COMPARISON OF ESTIMATION METHODS FOR FITTING WEIBULL DISTRIBUTION TO THE NATURAL STAND OF OLUWA FOREST RESERVE, ONDO STATE, NIGERIA

Ogana, F.N.
Department of Forest Resources Management, University of Ibadan, Ibadan, Nigeria
E-mail: Ogana_fry@yahoo.com; Phone: +234 (0)7037679328

Gorgoso-Varela, J.J.
University of Oviedo, Mieres, Spain

ABSTRACT

The relative performance of any distribution function truly depends on the estimation methods and where this is wrongly chosen poor fit is inevitable. This may mislead forest managers and thus thwart effort towards sustainable forest management. This study therefore compared estimation methods for fitting 3-parameter Weibull distribution to the natural stand of Oluwa Forest Reserve, Ondo State, Nigeria with a view to enhancing sustainable management of the tree resources. Systematic sampling technique was used in the laying of eight (8) temporary sample plots (TSPs) of size 50m x 50m in the natural forest. Three fitting methods were used that based on maximum likelihood, moments and percentile. Comparison was based on Kolmogorov-Smirnov statistic (K-S), bias, mean absolute error (MAE) and mean square error (MSE). The result revealed that maximum likelihood method was more accurate in fitting the Weibull distribution to the natural stand. It had the smallest mean bias and MSE values of 0.00009 and 0.00021, respectively. Maximum likelihood method is therefore recommended for fitting the 3-parameter Weibull distribution to natural stand of the reserve.

Keywords: Weibull distribution, maximum likelihood, moments, percentile, natural stand

INTRODUCTION

Tree diameter characterisation using probability distribution functions is essential for determining the structure of forest stands. This has been an intrinsic part of forest management planning, decision-making and research in recent times. The distribution of species and tree size in a forest area gives the structure of the stand. Usually, it is as a result of the growth habit of the species, environmental condition and the management practices under which the species has developed.

A number of diameter distribution models for describing the structure of forest stands have been published. The beta function (Gorgoso*et al*., 2008, 2012; Ogana*et al*., 2015), Johnson S_B function (Knoebel and Burkhart, 1991) and the Weibull function (Bailey and Dell, 1973; Zhang *et al*., 2003; Palahi*et al.,* 2007; Ajayi*et al*., 2013) are the most commonly used distribution functions in quantitative forest studies in Nigeria and other part of the world at large. The Weibull distribution has gained prominence because of the simplicity in estimating itsparameters and it flexibility in fitting wide varieties of unimodal shapes. More so, several studies have shown that the Weibull is more appropriate for estimating tree diameter distribution in many cases (e.g. Gorgoso*et al*., 2012; Ogana*et al*., 2015).

Generally, the parameters of these distributions are estimated by maximum likelihood, moments or percentile method. Comparison of these estimation methods for fitting Weibull and other distribution to forest plantation have been adequately researched. For example, Shiver (1988) found that maximum likelihood estimation had the best fit for 3-parameter Weibull distribution. In Ghana, Nanang (1998) reported that Weibull distribution fitted with moment was more appropriate for mixed age group. Such study is yet to receive significant leap in natural forest, this may be due to the complex nature of the tropical forest which is characterised by diverse species composition and indeterminate age structure. Estimation method chosen in preference of a study may be inappropriate and misleading; as such valuable information on the forest stand structure may be wasted. It is therefore necessary to compare parameter estimation methods for fitting Weibull distribution to natural forest data in other to have reliable inventory and thus facilitate productive and sustainable management of tree resources.

