International Journal of Advanced Statistics and Probability, 10 (1) (2022) 1-8

International Journal of Advanced Statistics and Probability

Website: www.sciencepubco.com/index.php/IJASP

Research paper

Modeling environmental tree species' volume using some selected skewed distributions

F. A. Hassan [1]*, H. O. Ilo[1], N. O. Afolabi[1], K. A. Akintola[2]

[1]Department of Statistics, Ogun State Institute of Technology, Igbesa, Ogun State
[2]Department of Statistics, Oyo State College of Agriculture and Technology, Igbo-ora, Oyo State.
**Corresponding author E-mail:hasfat78@gmail.com*

Abstract

Ecological requirements' knowledge in determining tree species' distributions is a precondition for sustainable forest management. Tree species in all regions are threatened by climate change but some are more vulnerable than others. Rightly skewed distributions were used to take care of the environmental data set obtained from FRIN using five species which are: Beech Wood, White Afara, Opepe, Afon and Teak.
Appropriate statistical tools distributions like descriptive analysis, Akaike Information Criterion (AIC), Goodness of Fit, Probability Functions, Kolmogorov Smirnovt-test, Gamma, Weibull, Log-normal and Beta-weibull were carried out to determine the best distribution for each selected specie in this research.
It was established from the distributional pattern of the tree species' volume used in this research that the Gamma distribution was a better fit for the Beech wood with the AIC of 617.21, Beta-weibull distribution was a better _t for the White Afara, Opepe and Teak Species with the AIC values of 580.772, 630.84 and 733.60 respectively while the Weibull distribution was a better fit for the Afon specie with the AIC value of 752.07.
Conclusively, the implication of the analysis is that the Beta-weibull distribution described the tree species' volume best among the four distributions used. In line with the findings, it is recommended that Beta-weibull should be the appropriate distribution to model forest specie when it comes to modelling because of its four (4) parameters and its goodness when fitting.

Keywords:*Ecology; Environmental Statistics; Tree Species; Distribution; Goodness of Fit (GOF).*

1. Introduction

Tree species in all regions are threatened by climate change but some are more vulnerable than others. It is therefore expected that predicted future climate change will have a significant impact on the distribution of species, which will be continuous, cumulative, and interactive (Pearson and Dawson, 2003; Williamson et al., 2009). Climate change will create changes in microclimates, local site conditions, disturbances (such as fire, insects and disease attacks, drought, extreme storms), phenology (i.e. the timing of biological activity over a year in relation to climate), and the diversity, distribution, abundance, and ecosystem interactions of species, all of which could lead to increased tree mortality and changes in competitive interrelationships (including the potential for the introduction of exotics). Tree species and genotypes will acclimatize, adapt, and migrate; however, in many cases, the rate and magnitude of future climate change may significantly exceed the ability of tree species to naturally adjust (Johnston et al., 2009; Shugart et al., 2003).
In Nigeria, there are over 66 research institutes majority of which are established and funded by the federal government. These research institutes were established with the aim at discovering better ways of doing things, improving the knowledge base and raising standards in specific fields. These research institutes also have mandate to make research findings, technology and knowledge adaptable to local Nigeria situations. These are several open grown trees within these institutes which offer several benefits such as beautification, reduction of urban heat and cooling, reduction of storm water run-off, reduction of air pollution, reduction of energy costs through increased shade over buildings, enhancement of property values and improvement of the overall urban environment impact. Wilson, (2011).
About 146 species on the International Union for Conservation of Nature (IUCN) list of threatened species are found in Nigeria out of which 18 fall under the category 'endangered' and under the category 'critically endangered' (Federal Ministry of Environment, 2006). However, there is a wide gap in the knowledge of the genetic diversity of tree species owing to lack of up-to-date documentation of plant genetic resources. The diversity of tree species is decreasing due to the rate of habitat destruction and over-exploitation which are far greater than the rate of genetic diversity collection and conservation (National Centre for Genetic Resources and Biotechnology, 2008).

