Regular Article
pISSN: 2288-9744, eISSN: 2288-9752
Journal of Forest and Environmental Science
Vol. 34, No. 6, pp. 451-456, December, 2018
https://doi.org/10.7747/JFES.2018.34.6.451

JFES
Journal of Forest and Environmental Science

Application of Finite Mixture to Characterise Degraded *Gmelina arborea* Roxb Plantation in Omo Forest Reserve, Nigeria

Friday Nwabueze Ogana*
Department of Social and Environmental Forestry, University of Ibadan, Ibadan 200284, Nigeria

Abstract

The use of single component distribution to describe the irregular stand structure of degraded forest often lead to bias. Such biasness can be overcome by the application of finite mixture distribution. Therefore, in this study, finite mixture distribution was used to characterise the irregular stand structure of the *Gmelina arborea* plantation in Omo forest reserve. Thirty plots, ten each from the three stands established in 1984, 1990 and 2005 were used. The data were pooled per stand and fitted. Four finite mixture distributions including normal mixture, lognormal mixture, gamma mixture and Weibull mixture were considered. The method of maximum likelihood was used to fit the finite mixture distributions to the data. Model assessment was based on negative loglikelihood value (-$\Lambda\Lambda$), Akaike information criterion (AIC), Bayesian information criterion (BIC) and root mean square error (RMSE). The results showed that the mixture distributions provide accurate and precise characterisation of the irregular diameter distribution of the degraded *Gmelina arborea* stands. The -$\Lambda\Lambda$, AIC, BIC and RMSE values ranged from -715.233 to -348.375, 703.926 to 1433.588, 718.598 to 1451.334 and 3.003 to 7.492, respectively. Their performances were relatively the same. This approach can be used to describe other irregular forest stand structures, especially the multi-species forest.

Key Words: finite mixture distribution, gamma mixture, normal mixture, lognormal mixture, Weibull mixture.

Introduction

The diameters of trees, and mainly its structure characterise any forest stand very well. Diameter structure is a very significant stand characteristic based on which we can evaluate the growth and volume production of stand as well as the structure of assortments, maturity, etc. Diameter structure may be expressed by frequency of tree diameters that quantify their distribution in diameter classes (Petras et al. 2010). Several probability distribution functions have been used to describe the diameter and/or height distributions in natural and forest plantation (manmade) forests e.g. beta, exponential, gamma, Johnson's SB, logit-logistic, lognormal, normal, Weibull distributions etc.

Modelling the diameter distributions of pure monospecific even-aged stands is a simple issue and this might be achieved by using any of the aforesaid statistical distributions. However, the diameter structure of forest is often affected by disturbance caused by natural disaster (e.g. wind throw) and/or anthropogenic activities (e.g. thinning, illegal logging, bush burning etc.) (Tsogt and Lin 2012). This may result to irregular, bimodal or heavily skewed diameter distributions. Under such condition, a single univariate distribution function may not uniquely de-

Received: April 30, 2018. Revised: November 29, 2018. Accepted: December 4, 2018.

Corresponding author: Friday Nwabueze Ogana

Department of Social and Environmental Forestry, Faculty of Renewable Natural Resources, University of Ibadan, Rm 115 Ibadan 200284, Oyo State, Nigeria
Tel: +2347037679328, E-mail: ogana_fry@yahoo.com

scribe the diameter structure of the forest stand. The use of single component distribution may lead to oversimplification of the forest stand structure (Maltamo et al. 2000; Liu et al. 2002). To capture this irregularity, a finite mixture model has been recommended for modelling stand with two or more modes (bimodal or multimodal) diameter distributions (Bailey and Dell 1973; Zhang et al. 2001).

Finite mixture modelling is an approach applied in the classification of observations and to model unobserved heterogeneity. In this approach, the observed data are generally assumed to come from unobserved subpopulations called classes or groups; as such, mixtures of probability densities e.g. normal, Weibull etc. are used to model the variable of interest. One of the most reliable and preferred methods of fitting mixtures of distribution is the method of maximum likelihood through the expectation maximization (EM) algorithm. This algorithm has "reliable convergence properties, however, convergence can be slow if the component of the mixture are not well separated and the sample size is small" (Izenman and Sommer 1988).