The Study Area

This study was carried out in Oluwa Forest Reserve located in the moist tropical rainforest zone of Nigeria. It occupies an area of about 629km^2 with much of it lying approximately between 300 and 600m above sea level (Ogunjemiteet *al.,* 2006). The natural forest covers about 8km^2 (approximately 800ha) of the Forest Reserve. The Reserve is situated in Odigbo Local Government Area of Ondo State, Nigeria and lies between Latitude 6.83° - 6.91°N and Longitude 4.51° - 4.59°E (See Fig. 1 below). Annual rainfall ranges from 1700 to 2200 mm. Annual mean temperature in Oluwa is 26 °C. The relative humidity is high and uniform, ranging from 75% (afternoon) to 95% (morning). Soils are predominantly ferruginous tropical. The natural vegetation of the area is tropical rainforest characterised by emergent with multiple canopies and lianas.

METHODOLOGY

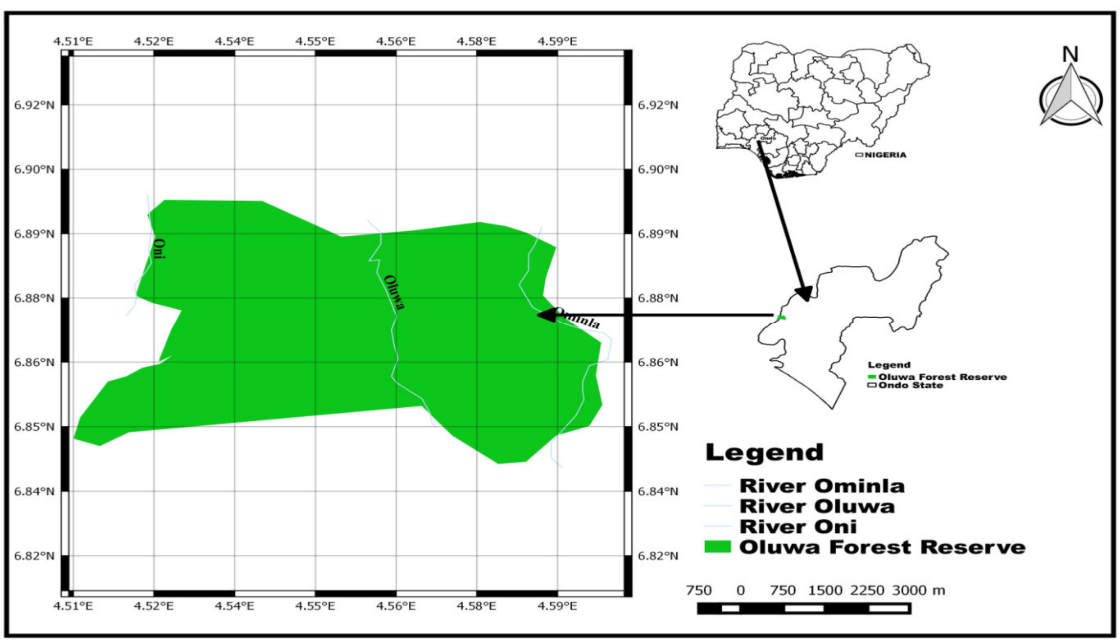


Fig. 1: Map of Oluwa Forest Reserve located in Ondo State, Nigeria (Source: Ogana, 2015)

Sampling Procedure, Data Collection and Processing

In this study, systematic sampling technique was used in the laying of the temporary sample plots (TSPs) in the 8km^2 natural forest. Two transects of 500m in length with a distance of 200m between the two parallel transects were laid. Sample plots of 50m x 50m in size were established in alternate position along each transect at 100m interval; summing up to 4 sample plots per 500m transect and a total of 8 sample plots in the study area (see Fig. 2). Living trees with Dbh ≥10.0cm in the selected plots were measured. The data collected were grouped into species and families, and the following stand variables were computed from the inventory data: mean diameter, minimum diameter, maximum diameter, number of trees per hectare and basal area. The summary statistics of the dataset used for this study are presented in Table 1.