Copyright ©F. A. Hassan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

At University of Ibadan, there is renewed interest in the heterogeneity of tree species composition, structure (Adeyemi and Adesoye, 2012; Onefeli et al., 2012) and the increasing demand for both timber and non-timber products has necessitated the need for the conservation of tree genetic resources as the wild population of most of these vulnerable tree species have been depleted in natural forests. Therefore, this study aimed at modelling Tree forest using some selected distributions.

Hence, this research seeks to know the most suitable distribution for environmental forest species' volume in order to fit the rightly skewed distributions to the data, to estimate the parameters of the distributions, to predict the future volume of the tree species and to use the model selection criterion to pick the best model among the distributions.

1.1. Concept of trees' measurements (plantation)

Measuring individual Trees or Logs or Forest as a whole is an essential part of any Forest management. Measuring the amount of wood present in tree forest is imperative for sales and purchases of wood in deciding kind and degree of management practices to be adopted, to know the effect of a particular treatment on either increase or decrease in volume/height/diameter and also in predicting amount of return which a piece of land can give in future. This measurement is a concern for not only the Foresters, but also for contractors, sawyers, Transporters, wood using public and Researchers.

2. Literature review

According to Olajuyigbe et al., (2015); Trees, which are important for the sustenance of life and the health of our planet, are disappearing at an alarming rate. Consequently, the need for actions to develop effective strategies to conserve them is receiving considerable attention worldwide. Forest genetic resources are fast becoming depleted in most natural forests due to the pressures of deforestation, urbanization, poor management and a regeneration programme that is virtually non-existent. In Nigeria, the impacts of climate change will further aggravate the plights of many indigenous and exotic tree species as climatic variability may limit the ability of forest trees to quickly adapt to the changing climate. The huge presence of various indigenous and exotic tree species on the University of Ibadan campus and the fact that some of these trees are no longer found in most natural forests underscores the potentials of the campus as an important live gene bank. There is little or no information on the taxonomy, diversity and growth characteristics of many of the trees on campus. This information is very important for their conservation and sustainable management. There is therefore, an urgent need for their identification, conservation and management.

Gauss et al., (2008) introduced a generalization of the Weibull distribution, termed the beta Weibull distribution. According to them, the Weibull distribution has been extensively used over the past decades for modeling data in reliability engineering and biological studies, and as it represents only a special case of the new distribution, we hope that this generalization shall attract wide application. They provided a comprehensive treatment of the mathematical properties of the beta Weibull distribution and derive expressions for its moment generating function and the rth generalized moment. They also discussed maximum likelihood estimation and provide formulae for the elements of the Fisher information matrix. We also demonstrate its usefulness on a real data set.

Boikanyo et al., (2018) developed a new family of generalized distributions called the beta Weibull-G (BWG) distribution is proposed and developed. This new class of distributions has several new and well known distributions including exponentiated-G, Weibull-G, Rayleigh- G, exponential-G, beta exponential-G, beta Rayleigh-G, beta Rayleigh exponential, beta exponential-exponential, Weibulllog-logistic distributions, as well as several other distributions such as beta Weibull-Uniform, beta Rayleigh-Uniform, beta exponential-Uniform, beta Weibull-log logistic and beta Weibull-exponential distributions as special cases. Series expansion of the density function, hazard function, moments, mean deviations, Lorenz and Bonferroni curves, Renyi entropy, distribution of order statistics and maximum likelihood estimates of the model parameters are given. Application of the model to real data set was presented to illustrate the importance and usefulness of the special case beta Weibull-log-logistic distribution.

Oluyede, P. et al., (2018) introduced the gamma Weibull-G family of distributions by combining the gamma generator with the Weibull-G family of distributions which was defined by Bourguignon et al., (2014). According to them, this new class of distributions is exile in accommodating all forms of hazard rate functions and contains several well-known and new sub-models such as Weibull, Rayleigh, exponential, modified Weibull, gamma-modified Weibull, gamma modified exponential, gamma-Weibull and gamma-Rayleigh distributions.

3. Research methodology

This section elicits on the methodologies used in tackling the set objectives. Several probability distribution approaches were engaged in providing answers to certain questions and hypothesis formulated as follows with the derivations of tree species' volume:

3.1. Volume determination of individual trees

Determination of Volume of Logs

For calculation of logs (portion of log which has been cut into different sizes depend upon tapers or the need of the length) following formulas are in use:

- If the log is almost cylindrical.
- If the log has some taper then its volume can be obtained by taking the basal area at 1 or 2 more points.

