquad (2.5)$$

Eliminating β between these two equations and simplifying we have

$$\frac{\sum_{i=1}^{n} xi^{\alpha} \ln xi}{\sum xi^{\alpha}} - \frac{1}{\alpha} - \frac{1}{n}\sum_{i=1}^{n} \ln xi = 0 \qquad (2.6)$$

This may be used to get $\hat{\mu}_k = \alpha$ using Newton –Raphson method. Once α is determined, β can be estimated using equation (2.6) as

$$\beta = \frac{\sum_{i=1}^{n} xi^{\alpha}}{n} \quad (2.7)$$

2.3 Probability Density Function (PDF) of Gamma Distribution

Gamma cumulative distribution function is defined as

$$F(x) = \frac{\gamma\left[\!\left[\alpha, \frac{x}{\beta}\right]\!\right]}{\Gamma(\alpha)} \quad (2.8)$$

While the two-parameter gamma probability density function is given as

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases} \quad (2.9)$$

The two parameter Weibull distribution is a special case of the three parameter Weibull probability distribution

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases} \quad (2.10)$$

$$l(\alpha,\beta) = \sum_{i=1}^{n} [-\log\Gamma(\alpha) - \alpha\log\beta + (\alpha-1)\log X_i - X_i/\beta]$$

(2.11)

Finding the partial derivatives, we have

$$\frac{\partial l}{\partial \alpha} = \sum_{i=1}^{n} \left[-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \log\beta + \log X_i \right]$$
$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^{n} \left[-\frac{\alpha}{\beta} + \frac{X_i}{\beta^2} \right] \quad (2.12)$$

Setting the second partial derivative equal to zero, we obtain

$$\hat{\beta}_{MLE} = \frac{\sum_{i=1}^{n} X_i}{n\hat{\alpha}_{MLE}} \qquad (2.13)$$

When this solution is substituted into the first partial derivative, we obtain a nonlinear equation for the MLE of α :

$$-n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - n\log\frac{\sum_{i=1}^{n} X_i}{n} + n\log\hat{\alpha}_{MLE} + \sum_{i=1}^{n}\log X_i = 0 \quad (2.14)$$

This equation cannot be solved in closed form. Newton's method or another iterative method can be used.

However, $\psi(\alpha) = ln\alpha - ln\bar{x} - \frac{1}{n}\sum_{i=1}^{n} lnx_i$ (2.15)

2.4. Probability Density Function (PDF) of Lognormal Distribution

Lognormal cumulative distribution function is defined as

F(x) = $\phi\left(\frac{lnx-\mu}{\sigma}\right)$ for > 0 , $\mu\epsilon R$,$\sigma > 0$ (2.16)

The likelihood function of the lognormal distribution is derived by taking the product of the probability densities of the individual X_is:

L(μ, σ²|X) = $(2\pi\sigma^2)^{-n/2} \prod_{i=1}^{n} X^{-1}{}_i \exp \sum_{i=1}^{n} -(\ln(Xi) - \mu)^2$ (2.17)

The log-likelihood function of the lognormal distribution is then derived by finding the natural log of the likelihood function:

L(μ, σ²|X) = $\ln\left[(2\pi\sigma^2)^{-n/2} \prod_{i=1}^{n} X^{-1}{}_i \exp\left[\sum_{i=1}^{n} \frac{-(\ln(Xi) - \mu)^2}{2\sigma^2}\right]\right]$ (2.18)

$$= -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n}\ln(Xi) - \frac{\sum_{i=1}^{n} -(\ln(Xi) - \mu)^2}{2\sigma^2} \qquad (2.19)$$

$$= -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n}\ln(X_i) - \frac{\sum_{i=1}^{n} -(\ln(Xi)^2 - 2\ln(Xi)\mu + \mu)^2}{2\sigma^2} \qquad (2.20)$$

$$= -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n}\ln(X_i) - \frac{\sum_{i=1}^{n} -(\ln(Xi)^2}{2\sigma^2} + \frac{\sum_{i=1}^{n} 2(\ln(Xi)\mu}{2\sigma^2} - \frac{\sum_{i=1}^{n}\mu^2}{2\sigma^2} \qquad (2.21)$$

$$= -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n}\ln(X_i) - \frac{\sum_{i=1}^{n} -(\ln(Xi)^2}{2\sigma^2} + \frac{\sum_{i=1}^{n} \ln(Xi)\mu}{\sigma^2} - \frac{n\mu^2}{2\sigma^2} \qquad (2.22)$$