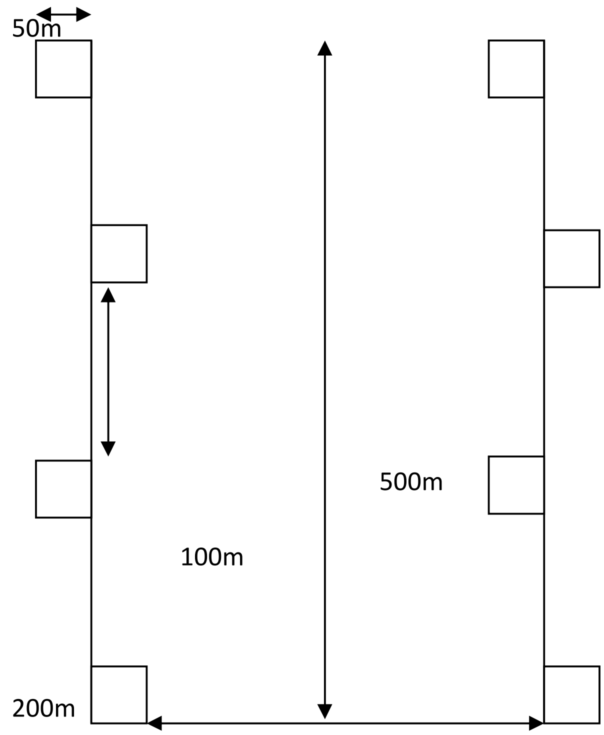


Fig 2: Plot layout with systematic line transects sampling technique

Table 1. Summary statistics of the data from the sample plots

	Statistics			
Variables	**Mean**	**Maximum**	**Minimum**	**Standard deviation**
No of Species = 58				
No of Family = 26				
Dbh (cm)	24.7	118.5	10.0	16.2
Basal area (m^2/ha)	18.28	31.72	8.12	7.31
Density (tree/ha)	267.5	352.0	196	60.0
Dominant Ht (m)	33.3	46.9	26.7	8.3

The Weibull function

The 3-parameters Weibull distribution (Weibull 1951) was used for this study. It is expressed as:

$$f(x) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} exp\left[-\left(\frac{x-a}{b}\right)^{c}\right] \quad \text{Eq. (1)}$$

Where: x = tree diameter, a, b and c are the location, scale and shape parameters of the distribution respectively. The location parameter of the Weibull distribution was taken as the minimum inventoried diameter per plot.The Weibull cumulative distribution function is obtained by integrating its density function in equation 1 above:

$$F(x) = \int_0^x \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} exp\left[-\left(\frac{x-a}{b}\right)^c\right] dx \quad \text{Eq. (2)}$$

$$= 1 - exp\left[-\left(\frac{x-a}{b}\right)^c\right] \quad \text{Eq (3)}$$

Where: F(*x*) is the cumulative distribution function.

2.4 Fitting Methods

Three methods of estimating the 3-parameter Weibull distribution were compared in this study. This include: maximum likelihood, moments and percentile methods.

2.4.1 *Maximum Likelihood (ML)*

The maximum likelihood estimation method used by Nanos and Montero (2002) and Gorgosoet *al*., (2012) was used. The distribution parameters were calculated with the following equations:

$$\frac{\sum_{i=0}^{n}(x_i)^c \ln(x_i - a)}{\sum_{i=1}^{n}(x_i - a)^c} - \frac{1}{c} = \frac{1}{n}\sum_{i=1}^{n}(x_i - a) \quad Eq\ (4)$$

$$b = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i - a)^c\right)^{\frac{1}{c}} \quad Eq\ (5)$$

Where: *n* is the number of sample observation and *x* is the diameter of the tree. The LIFEREG procedure in SAS/STAT™ (SAS Institute Inc., 2001) was used to estimate the shape and scale parameters.