The concept of finite mixture modelling was first used in tree diameter distributions by Zhang et al. (2001), Liu et al. (2002) and Zasada and Cieszewski (2005). Zhang et al. (2001) used a finite mixture of two Weibull functions to model the diameter distributions of uneven-aged stands. The authors reported a better result with mixture Weibull compared to a single Weibull and exponential distributions. Liu et al. (2002) also used a finite mixture model to characterise the diameter distributions of mixed-species forest stands. Zasada and Cieszewski (2005) showed the applicability of finite mixture distribution approach for characterising tree diameter distributions by natural social class in pure even-aged stand. Some of the recent studies on finite mixture modelling in tree diameter distribution include Zhang and Liu (2006), Podlaski (2010a, 2010b, 2017), Jaworski and Podlaski (2012), Tsogt and Lin (2012), Lui et al. (2014). In these studies, the authors found finite mixture modelling as a promising method for modelling irregular, bimodal or multimodal diameter distributions.

To date, most distribution studies have relied on the use of a single component distribution to model irregular stand structure of degraded forest. The *Gmelina arborea* stands in Omo forest reserve have suffered from severe illegal logging activities, bush burning and agricultural encroachment (Ogana et al. 2017). This has exposed the forest estate to wind-throw. In consequence, result to irregular stand structure. Therefore, the main objective of this study is to apply the finite mixture distributions to characterise the irregular diameter distribution of the degraded *Gmelina arborea* stands in Omo forest reserve, Nigeria.

Materials and Methods

Data

The data used for this study came from the *Gmelina arborea* plantation in Omo Forest Reserve. Its lies between Latitudes 6°35′ and 7°05′N and Longitudes 4°19′ and 4°10′E of Ogun State, Nigeria and occupies an area of 130,500 ha (Ogana et al. 2017). A total of 30 sample plots of 0.04 ha from three distinct stands established in 1984, 1990 and 2005 were used. Ten plots were laid in each stand. The descriptive statistics of the inventory data are presented in Table 1.

Mixture distribution

The finite mixture distribution for a random continuous variable x (e.g. tree diameter, height etc.) can be expressed as:

$$f(x,p)=\sum_{j=1}^{k}p_jf_j(x)=p_1f_{1j}(x)+p_2f_2(x)+\cdots+p_kf_k(x) \qquad (1)$$

Where $f_j(x)$ is the probability density function (pdf) of the jth individual component distribution, p is the mixing proportion of the individual components in the mixture distribution and must satisfy the condition: $0 \leq p_j \leq 1$ and

Table 1. Descriptive statistics of the data set

Stand	Tree variable	Statistics			
		Mean	Max	Min	Standard deviation
1984	Dbh (cm)	18.5	42.9	5.8	9.14
N=257	Height (m)	13.7	42.7	1.6	6.59
1990	Dbh (cm)	22.2	40.8	5.7	9.35
N=240	Height (m)	15.5	26.3	5.9	5.13
2005	Dbh (cm)	15.3	34.1	5.9	6.26
N=139	Height (m)	8.8	14.9	3.1	2.43

$p_k = \sum_{j=1}^{k} p_j$. In this study, the normal mixture, lognormal mixture, gamma mixture and Weibull mixture were used to characterise the irregular tree diameter distributions of the *Gmelina arborea* stands. The normal mixture consists of two component normal distributions. the lognormal mixture also has two component lognormal distributions. Similarly, the gamma mixture and Weibull mixture have two component gamma and two components Weibull distributions, respectively. The component distributions are expressed as:

$$\text{Normal pdf} \quad f(x) = \frac{\exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)}{\sigma\sqrt{2\pi}} \tag{2}$$

$$\text{Lognormal pdf} \quad f(x) = \frac{\exp(-\frac{1}{2}(\frac{lnx-\mu}{\sigma})^2)}{x\sigma\sqrt{2\pi}} \tag{3}$$