Hubers formula;

$$V = S_m \times L \tag{1}$$

Where

V = The Volume of tree.

S_m = Basal Area at the middle of the log.

L = The Tree length.

Smalians formula;

$$V = \frac{(S_1 + S_2) \times L}{2} \quad (2)$$

Where
V = The volume of trees
S_1 and S_2 are the Basal areas at the two end points of the log.
L = The length of the tree.

Newtons formula;

$$V = \frac{(S_1 + 4S_m + S_2) \times L}{6} \quad (3)$$

Where
V = The volume of the tree
S_1 and S_2 = The basal areas at the two end points of the log
4Sm = Basal areas at four different points in between the edges of the log.
Also
Newtons formula gives the best and more accurate volume but it is also cumbersome. To have more accurate results, Newtons formula can be further modified instead of taking only one middle point and multiplying it by four, one can take basal areas at four different points between S_1 and S_2.

The Modified Newtons formula;

$$V = \frac{\pi H\,(Db^2 + 4Dm^2 + Dt^2)}{24} \quad (4)$$

Where
V=volume
Db=Diameter of the tree base.
Dm= Diameter of the tree middle.
Dt= Diameter of the tree top.
H= The total height of the tree.
Also under measurement of diameter. To have a uniform standard measurement throughout the country, so that a result obtained after measurement at one site can be compared to those at other site in other part of the country. The diameters of the standing Tree are measured at 1.37m (4 feet 6 inches). This height is commonly called diameter at breast height, and abbreviated as DBH.

3.2. Descriptive statistics and explorative data analysis

Descriptive statistics is an aspect of Statistics that deals with describing, summarizing and explaining a data set using graphs, pictures and so on. It deals with techniques to describe and interpret statistic data, using numerical measures of center, location, and variation.
Arithmetic Mean;

$$Mean\left(\bar{x}\right) = \frac{\sum_{i=1}^{n} X_i}{n} \quad (5)$$

Median;

$$Median = L_m + \left[\frac{N/2 - Cf_m}{f_m}\right] C_m \quad (6)$$

Mode;

$$Mode = L_m + \left[\frac{d_1}{d_1 + d_2}\right] C_m \quad (7)$$

Standard Deviation (S.D);

$$S.D = \sqrt{\frac{\sum f\,(x - \bar{x})^2}{\sum f}}$$

$$OR \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2} \quad (8)$$

Coefficient of Variation (CV);

$$CV = \frac{\sigma}{\bar{x}} \times 100 \quad (9)$$

Skewness;

$$= \frac{Q_3 + Q_1 - 2\text{Median}}{Q_3 - Q_1} \quad (10)$$

Kurtosis

$$\beta_2 = \frac{\mu_4}{\mu_2{}^2} \quad (11)$$

3.3. Modelling the distributions of the forest specie

Gamma Distribution

The probability density function of gamma for the forest specie (x), with two parameters α and β is given by:

$$f(\text{x}, \alpha, \beta) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} \exp\left[-\frac{x}{\beta}\right]; 0 < x < \infty \quad (12)$$

Where x, α, β and Γ are the wind speed values, the shape parameter, rate or scale parameter and the gamma function respectively.

Weibull Distribution

The probability density function of weibull for the forest specie (x), with two parameters α and β is given by:

$$f(x, \alpha, \beta) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right]; 0 < x < \infty \quad (13)$$

Where α, β and x are the shape parameter, scale parameter and wind speed value respectively.

Log-Normal Distribution

The probability density function of log-normal the forest specie (x), with two parameters μ and σ is given by:

$$f(x, \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln(\text{x}) - \mu}{\sigma}\right)^2\right\}; 0 < x < \infty \quad (14)$$

Where σ, μ and v are the shape parameter, scale parameter and wind speed value respectively.