Method of Moments

The method of moment used by Stankova and Zlatanov, (2010); Gorgosoet *al.* (2012) and Oganaet *al.* (2015) was used to estimate the Weibull parameters. It is based on the relationship between the parameters and the first and second moment of the diameter distribution (i.e. arithmetic mean diameter and variance, respectively). Expressed as:

$$b = \frac{\bar{d} - a}{\Gamma\left(1 + \frac{1}{c}\right)} \quad \text{Eq (6)}$$

$$\sigma^2 = \frac{(\bar{d} - a)^2}{\Gamma^2\left(1 + \frac{1}{c}\right)}\left[\Gamma\left(1 + \frac{2}{c}\right) - \Gamma^2\left(1 + \frac{1}{c}\right)\right] \quad \text{Eq. (7)}$$

Where: *a* which is the location parameter was taken as the smallest diameter of the plot, *d* is the arithmetic mean diameter of the distribution, σ^2 is the variance and Γ*(i)* is the Gamma function.

2.4.3 *Method of Percentile*

The Dubey (1967) percentile method was used to estimate the parameters of the Weibull distribution. The values of the parameters were computed with the following expressions:

$$\ln\hat{b} = \frac{\ln P_r - \frac{\ln P_t \cdot \ln(-\ln(1-r))}{\ln(-\ln(1-t))}}{1 - \frac{\ln(-\ln(1-r))}{\ln(-\ln(1-t))}} \quad \text{Eq (8)}$$

$$\hat{c} = \frac{\ln\left[\frac{\ln(1-r)}{\ln(1-t)}\right]}{\ln\left(\frac{P_r - a}{P_t - a}\right)} \quad \text{Eq (9)}$$

The proposed values of r = 0.97 and t = 0.17 by Dubey (1967) were used in this study.

Method comparison

The consistency of the three methods was assessed by Kolmogorov-Smirnov (KS), bias, mean absolute error (MAE), and mean square error (MSE), with the following equations:

Kolmogorov-Smirnov (KS) test:this was used to compare the cumulative estimated frequency with the observed frequency. The most striking difference between the two distributions was the $D_n$statistic value of the KS test:

$$D_n = Supx|F(x_i) - F_0(x_i)| \qquad \text{Eq.} \quad (10)$$

Where:Supx is the supremum value, $F(x_i)$ is the cumulative frequency distribution observed for the sample x_i (i = 1, 2,..., n)

$F_0(x_i)$ is the probability of the theoretical cumulative frequency distribution.Diameter classes of 1cm intervals were selected.

Bias:

$$Bias = \frac{\sum_{i=1}^{N} Y_i - \hat{Y}_i}{N} \qquad \text{Eq.} \quad (11)$$

Mean Absolute Error (MAE):

$$MAE = \frac{\sum_{i=1}^{N} |Y_i - \hat{Y}_i|}{N} \qquad \text{Eq.} \quad (12)$$

Mean Square Error (MSE):

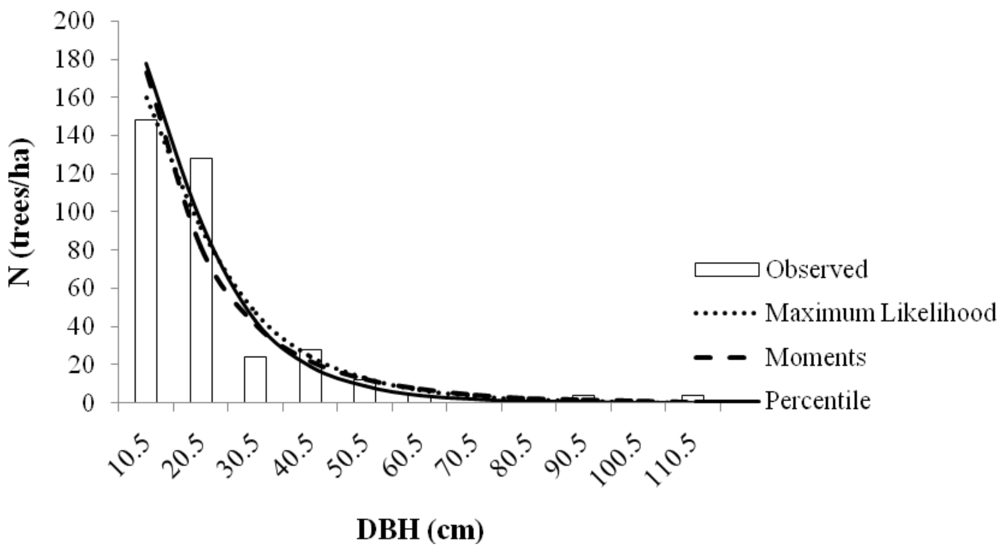
$$MSE = \frac{\sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2}{N} \qquad \text{Eq.} \quad (13)$$