Where μ and σ are the location and scale parameters of the normal and lognormal function, x=diameter at breast height (Dbh)

$$\text{Gamma pdf} \quad f(x) = \frac{x^{\alpha-1}}{\beta^{\alpha}\tau(\alpha)}\exp(-\frac{x}{\beta}) \tag{4}$$

2-parameter Weibull pdf (Weibull 1951):

$$f(x) = \frac{\alpha}{\beta}(\frac{x}{\beta})^{\alpha-1}\exp[-(\frac{x}{\beta})^{\alpha}] \tag{5}$$

Where α and β are the shape and scale parameters, respectively. $\Gamma(\alpha)$ is the gamma function with parameters α where $\alpha > 0$, $\beta > 0$ and $\Gamma(\alpha) = (\alpha - 1)!$

The method of maximum likelihood through the expectation maximization (EM) algorithm was used to fit the mixture distributions to the data. The R software (R Core Team 2017) was used in this study. Due to anthropogenic activities in the study site, few trees were encountered in the plots across the 3 stands; as such, the plot data were pooled per stand and fitted. Model assessment was based on negative loglikelihood value, Akaike information criterion (AIC), Bayesian information criterion (BIC) and root mean square error (RMSE). The smaller the values are, the better the model.

Result and Discussion

Finite mixture distributions have been used to described the diameter distributions of the disturbed *Gmelina arborea* stands in Omo Forest Reserve. The estimated parameters of the finite mixture distributions are presented in Table 2. The mixing parameter p gives the contribution of each component to the overall diameter distribution of the forest

Table 2. Estimated parameters of the four finite mixture distributions for the three stands

Stand	Distributions	Parameters					
	2 Normal mix	p_1	μ_1	σ_1	p_2	μ_2	σ_2
1984		0.342	9.674	1.995	0.658	23.170	7.647
1990		0.412	13.052	4.361	0.588	28.736	5.542
2005		0.472	10.434	2.929	0.528	19.519	5.145
	2 Lognormal mix	p_1	μ_1	σ_1	p_2	μ_2	σ_2
1984		0.390	2.265	0.223	0.610	3.139	0.300
1990		0.526	2.647	0.421	0.474	3.385	0.164
2005		0.393	2.235	0.286	0.607	2.896	0.274
	2 Gamma mix	p_1	α_1	β_1	p_2	α_2	β_2
1984		0.356	23.326	0.412	0.644	9.864	2.382
1990		0.520	6.344	2.381	0.480	38.126	0.788
2005		0.484	11.358	0.918	0.516	16.274	1.212
	2 Weibull mix	p_1	α_1	β_1	p_2	α_2	β_2
1984		0.315	5.555	10.311	0.685	3.154	25.385
1990		0.399	3.190	14.586	0.601	5.584	30.747
2005		0.415	3.983	11.479	0.585	3.634	20.677

stand. This parameter adds up to 1 for each distribution for the individual stand. This is a major condition for finite mixture model. For example, the mixing parameter values of the 2 component normal mix for stand 1984 were 0.342 and 0.658 -which summed up to 1. This shows that the diameter data of the stands have two distinct groups. The number of components mixtures to include in a finite mixture distribution could be predetermined especially when other information on the data set are available. During forest inventory, trees are usually classified or grouped based on species, tree position etc. (Zasada and Cieszewski 2005). This information can be used to determine the number of component mixtures in the distributions. Jaworski and Podlaski (2012) noted that "the chance of lack of convergence during parameter estimation process increases with the number of components of the finite mixture model". However, in this study, an iterative search was used to determine the optimum number of component mixture with the smallest fit indices. This method was also recommended by Zhang et al. (2001) and Liu et al. (2002).