Beta - Weibull Distribution

The probability density function of beta-weibull for the forest specie (x), with four parameters a, b, α and β is given by:

$$f(x, \text{a}, \text{b}, \alpha, \beta) = \frac{\alpha\beta^{\alpha}}{B(a,b)} x^{\alpha-1} \exp\left\{-b(\beta x)^{\alpha}\right\}\left[1 - \exp\left\{-(\beta x)^{\alpha}\right\}\right]^{\text{a}-1}; 0 < x < \infty \quad (15)$$

Kolmogorov-Smirnov Test (Goodness of Fit)

$$D = \max\left\{\left| \text{F}_0(\text{x}) - \text{H}(\text{x}) \right|\right\} \forall x \quad (16)$$

4. Results and discussions

4.1. Descriptive statistics

Table 1: Preliminary Analysis of Forest Species

S/N	Specie	Min	Mean	Median	Max	SD	CV	Skew	Kurt
1	Beech Wood	3.46	63.77	53.62	288.29	53.84	0.84	1.98	8.26
2	White Afara	3.07	46.58	31.99	163.51	39.43	0.85	1.08	3.46
3	Opepe	3.86	70.76	49.31	317.01	59.34	0.84	1.64	6.52
4	Ofon	3.27	194.19	143.14	737.43	155.11	0.80	1.19	4.29
5	Teak	1.85	163.38	126.62	587.84	130.43	0.80	0.96	3.57

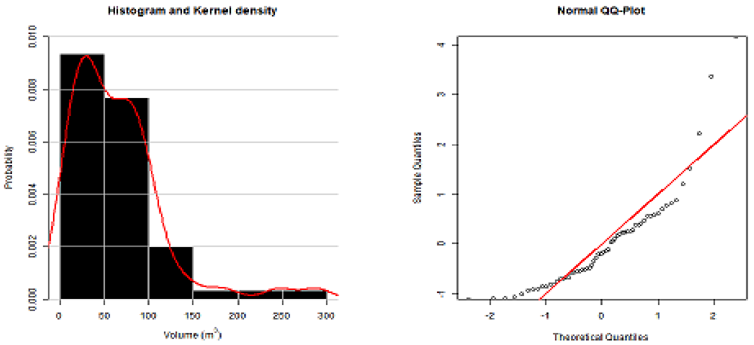

Fig. 1: Exploratory Analysis of Beech Wood.

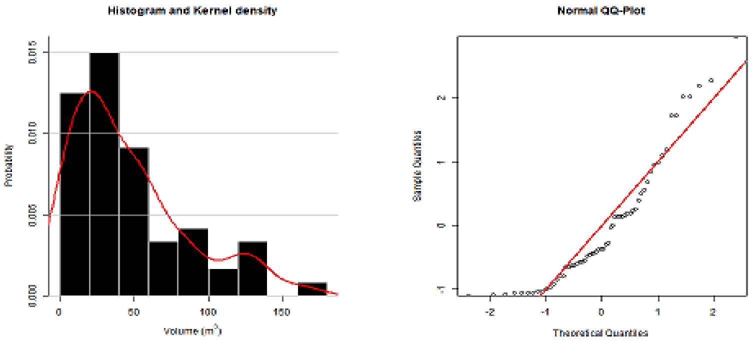

Fig. 2: Exploratory Analysis of White Afara.

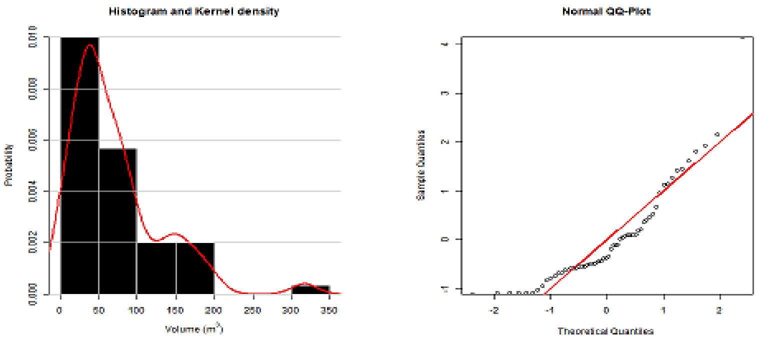

Fig. 3: Exploratory Analysis of Opepe.