Where: Y_i is the observed value, $\ddot{Y}_i$ is the theoretical value predicted by the model and N is the number of data points.

The bias, mean absolute error (MAE) and mean square error (MSE) were computed for each fit in mean relative frequency of trees per one for all diameter classes and plots.

RESULT

The comparison of maximum likelihood, moments and percentile methods for fitting the 3-parameter Weibull distribution to the natural stand data of Oluwa Forest Reserve have been made and the results are shown below. Graphical analyses of the observed frequency of trees(trees/ha) and the predicted frequency by Weibull distribution was no doubt typical of a natural forest, where a larger proportion of trees are found in the smallest diameter classes with decreasing frequency as the diameter increases; given rise to reverse J-shaped structure (see Fig. 3). The expected frequency of trees produced by3-parameter Weibull distribution fitted with maximum likelihood, moments and percentile methods showed slight variation with the observed diameter distribution; as the three fitting methods predicted larger values than the observed distribution for the smaller diameter classes of 20.5cm and 30.5cm.



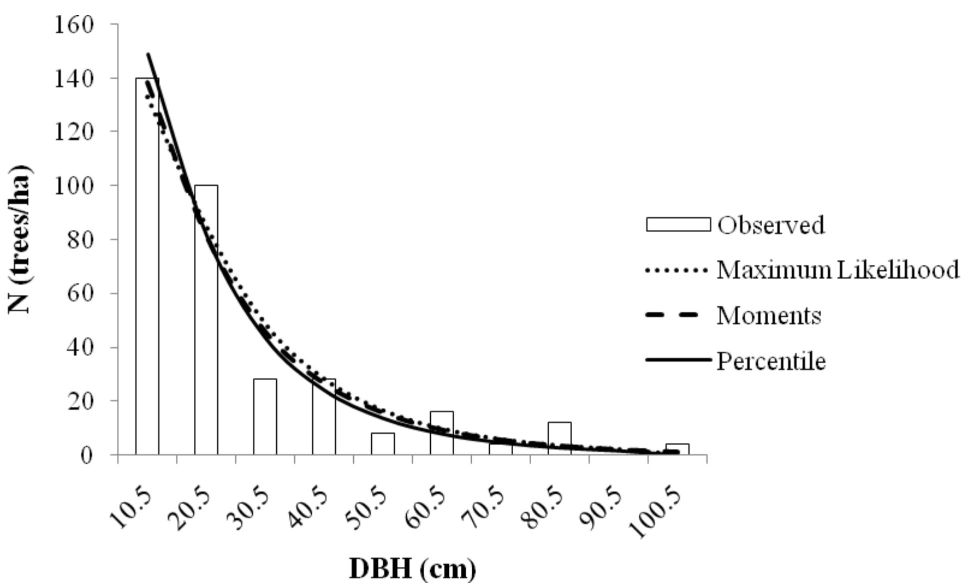


Fig. 3: Observed diameter distributions, fitted 3-parameter Weibull distribution by maximum likelihood, moments and percentile methods in number of trees per ha for two plots.

The overall ranking in terms of mean values of bias, mean absolute error (MAE), mean square error (MSE) and Kolmogorov-Smirnov (K-S) statistic summarizes the overall accuracy of the fitting methods as comparison criteria (Table 2). The results showed that there were little or no variation in the fitting methods (i.e. maximum likelihood, moments and percentile) considered in this study. Nevertheless, maximum likelihood method was more consistent than moments and percentile based on the result of the goodness-of-fit statistics; as such, ranked best.

