Research paper

A percentile-based estimator for the log-logistic function: Application to forestry

Friday Nwabueze Ogana

Ogana, F.N. 2020. A percentile-based estimator for the log-logistic function: Application to forestry. – Forestry Studies | Metsanduslikud Uurimused 72, 107–120, ISSN 1406-9954. Journal homepage: http://mi.emu.ee/forestry.studies

Abstract. Developing a simplified estimation method without compromising the performance of the distribution is germane to forest modelling. Few estimation methods exist for the Log-Logistic distribution and are relatively complex. A simplified estimator for the Log-Logistic parameters will increase its application in diameter distribution yield systems. Therefore, in this study, a percentile-based estimator was applied for the Log-Logistic distribution. The Kolmogorov-Smirnov, Anderson-Darling and Cramer-von Mises statistics were used to evaluate the method in two natural forest stands and two monospecific plantations of *Gmelina arborea* Roxb. and *Tectona grandis* Linn. f. in Nigeria. The parameter recovery model (PRM) and parameter prediction model (PPM) were used to predict the diameter distributions of independent stands of *G. arborea* and *T. grandis*. The results showed that the percentile estimator did not compromise the quality of fits of the Log-Logistic function across the four forest stands and are comparable to the maximum likelihood estimator. The 25th and 75th, and 40th and 80th were the best sample percentiles for the estimator. The predicted diameter distributions of *G. arborea* and *T. grandis* stands from the PRM and PPM were reasonable and compare well with the observed distribution. Thus, either of the models can be incorporated into the growth and yield system of forest stand management.

Key words: percentile method, parameter prediction model, *Gmelina arborea*, *Tectona grandis.*

Author's address: Department of Social and Environmental Forestry, University of Ibadan, Nigeria; *e-mail: ogana_fry@yahoo.com; ORCID: 0000-0002-8388-204X

Introduction

Statistical functions are often used to model diameter distributions of forest stands (e.g., Poudel & Cao, 2013; Gorgoso-Varela & Rojo-Alboreca, 2014; Ogana *et al*., 2017), and have become an important part of forest management and planning. Diameter distribution forms the link between stand level and tree level models (Nord-Larsen & Cao, 2006). It can be used to provide information on the growth, value and volume production of forest stands (Gorgoso *et al*., 2012; Ogana *et al*., 2017).

Some important statistical functions that have been adequately used to model diameter distribution of forest stands include: the beta, gamma, Johnson's SB and Weibull; more recently the Burr XII and Logit-Logistic functions. Extensive forestry literature is available on the application of these statistical functions. For example, Poudel & Cao (2013) compared several methodologies to estimate the parameters

DOI: 10.2478/fsmu-2020-0009

of Weibull for tree diameter characterisation. Similarly, Gorgoso-Varela & Rojo-Alboreca (2014) evaluated different estimation methods for fitting the Johnson's SB and Weibull functions to pedunculate oak (*Quercus robur* L.) and birch (*Betula* sp.) stands in Northwest Spain. In Nigeria, Ogana *et al*. (2017, 2018) modelled the diameter distributions of *Gmelina arborea* Roxb. and *Eucalyptus camaldulensis* Dehn. plantations, respectively. Recently, Sun *et al*. (2019) characterised tree diameter distributions for mixed stands of oak and pine in China. Other studies include: Zhang *et al*. (2003), Cao (2004), Wang & Rennolls (2005), Palahi *et al*. (2007), Zheng & Zhou (2010), Gorgoso *et al*. (2008, 2012), etc.

Although research on diameter distribution has spanned several decades in forestry, there are however, some simplified statistical distributions such as the Log-Logistic which have not been sufficiently applied to forestry. The two-parameter Log-Logistic distribution (hereafter referred to as LogL) is a continuous probability distribution characterised by shape (α) and scale (β) parameters. The LogL has been applied to several disciplines, including actuarial science, computer science, economics, hydrology, medicine, etc. The LogL also known as Fisk distribution in economics was applied to model income distribution (Kleiber & Kotz, 2003). It has also been used in survival analysis to model accelerated failure time (Bennett, 1983). In hydrology, Ashkar & Mahdi (2006) modelled precipitation and the rates of stream flow using the LogL function. The application of LogL to forestry has been limited. Ogana & Dau (2019) used the LogL function to derive crown distributions from diameter at breast height for *Parkia biglobosa* Benth. stands in Nigeria.