We now find ^μ and σ^2, which maximize L (μ, σ^2|X). To do this, we take the gradient of L with respect to μ and σ^2 and set it equal to 0: with respect to μ,

$$\frac{\delta L}{\delta \mu} = \frac{\sum_{i=1}^{n} \ln(Xi)}{\sigma^2} - \frac{2n\mu}{2\sigma^2} = 0 \qquad (2.23)$$

$$\frac{n\mu}{\sigma^2} = \frac{\sum_{i=1}^{n} \ln(Xi)}{\sigma^2} \qquad (2.24)$$

$$\mu = \frac{\sum_{i=1}^{n} \ln(Xi)}{n} \qquad (2.25)$$

;

With respect to σ^2,

$$\frac{\delta L}{\delta \sigma^2} = \frac{n}{2}\frac{1}{\sigma^2} - \frac{\sum_{i=1}^{n} (\ln(Xi) - \mu)^2}{2} (-\sigma^2)^{-2} \qquad (2.26)$$

$$= -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^{n} (\ln(Xi) - \mu)^2}{2\sigma^2} = 0 \qquad (2.27)$$

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 17
July 2018

$$= \frac{n}{2\sigma^2} = \frac{\sum_{i=1}^{n} (\ln(Xi) - \mu)^2}{2\sigma^4} \qquad (2.28)$$

$$= \text{n} = \frac{\sum_{i=1}^{n} (\ln(Xi) - \mu)^2}{\sigma^2} \qquad (2.29)$$

$$= \sigma^2 = \frac{\sum_{i=1}^{n} (\ln(Xi) - \mu)^2}{n} \qquad (2.30)$$

$$= \sigma^2 = \frac{\sum_{i=1}^{n} (\ln(Xi) - \frac{\sum_{i=1}^{n} (\ln(Xi)}{n})^2}{n} \qquad (2.31)$$

So, the maximum likelihood estimators of lognormal density function are

$$\hat{\mu} = \frac{1}{n}\sum lnx \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n}\sum(lnx - \frac{1}{n}\sum lnx)^2 \qquad (2.32)$$

2.5 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov, which is also known as K-S statistics belongs to the Supremum class of EDF statistics and this category of statistics is based upon the largest absolute difference between the hypothesized and empirical distribution (Conover, 1999). Given n order data points, $_1 < x_2 < \ldots < x_n$, The test statistics proposed by Kolmogorov (1993) was defined by Conover (1999) as,

$$T = \text{Sup}_x |F^*(x) – F_n(x)|$$

.Where ‘sup’ stands for supremum which means the greatest. $F^*(x)$ is the hypothesized distribution function and $F_n(x)$ is the EDF estimated based on the random sample. In K-S test of normality, $F^*(x)$ is taken to be a normal distribution with known mean, μ, and standard deviation σ,

The K-S test statistic is specifically for hypotheses testing,

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 17
July 2018

H_o; F(x) for all x from -∞ to ∞ (The data follow a specified or known distribution)

H_a; F(x) ≠ F*(x) for at least one value of x (the data do not follow the specified or known distribution)