The maximum likelihood method had the smallest mean value of bias of 0.00009; this was followed by moments and percentile methods, respectively. In the case of mean absolute error, moments had the smallest mean value of 0.00847, while percentile and maximum likelihood had 0.00859 and 0.00872, respectively. Also, maximum likelihood had the smallest mean value of mean square error of 0.00021, whereas moments and percentile methods had the same values of 0.00022. However, the Kolmogorov-Smirnov statistics showed that percentile had the smallest value of 0.11132, this was followed by moments with 0.11449; and lastly, by maximum likelihood with 0.1435.

Table 2. Mean values of bias, mean absolute error, mean square error in number of trees per one and K-S test for the three fitting methods for 3-paramter Weibull distribution

Fitting method	**Bias**	**MAE**	**MSE**	**K-S (D_n)**
Maximum likelihood	0.00009	0.00872	0.00021	0.14354
				[0.01911]
Moments	0.00015	0.00847	0.00022	0.11449
				[0.04639]
Percentiles	0.00018	0.00859	0.00022	0.11132
				[0.04272]

Standard deviation is enclosed in square brackets

The values of bias and MSE for each diameter class obtained with the three methods of fitting are shown in Fig. 4 and 5, respectively. It can be observed from Fig. 4 that the lower diameter classes were characterised by high values of bias up to a diameter class of 24.5cm with maximum likelihood, moments and percentile methods, and then with sudden decreased in values which subsequently became a bit stabilized.

Similarly, the maximum likelihood, moments and percentile method provided high MSE values up to 26.5cm and thereafter became stabilized as the diameter increased (see Fig. 5).

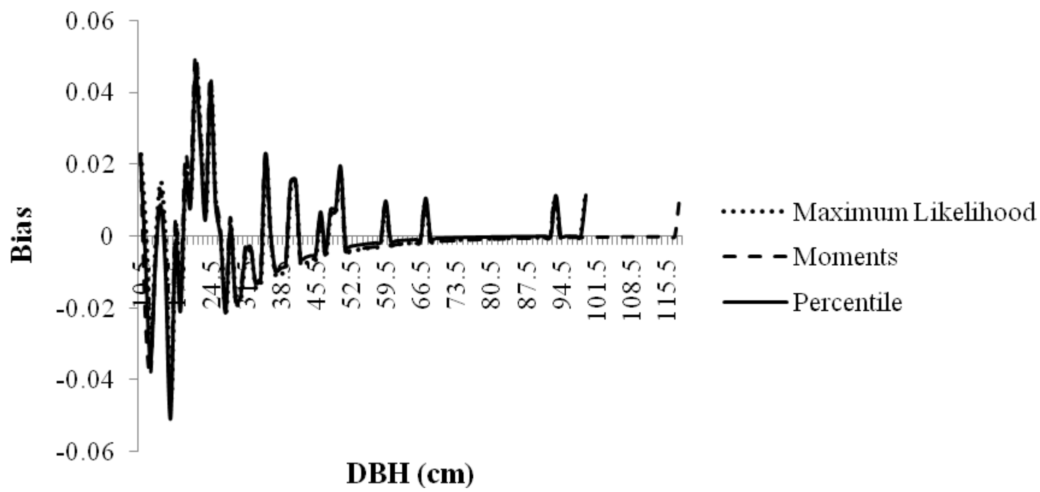


Fig. 4: Mean values of bias in number of trees per one in each diameter class obtained by three fitting methods of the 3-parameter Weibull distribution.