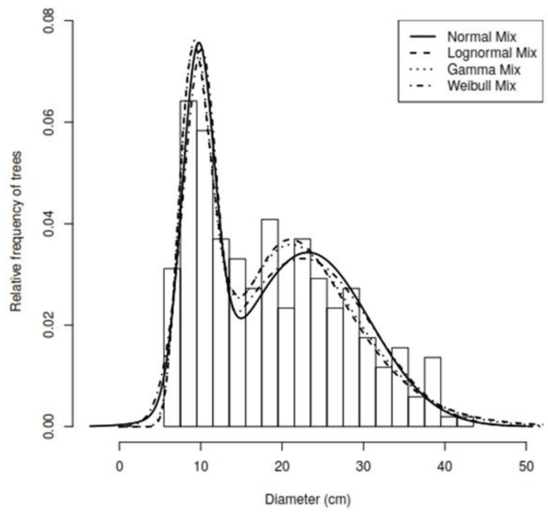

Fig. 1. Observed relative frequency of trees per Dbh class and fitted mixture distributions for stand 1984.

The diameter distributions of the three disturbed *Gmelina arborea* forest stands in Omo Forest Reserve are presented in Fig. 1-4. The graphs showed the observed relative frequency of trees with the fitted normal mix, lognormal mix, gamma mix and Weibull mix by diameter classes for stand 1984, 1990 and 2005 (Fig. 1-3). The mixture distributions described the irregular diameter structure of the forest stands very well compared to the single component distributions. For comparison, the single component of normal lognormal, gamma and Weibull distributions were fitted to the data (see Fig. 4-6). The graph showed that the single component distribution underestimated the frequency in the 10cm diameter class and overestimated in the 18 and 22 cm classes in stand 1984. In addition, the single component distributions resulted in the overall simplification of the forest stand which is in con-

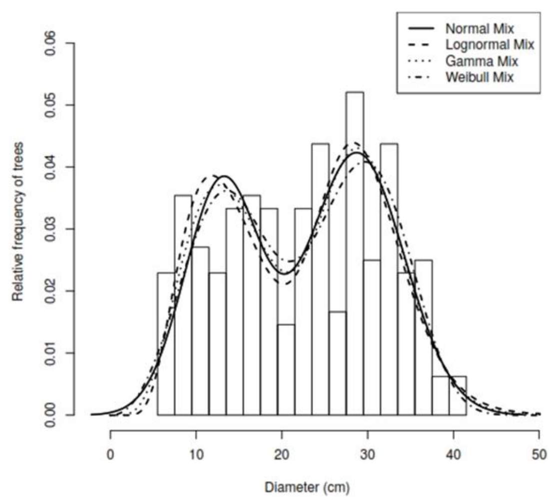

Fig. 2. Observed relative frequency of trees per Dbh class and fitted mixture distributions for stand 1990.

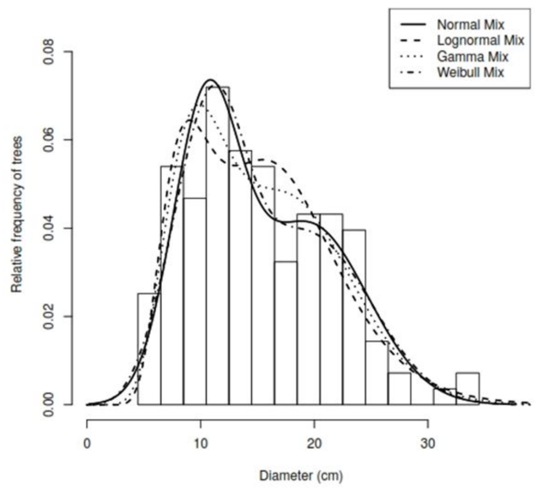

Fig. 3. Observed relative frequency of trees per Dbh class and fitted mixture distributions for stand 2005.

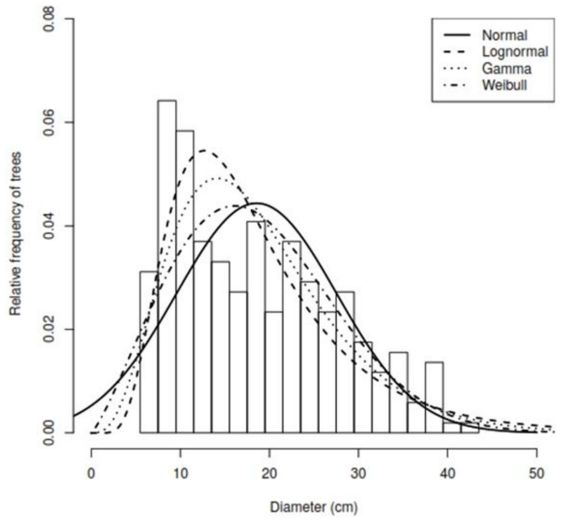

Fig. 4. Observed relative frequency of trees per Dbh class and fitted single component distributions for stand 1984.