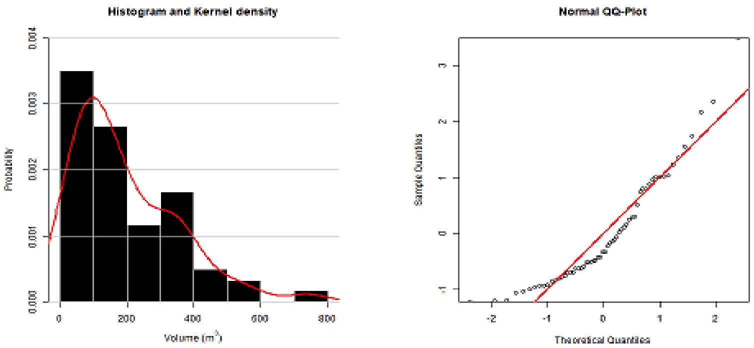

Fig. 4: Exploratory Analysis of Beech Afon.

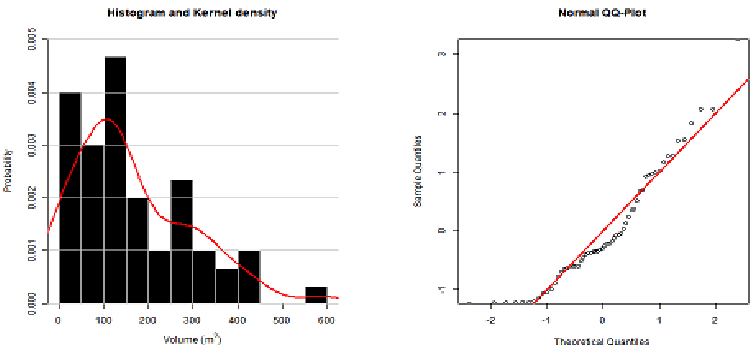

Fig. 5: Exploratory Analysis of Teak.

From the figures 1 to 5, the histogram and kernel probability distribution of forest specie volume revealed that the data exhibit significant skewness with long tails.

4.2. Parameter estimation and goodness of fit

Table 2: Parameter Estimates and Goodness-of-Fit Test for Each Distribution and Tree Types

	Distribution	Shape	Scale	Shape 2	Scale 2	D – stat	P – value	AIC	Fit
Beech Wood	Gamma	1.43	44.70			0.077	0.505	617.21	Yes
	Weibull	1.26	68.77			0.078	0.555	617.66	Yes
	Lognormal	3.79	0.95			0.110	0.159	623.31	Yes
	Beta-weibull	1.61	1.51	0.97	68.76	0.075	0.890	621.03	Yes
White Afara	Gamma	1.42	32.82			0.087	0.00074	583.34	No
	Weibull	1.17	49.16			0.071	0.003	582.79	No
	Lognormal	3.40	1.05			0.118	0.000009	588.55	No
	Beta-weibull	0.90	0.91	1.24	49.16	0.067	0.950	580.772	Yes
Opepe	Gamma	1.45	48.94			0.088	0.0008	630.86	No
	Weibull	1.24	75.88			0.087	0.0009	631.01	No
	Lognormal	3.87	0.99			0.140	0.00007	638.11	No
	Beta-weibull	1.29	1.25	1.07	75.87	0.088	0.740	630.84	Yes
Afon	Gamma	1.59	121.84			0.081	0.806	753.40	Yes
	Weibull	1.25	208.02			0.073	0.778	752.07	Yes
	Lognormal	4.87	2.06			0.090	0.916	765.81	Yes
	Beta-weibull	0.81	0.83	1.41	208.02	0.069	0.939	755.934	Yes
Teak	Gamma	1.60	102.38			0.109	9.798e-20	743.25	No
	Weibull	1.10	168.51			0.115	0.00003	734.78	No
	Lognormal	4.55	1.37			0.211	3.48e-08	757.77	No
	Beta-weibull	0.38	0.41	1.99	168.47	0.101	0.578	733.60	Yes