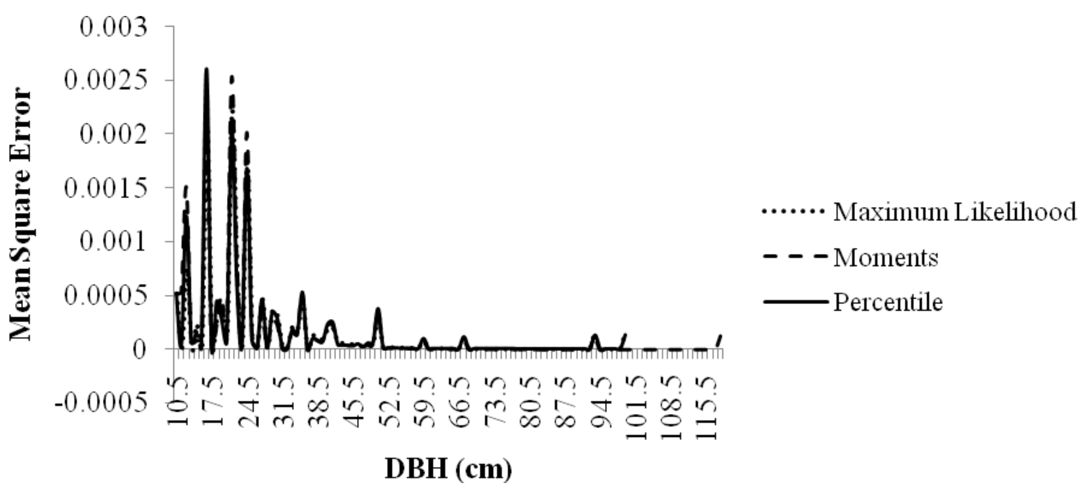


Fig. 5: Mean values of mean square error (MSE) in number of trees per one in each diameter class obtained by three fitting methods of the 3-parameter Weibull distribution.

DISCUSSION

The effectiveness of maximum likelihood, moments and percentile methods were compared in pursuant of the best estimation method that could fit the 3-parameter Weibull distribution to the natural stand data. The assessment of the fitting performance as adjudged by Kolmogorov-Smirnov, bias, mean absolute error and mean square error revealed that the estimation methods were appropriate in fitting the Weibull distribution to the data. This implies that any of the three fitting methods can be used to fit the 3-parameter Weibull distribution to the natural forest data. But maximum likelihood method was more consistent than moments and percentile methodsas revealed by its smallest values of mean bias and mean square error. The method of moments and percentile were slightly better than maximum likelihood in terms of MAE and K-S values.

This study is much in tandem withShiver (1988) who reported that maximum likelihood was more accurate than moments and percentile methods for fitting the 3-parameter Weibull distribution to diameter in unthinned slash pine plantation.However, he concluded that if the estimated distribution is to have less than 10% error in any one class, the approximate number of sample trees need is 50. This make maximum likelihood method most suitable as the sampled trees used in this study was more than the proposed minimum value by Shiver.Similarly, Zhang *et al*. (2003) obtained better results with maximum likelihood method than moments and percentile for fitting the 3-parameter Weibull distribution to mixed spruce-fir stand in northeastern North America. However, Nanang (1998) reported that the method of moments was appropriate for fitting the Weibull distribution to *Azadirachtaindica*plantation in Ghana.Gorgoso*et al*. (2007) and Carretero and first quadrant, 7 in the second, 8 in the third, result with non-linear regression approach than maximum likelihood, moments and percentile methods. However, the non-linear regression was not used in this study; as such warrant further research.

In conclusion, the maximum likelihood, moments and percentile methods performed creditably well in fitting the 3-parameter Weibull distribution to the natural forest data. However, the complexity of estimation varies with the percentile method exhibiting more simplicity. Thus, when simplicity, vis-à-vis the ease of estimation is the key focus of fitting the Weibull distribution, percentile method can be viewed as a workable tool to be used. But maximum likelihood method ranked best in this study; as such we recommend it for fitting the Weibull distribution to the natural forest stand of Oluwa Forest Reserve.

REFERENCES

Ajayi S. 2013. Diameter distribution for *Gmelinaarborea*plantations in Ukpon River Forest Reserve, Cross River State, Nigeria, AFRREV STECH, 2(1): 64-82.