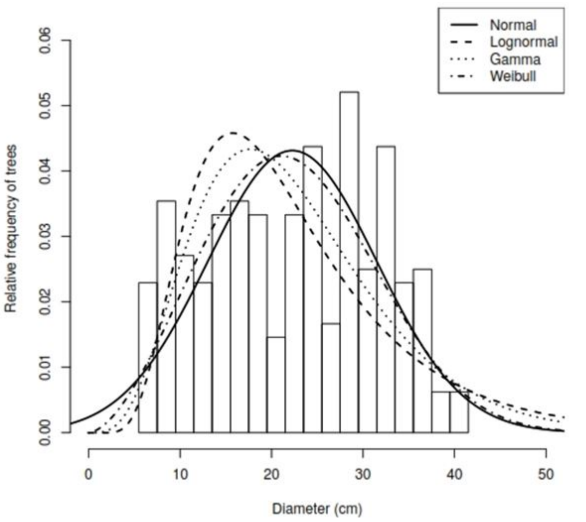

Fig. 5. Observed relative frequency of trees per Dbh class and fitted single component distributions for stand 1990.

trasts with reality. However, the mixture distributions fitted to the same stand revealed the irregularity in the forest stands. Similar observation was reported by Lui et al. (2002), Zhang and Liu (2006) and Lui et al. (2014). The irregularities in the structure of the *Gmelina arborea* stands can be attributed to the illegal logging activities and agricultural activities. This has been a major constraint of the forest sector, especially in Nigeria.

The negative loglikelihood (-ΛΛ), AIC, BIC and

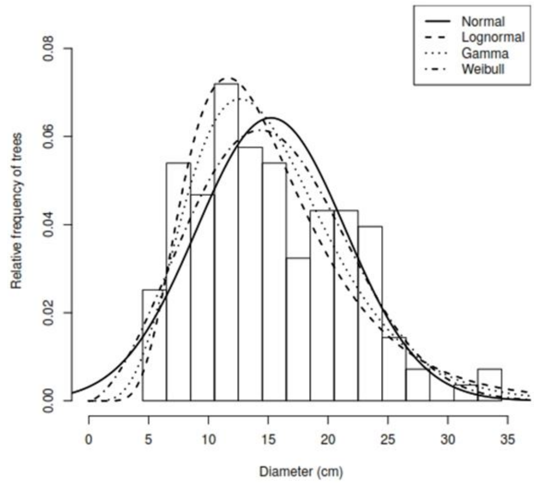

Fig. 6. Observed relative frequency of trees per Dbh class and fitted single component distributions for stand 2005.

RMSE values of the normal mix, lognormal mix, gamma mix and Weibull mix were relatively the same for the three stands (Table 3). The -ΛΛ, AIC, BIC and RMSE values ranged from -715.233 to -348.375, 703.926 to 1433.588, 718.598 to 1451.334 and 3.003 to 7.492, respectively. The fit indices values for the single component distributions were larger than the mixture distributions (though not documented). This shows the inadequacy of the single component distributions in describing the irregular diameter distribution of the degraded forest stand. The study is in line with Liu et al. (2014) who reported lower RMSE value for the mixture Weibull distribution than the single component Weibull in their study on "modelling diameter distributions of mixed-species forest stands". Similarly, Jaworski and Podlaski (2012) found a better performance with mixture gamma than the single component gamma distribution. It is obvious that finite mixture modelling remains the most effective and efficient technique that can be used to model multi-species, irregular or complex forest stands. The fitting method especially, maximum likelihood through the EM algorithm is relative simple, and can be achieved with different statistical tools.