Table 3 above shows the values of the estimates for the parameters of each distribution and location. The parameters of the gamma distribution was estimated using the moment estimator; the weibull, lognormal and beta-weibull parameters were estimated using the method of maximum likelihood. Also, values for the kolmogorv-smirnov test statistic (D-stat) was obtained with the respective p-value. The Akaike Information Criterion (AIC) value was also obtained. The distribution whose p-value was greater than 0.05 was ascribed a good fit for the volume of tree for a particular tree type and the best fit was taken as one with the lowest AIC value.
It was observed that for the Beech wood tree volume, the Gamma distribution outperformed other considered distributions based on AIC; all distributions considered were god fit for Afon tree volume but was best described by the weibull distribution. The beta-weibull was best among the distributions considered to model the volume for White Afara, Opepe and Teak Trees.

4.3. Probability of exceedance and return period

Table 3: Probability of Exceedance and Return Period of the Forest Specie

S/N	Species	Distribution	Parameters	Max	P_e	R_p
1	Beech Wood	Gamma	(1.43, 44.70)	288.29	0.0042	237
2	White Afara	Beta-weibull	(0.90, 0.91, 1.24, 49.16)	163.51	0.016	63
3	Opepe	Beta-weibull	(1.29, 1.25, 1.07, 75.87)	317.01	0.0042	240
4	Afon	Weibull	(1.25, 208.02)	737.43	0.0077	130
5	Teak	Beta-weibull	(0.38, 0.41, 1.99, 168.47)	587.84	0.0041	245

Table 3 shows the probability that the volume of trees taken subsequently would exceed the maximum recorded volume for different Tree types. The best fit distributions were used in obtaining these results. The return periods $R_p$indicates the number of trees that need to be subsequently sampled before we record a volume greater than the maximum already recorded.
Probability Density and Return Level Plot
Figures 6 to 10 below shows the respective probability density and return plots for volume of trees for respective tree types. The best _t distribution was used in achieving these.

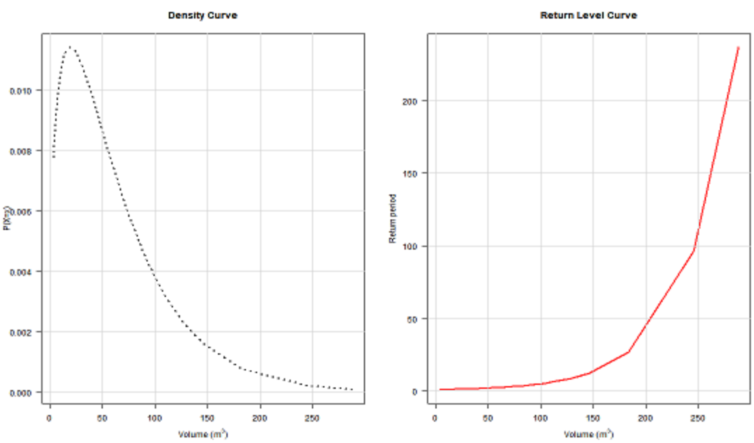

Fig. 6: Probability Density and Return Period Plot - Beech Wood.

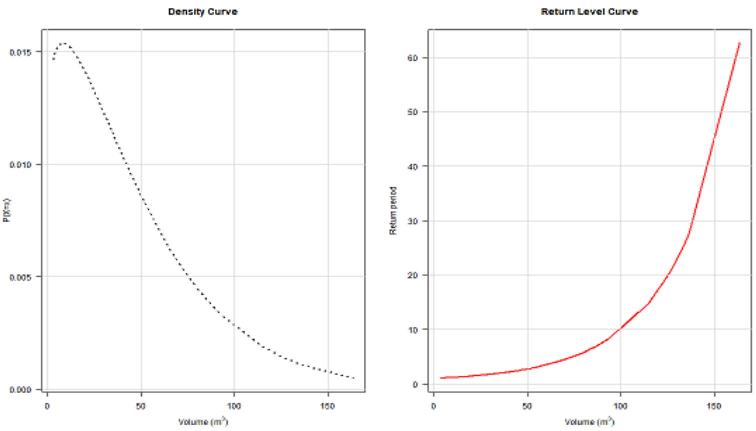

Fig. 7: Probability Density and Return Period Plot – White Afara.