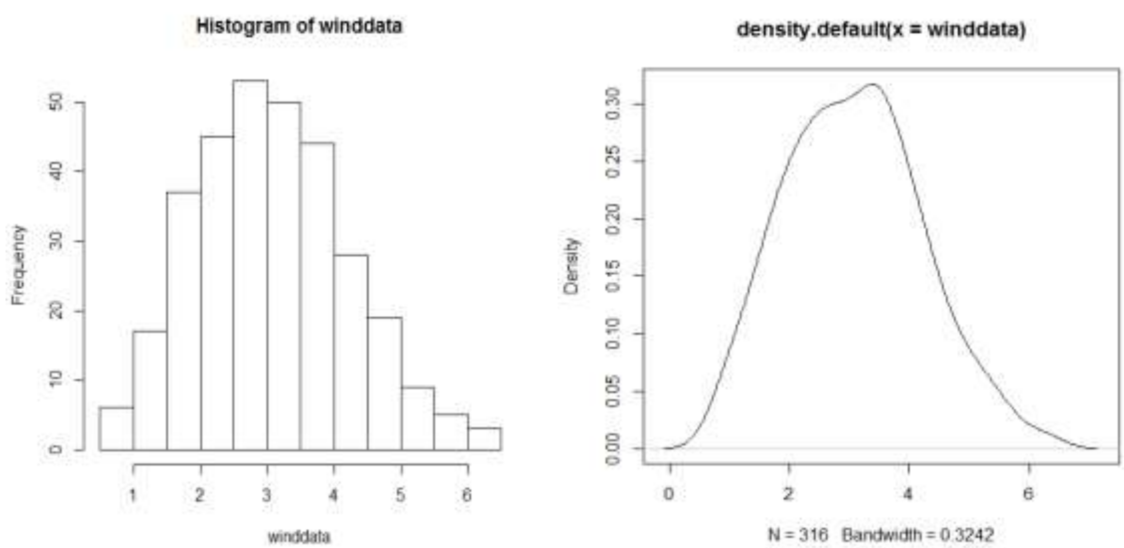
Decision rule states that if T is greater than the 1 – α quantile as given by the table of qualities for the Kolmogorov test statistic, then we reject H_0 at a specified level of significance, α.

3.0 Results and Discussion

3.1 Data Description

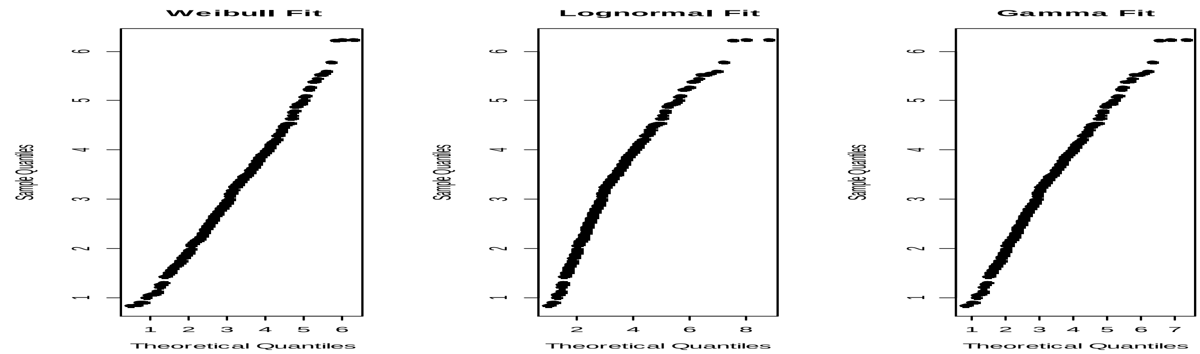
The data for the analysis is wind speed data adapted from Saporu and Esbond (2015) recorded at the Nigerian Meteorological Agency (NiMET) Office in Maiduguri, at hub height of 10 meters. The period covered by the data is September, 1985 to December, 2011. The R software was used to perform the estimation of the parameters for the graphical and analytical procedures.

Figure 1: Graphics of Histogram and Probability Density Function of the Wind Speed Data



e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 17
July 2018

Figure 2: Graphics of Q-Q Plots for Weibull, Gamma and Lognormal Distribution of the Wind