In conclusion, the application of finite mixture of normal, lognormal, gamma and Weibull distributions provides more accurate and precise characterisation of the irregular diameter distribution of the degraded *Gmelina arborea* plantation; in consequence, the forest stand structure. This

Table 3. Fit indices for the mixture distribution

Stand	Distribution	-ΛΛ	AIC	BIC	RMSE
1984	Normal mix	715.223	1440.447	1458.192	7.242
	Lognormal mix	712.485	1434.969	1452.714	6.188
	Gamma mix	711.794	1433.588	1451.334	7.302
	Weibull mix	714.764	1439.529	1457.274	7.492
1990	Normal mix	690.534	1391.067	1408.470	5.699
	Lognormal mix	689.097	1388.194	1405.597	5.193
	Gamma mix	688.238	1386.477	1403.879	4.965
	Weibull mix	688.037	1386.073	1403.477	5.455
2005	Normal mix	348.371	706.742	721.415	3.003
	Lognormal mix	347.437	704.874	719.546	4.220
	Gamma mix	346.963	703.926	718.598	3.006
	Weibull mix	348.375	706.749	721.422	3.441

-ΛΛ=negative loglikelihood (the larger the value in negative sign, the better the model).

approach can be used to describe other irregular forest stand structures, especially the multi-species forest in Nigeria.

References

Bailey RL, Dell TR. 1973. Quantifying diameter distributions with the Weibull function. For Sci 19: 97-104.

Izenman AJ, Sommer CJ. 1988. Philatelic mixtures and multimodal densities. J Am Stat Assoc 83: 941-953.

Jaworski A, Podlaski R. 2012. Modelling irregular and multimodal tree diameter distributions by finite mixture models: an approach to stand structure characterisation. J For Res 17: 79-88.

Liu C, Zhang L, Davis CJ, Solomon DS, Gove JH. 2002. A finite mixture model for characterizing the diameter distribution of mixed-species forest stands. For Sci 48: 653-661.

Liu F, Li F, Zhang L, Jin X. 2014. Modeling diameter distributions of mixed-species forest stands. Scand J For Res 29: 653-663.

Maltamo M, Kangas A, Uuttera J, Torniainen T, Saramäki J. 2000. Comparison of percentile based prediction methods and the Weibull distribution in describing the diameter distribution of heterogeneous Scots pine stands. For Ecol Manag 133: 263-274.

Ogana FN, Itam ES, Osho JSA. 2017. Modeling diameter distributions of Gmelina arborea plantation in Omo Forest Reserve, Nigeria with Johnson's SB. J Sustain For 36: 121-133.

Petráš R, Mecko J, Nociar V. 2010. Diameter structure of the stands of poplar clones. J For Sci 56: 165-170.

Podlaski R. 2010. Diversity of patch structure in Central European forests: are tree diameter distributions in near-natural multilayered Abies-Fagus stands heterogeneous? Ecol Res 25: 599-608.

Podlaski R. 2010. Two-component mixture models for diameter distributions in mixed-species, two-age cohort stands. For Sci 56: 379-390.

Podlaski R. 2017. Forest modelling: the gamma shape mixture model and simulation of tree diameter distributions. Ann For Sci 74: 29.

R Core Team. 2017. R: A language and environment for statistical computing. http://www.R-project.org/.

Tsogt K, Lin C. 2014. A flexible modeling of irregular diameter structure for the volume estimation of forest stands. J For Res 19: 1-11.

Weibull W. 1951. A statistical distribution function of wide applicability. J Appl Mech 18: 293-297.

Zasada M, Cieszewski CJ. 2005. A finite mixture distribution approach for characterizing tree diameter distributions by natural social class in pure even-aged Scots pine stands in Poland. For Ecol Manag 204: 145-158.

Zhang L, Gove JH, Liu C, Leak WB. 2011. A finite mixture of two Weibull distributions for modeling the diameter distributions of rotated-sigmoid, uneven-aged stands. Can J For Res 31: 1654-1659.

Zhang L, Liu C. 2006. Fitting irregular diameter distributions of forest stands by Weibull, modified Weibull, and mixture Weibull models. J For Res 11: 369-372.