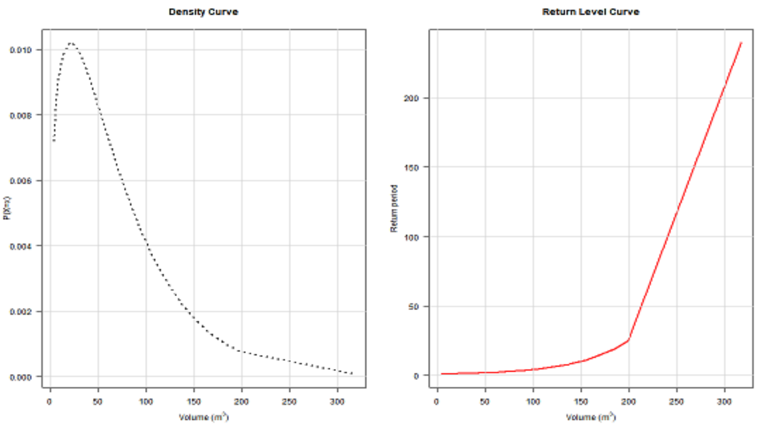

Fig. 8: Probability Density and Return Period Plot – Opepe.

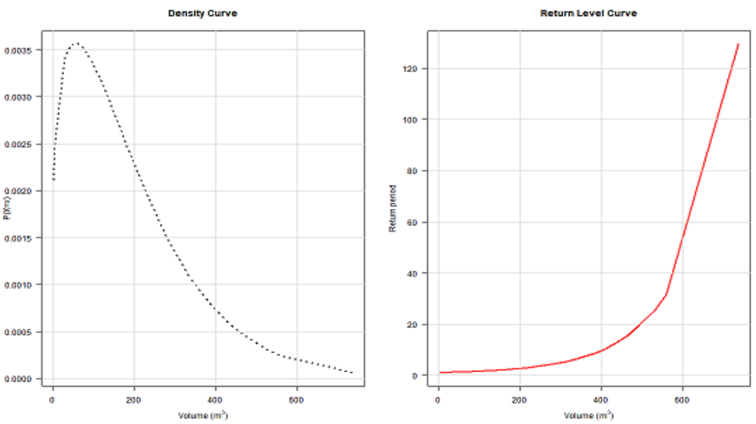

Fig. 9: Probability Density and Return Period Plot – Afon.

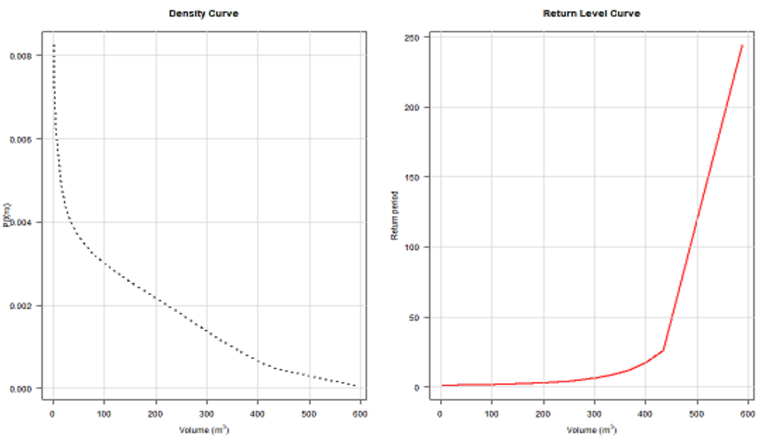

Fig. 10: Probability Density and Return Period Plot – Teak.

5. Concluding remarks

In this research, the EDA via histogram and QQ-plot indicated that the data exhibit significant skewness with long tails and that they do not follow the normal distribution which suggested that the volume is best described by a family of rightly skewed distributions.

From the distributional pattern of the tree species' volume of the five different types, it was established that the Gamma distribution was a better fit for the Beech wood with the AIC of 617.21, Beta-weibull distribution was a better _t for the White Afara, Opepe and Teak Species with the AIC values of 580.772, 630.84 and 733.60 respectively while the Weibull distribution was a better fit for the Afon specie with the AIC value of 752.07.

This implies that Beta-weibull distribution described the tree species' volume best among the four distributions used.