Speed Data

Figure 3: Graphics of fitted MLE for Weibull distribution

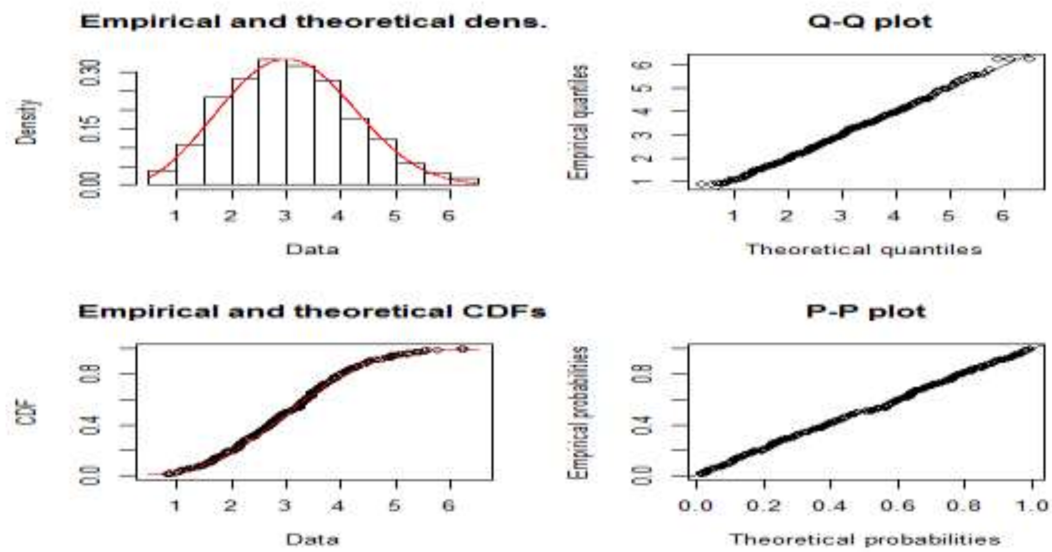


Figure 4: Graphics of fitted MLE for Gamma distribution

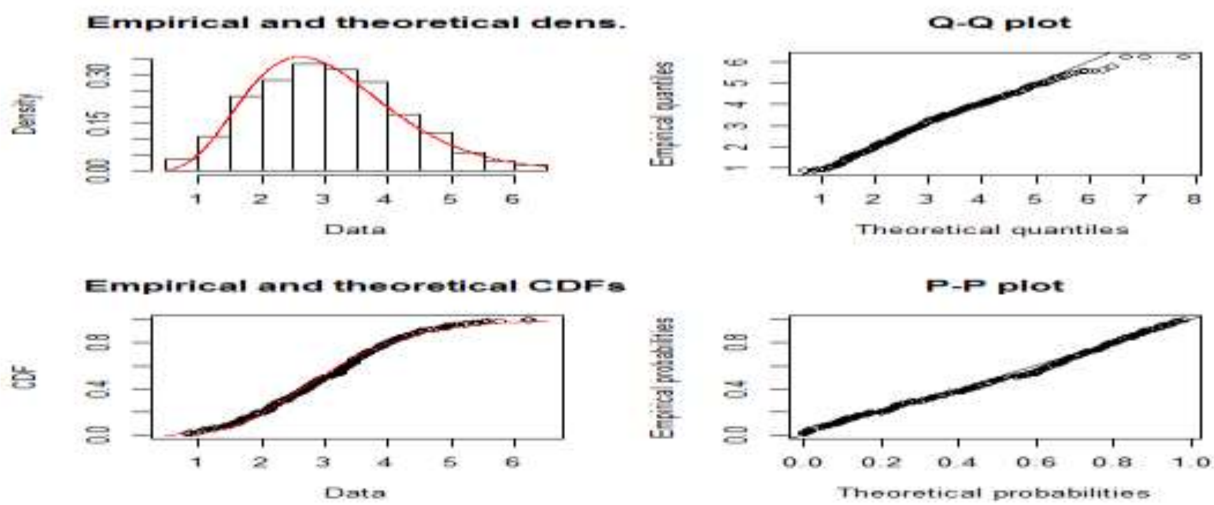
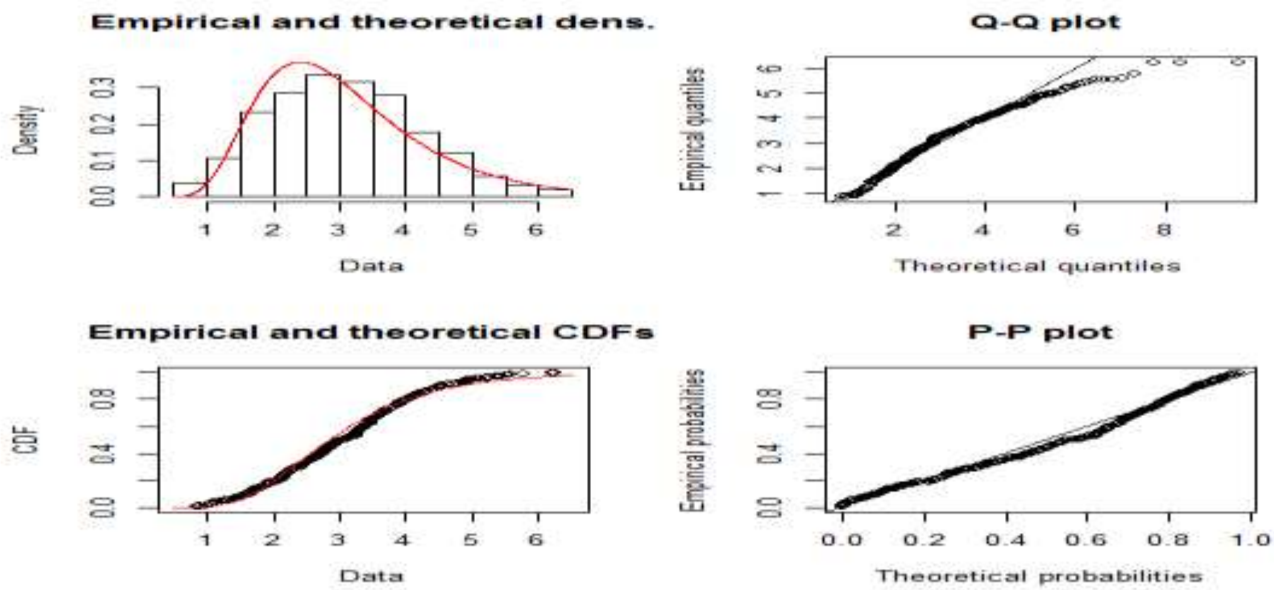
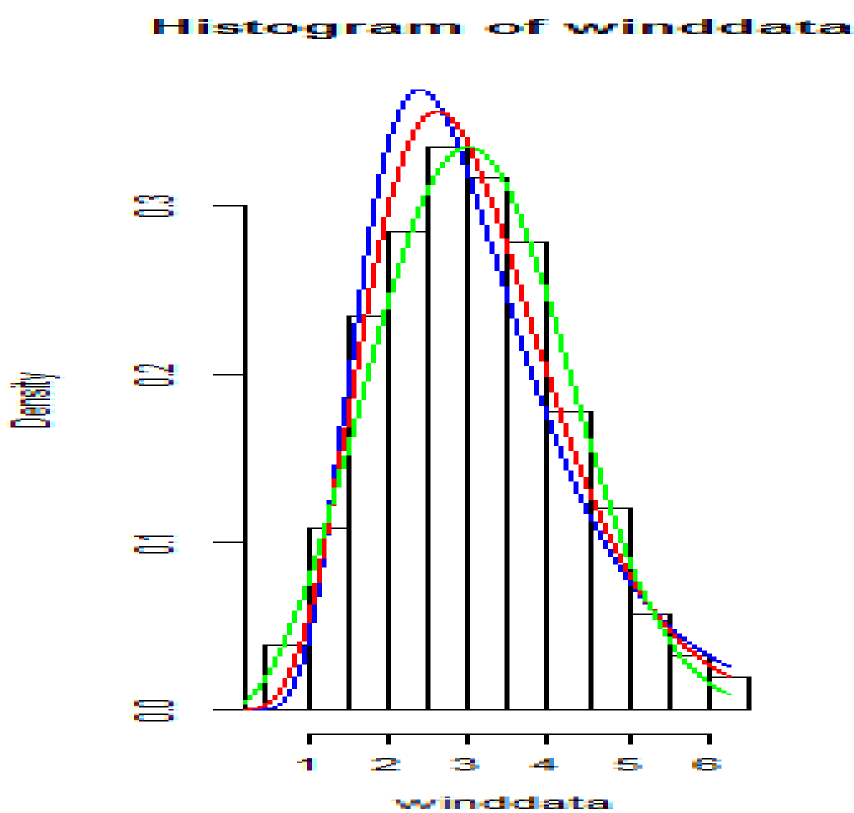


Figure 5: Graphics of fitted MLE for Lognormal distribution





In the histogram shown above on wind data comparing the fits of the three distribution Weibull(green), Gamma (red), Lognormal (blue). It is apparent that Weibull density function is the best fit on the histogram plotted using the wind data. In addition, Gamma density function follows that of Weibull in fitting the wind data. However, from Figure 3, 4 and 5 of the fitted MLE for Weibull, Gamma and Lognormal distributions, it is seen that the graphics of Weibull distribution is the best fit followed by Gamma distribution. The plot of the empirical distribution and the estimated Weibull distribution in Figure 1 reveals that shape and scale parameters' values estimated with MLE best fit with measured data set.