Bailey R.L,and Dell T.R. 1973. Quantifying diameter distributions with the Weibull function, *For. Sci*. 19:97-104.

Carretero A.C.and Alvarez E.T. 2013. Modelling diameter distributions of *Quercussuber*L. Standsin “Los Alcornocales” Natural Park (Cádiz-Málaga, Spain) by usingthe two parameter Weibull functions.*Forest Systems,*22(1): 15-24

Dubey S.D.1967. Some percentile estimators for Weibull parameters.*Technometrics* 9: 119-129

Gorgoso J.J, Rojo A, Camara-Obregon A. andDieguez-Aranda U. 2012. A comparison of estimation methods for fitting Weibull, Johnson’s S_B and beta functions to *Pinuspinaster, Pinus radiate* and *Pinussylvestris*stands in northwest Spain. *Forest Systems,* 21(3): 446-459.

Gorgoso-Varela J.J, Rojo-Alboreca A, Afif-Khouri E. and Barrio-Anta M.2008.Modelling diameter distributions of birch (*Betulaalba*L.) and Pedunculate oak (*Quercusrobur*L.) stands in northwest Spain with the beta distribution. Investigación Agraria: *Sistemasy Recursos Forestales* 17(3): 271-281.

Gorgoso J.J, Álvarez-González J.G, Rojo A. andGrandas-arias J.A. 2007. Modelling diameter distributions of *Betulaalba*L. stands in northwest Spain with the two parameter Weibull function. *Invest. Agrar.:Sist. Recur. For.*16(2): 113-123.

Knoebel B.R, Burkhart H.E. andBeck D.E.1986. A growth and yield of thinned yellow poplar, *For. Sci*. *Monograph* 27, 41 p.

Nanang D.M. 1998. Suitability of the Normal, Log-normal and Weibull distributions for fitting diameter distributions of neem plantations in Northern Ghana.*For Ecol Manage*. 103: 1-7.

Nanos N. and Montero G. 2002.Spatial prediction of diameter distribution models.*Forest Ecology and Management*, 161: 147-158.

Ogana F.N, Osho J.S.A. andGorgoso-Varela J.J. 2015.Comparison of beta, Gamma and Weibull distributions for characterising tree diameter in Oluwa Forest Reserve, Ondo State, Nigeria.*Journal of Natural Sciences Research*, 5(4): 28-36.

Ogunjemite B.G, Agbelusi E.A, Afolaya, T.A. andOnadeko S.A. 2006."Status Survey of Chimpanzee in the Forest Zone of Southwestern Nigeria", *Acta Zoological Sinica*, 52(4): 1009–1014.

Palahí M, Pukkala T, Blasco E. andTrasobares A. 2007. Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (north-east of Spain). *Eur. J. Forest Res* 126: 563-571.

SAS 2004.SAS Online Doc 9.1.3.SAS Institute Inc., Cary, North Carolina, USA.http://support.sas.com/onlinedoc/913/docMainpage.jsp

SAS Institute Inc., 2001. SAS/STATTM User's Guide, Version 8.2. Cary, North Carolina.

Shiver B.D. 1988. Sample sizes and estimation methods for the Weibull distribution for unthinned slash pine plantations diameter distributions.*For. Sci.* 34 (3): 809–814.

Stankova T.V. andZlatanov TM.2010.Modeling diameter distribution of Austrian black pine (PinusnigraArn.) plantation: a comparison of the Weibull frequency distribution function and percentile-based projection methods, *Eur. J. Forest Res,* 129: 1169-1179

Weibull, W. (1951).A Statistical distribution function of wide applicability. *J. Appl. Mech*. 18: 293-297

Zhang L, Kevin C.P. andChuanmin L. 2003. A comparison of estimation methods for fitting Weibull and Johnson's SB distributions to mixed spruce-fir stands in north-eastern North America*. Can. J. For. Res*. 33: 1340-1347.