References

[1] Adeyemi, A. A. and Adesoye, P. O. (2012). Assessment of tree species in the Faculty of Agriculture and Forestry, University of Ibadan, Nigeria. In: Onyekwelu, J. C., Agbeja, B. O., Adekunle, V. A. J., Lameed, G. A., Adesoye, P. O. and Omole, A. O. (eds.). *De-reservation, Encroachment and Deforestation: Implications for the Future of Nigerian Forest Estate and Carbon Emission Reduction.* Proceedings of the 3rd Biennial National Conference of the Forests and Forest Products Society, Ibadan, Nigeria (3rd – 6th April, 2012): 56-66.

[2] Boikanyo M., Broderick O. O., Gofaone M., Shujiao H. and Adeniyi F. F. (2018). The Beta Weibull-G Family of Distributions: Model, Properties and Application

[3] Bourguignon, M., Silva, R. S., and Cordeiro, G. M. (2014). The Weibull-G Family of Distri-butions. Journal of Data Science, 12, 53-68. https://doi.org/10.6339/JDS.201401_12(1).0004.

[4] Federal Ministry of Environment, (2006). Nigeria First National Biodiversity Report, Federal Ministry of Environment, Abuja, Nigeria.

[5] Gauss M. C., Alexandre B. Simas and Borko D. S. (2010). The Beta Weibull Distribution.

[6] Johnston, M., Campagna, M., Gray, P., Kope, H., Loo, J., Ogden, A., O'Neill, G. A., Price, D. and Williamson, T. (2009). Vulnerability of Canada's tree species to climate change and management options for adaptation: An overview for policy makers and practitioners. Canadian Council of Forest Ministers. Pp. 12-21.

[7] National Centre for Genetic Resources and Biotechnology (NACGRAB), (2008). State of plant genetic resources for food and agriculture in Nigeria (1996-2008). Country Report Prepared by NACGRAB and Federal Department of Agriculture, Abuja Nigeria. Pp. 15-17.

[8] Obisesan K. O. and Dosumu (2016): On Comparing Parameter Estimation Procedures Using Rainfall Series. Science Focus Vol.21 (1) 2016 pp. 1-21 ISSN 1598-7026.

[9] Obisesan K. O., Lawal M., Bamiduro T.A. and Adelakun A.A. (2013): Data Visualization and Change Point-Detection in Environmental Data: The case of Water Pollution in Oyo State. Journal of Scienti_c Research 12:181-190.

[10] Olajuyigbe S. O., Akinyele A. O., Jimoh S. O. and Adegeye A. O. (2015). Tree species diversity in the Department of Forest Resources Management, University of Ibadan, Nigeria

[11] Oluyede, B. O., Pu, S., Makubate, B. and Qiu, Y. (2018). The Gamma Weibull-G Family of Distributions. Austrian Journal of Statistics.

[12] Onefeli, A. O., Isese, M. O. O. and Oluwayomi, T. I. (2012). Taxonomical classification and physical health assessment of avenue trees in the Faculty of Agriculture and Forestry, University of Ibadan, Nigeria. *Proceedings of the 3rd Biennial National Conference of the Forests and Forest Products Society, Ibadan, Nigeria* (3rd – 6th April, 2012): 75-84. https://doi.org/10.17713/ajs.v47i1.155.

[13] Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* 12: 361–371. https://doi.org/10.1046/j.1466-822X.2003.00042.x.

[14] Shugart, H. Sedjo, R and Sohngen, B. (2003). Forests and global climate change: potential impacts on US forest resources. Prepared for the Pew Centre on Global Climate Change. Pp. 1-24.

[15] Williamson, T. B., Colombo, S. J., Duinker, P. N., Gray, P. A., Hennessey, R. J., Houle, D., Johnston, M. H., Odgen, A. E. and Spittlehouse, D. L. (2009). Climate change and Canada's forests: from impacts to adaptation. Sustainable Forest Management Network and Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alta.

[16] Wilson, W. G. (2011). Constructed Climates: A primer on urban environments. *Chicago University and Parks Victoria.* https://doi.org/10.7208/chicago/9780226901473.001.0001.