TABLE 1: Maximum likelihood method of the parameter estimates and K-S test

Distributions	ML estimates of parameters.	One sample K-S test

Gamma	α =6.6162 (0.5136) β =2.1496 (0.1733)	D = 0.07194 p-value = 0.0759
Weibull	α = 2.949 (0.1286) β = 3.4530 (0.0694)	D = 0.03176 p-value = 0.9074
Lognormal	μlog = 1.0467 (0.0232) σlog = 0.4127 (0.0164)	D = 0.0908 p-value = 0.0108

Table 1 above present shape and rate estimation values of Gamma, shape and scale estimates values of Weibull distribution as well as the parameters of lognormal distribution computed by using maximum likelihood and K—S test methods.

The decision rule for K-S test statistic is stated as;

H_o; F(x) for all x from -∞ to ∞ (The data follow a specified distribution)

H_a; F(x) ≠ F*(x) for at least one value of x (the data do not follow the specified distribution)

In Table 1, it is seen that parameters and standard errors estimated per MLE for Lognormal has the smallest values followed by Weibull and Gamma distributions. The parameters and standard error estimated per MLE of Weibull are smaller than that of Gamma distributions. Table 3 also reveals that Weibull distribution best fit the wind data per the K-S test since it has the smallest D and p-value compare to Gamma and Lognormal distributions, since low value for D indicates that the selected curve is fairly close to our data. The p-value indicates the chance that D was produced by the null hypothesis.

CONCLUSION

Based on the outcome of the analysis with respect to estimated parameters using maximum likelihood method and Kolmogorov-Smirnov (K-S) for the Weibull, Lognormal and Gamma distributions. Weibull distributions perform best compared to the Gamma and Lognormal distributions. Therefore, Weibull distribution can be used as the distribution that adequately describes the considered wind speed data in Maiduguri.

REFERENCES

[1]. Aras, H., Yilmaz, V., Celik, H.E. (2003). Estimation of Monthly Wind Speeds of
Eskişehir, Turkey. The First International Exergy, Energy and Environment Symposium, Hotel Princess, Izmir, Turkey.

[2]. Auwera V., Meyer L.F., Malet, L. M. (1980). The use of the Weibull three parameter model for estimating mean wind power densities. *Journal of Apply Meteorology*, vol.19,pp. 819–825.

[3]. Carta,J. A and Ramírezes, P. (2007) Analysis of two-component mixture Weibull statistics for estimation of wind speed distributions, Renewable Energy,32 (3), 518-531.

Celik J. N., A (2003) statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey. Renewable Energy, 29, 593-604.

[4]. Conover, W. J. (1999). Practical Nonparametric statistics. Third Edition, john Wiley & Sons, Inc. New York, pp.428-433 (6.1).

[5]. Ramírez P and Carta J. A.(2005) Influence of the data sampling interval in the estimation of parameters of the Weibull wind speed probability density distribution a case study, Energy Conversion and Management, 46 (2005), 2419–2438.

[6]. Saporu, F.W.O and Esbond, G.I (2015) Wind Energy Potential of Maiduguri, Borno State, Nigeria. International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064.

[7]. Seguro, J. W and Lambert,T. W. (2000) Modern estimation of the parameters of the Weibull speed distribution for wind energy analysis, Journal of Wind Engineering and Industrial Aerodynamics, 85, 75-84.

[8]. Stevens ,J. and Smulders, P. T. (1979) The estimation of the parameters of the Weibull wind speed distribution for wind energy utilization purposes, Wind Engineering, vol 3(2), , 132-45.

[9]. Zhou, J., Erdem, E. G. Li and Shi, J. (2010) Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites, Energy Conversion and Management, 51, 1449-1458.