The choice of distribution function depends on its relative flexibility, simplicity in terms of model expression, ease of parameter estimation and ease of computing proportion of trees in diameter classes (Burkhart & Tomé, 2012). The LogL has a simplified expression and a closed-form cumulative distribution function (cdf), thus, it does not require numerical integration to estimate the proportion of trees in diameter classes. Furthermore, different estimation methods such as the maximum likelihood estimation (MLE), generalized moments, and the least square have been used to estimate the parameters of LogL with different levels of success (e.g., Kleiber & Kotz, 2003; Ashkar & Mahdi, 2006; Ogana & Dau, 2019). These methods vary in complexity; and as such, there is a need for developing a simplified method of estimating the parameters of LogL for characterising tree diameter distributions. The percentile estimator is a good substitute to the MLE (Clutter *et al*., 1983). This method has not been used to estimate the parameter of LogL in forestry. A simplified estimation method for the LogL function will increase its application in diameter distribution yield systems.

Generally, diameter distribution yield systems predict values for the parameters of distribution functions (Clutter *et al*., 1983). The parameters are related either directly with stand variables such as the quadratic mean diameter, basal area per ha, stand age, number of trees per ha, site index (or the average height of dominant trees), etc. (i.e., parameter prediction model) or diameter moment/percentiles derived from forest stand variables (parameter recovery model). This can be used for implicit prediction of future yield. To date, no published forestry literature exists that relate the parameters of LogL distribution with stand variables to the knowledge of the researcher. Therefore, the objectives of this study were to evaluate a percentile-based estimator for the fitting of the LogL function; and to develop models that could be used to derive the parameters of distribution using Nigerian forest stands as a case study.

Methodology

Data

The data used in this study were collected from two natural forest stands (multiple species composition with indeterminate age structure) and two monoculture stands of *Gmelina arborea* Roxb. and *Tectona grandis* Linn. f. One of the natural stands is situated in Ikrigon Forest Reserve, Cross River State, Nigeria and lies between Latitude 6°17.597′–6°17.862′ N and Longitude 8°35.597′–8°35.276′ E with an area of 542.7 ha. The second natural forest and part of the *G. arborea* stand are in Oluwa Forest Reserve, Ondo State, Nigeria. The natural forest lies between 6.83°–6.91° N and Longitude 4.52°–4.59° E with an area of 629 km^2. The *G. arborea* stand lies between Latitude 6°55′–7°20′ N and Longitude 3°45′–4°32′ E and occupies an area of 87,816 ha (Onyekwelu, 2001). *T. grandis* and *G. arborea* data were also collected from Omo Forest Reserve. Omo Forest Reserve is situated in Ogun State, Nigeria between Latitude 6°35′–7°05′ N and Longitude 4°10′–4°19′ E with an area of 130,500 ha.

Data were collected from 75 and 35 temporary sample plots (TSPs) in *G. arborea* and *T. grandis* stands, respectively; while 10 and 7 TSPs were used in Ikrigon and Oluwa natural forests, respectively. The plot sizes ranged from 0.04 to 0.25 ha, to achieve at least 30 trees in a plot. All living trees within the plots of the monoculture stands were measured, while in the natural forest, only trees with diameter at breast height (Dbh) ≥ 10 cm were measured. Details of the sampling procedures used for the data collection in Ikrigon natural forest, Oluwa natural forest, *G. arborea* and *T. grandis* can be found in T.E. Adeniyi (unpubl.) and Ogana *et al*. (2015, 2017), and Chukwu & Osho (2018), respectively. The descriptive statistics of the Dbh, total height (*Ht*), quadratic mean diameter (*dg*), dominant height (i.e., average height of the 100 tallest trees per ha, H_o), basal area per ha (*G*), density (*N*) and stand age (*A*) are presented in Table 1.

Log-Logistic (LogL)

The probability density function (pdf) and cumulative distribution function (cdf) of the LogL are expressed as:

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}\left[1+\left(\frac{x}{\beta}\right)^{\alpha}\right]^{-2} \quad (1)$$

$$F(x) = \left[1+\left(\frac{\beta}{x}\right)^{\alpha}\right]^{-1} \quad (2)$$

where $f(x)$ = pdf; $F(x)$ = cdf α = shape parameters ($\alpha > 0$); β = scale parameter ($\beta > 0$). The LogL has a closed-form cdf as to facilitate the estimation of tree proportion in different diameter classes without resorting to numerical integration.

The percentile estimator

The idea of the percentile estimator for the LogL stems from the principle developed for the Weibull distribution by Clutter *et al*. (1983). The idea is, if two sample percentiles are known, each can be related to its corresponding LogL cdf, and the resulting equations are iteratively solved to get the estimates of shape (a) and scale (β). Thus, the equations were developed as follows: "Let X_p represent the p-percentile value in the sample so that $100p$-percentile of the sample values are less than X_p". Therefore, from the definition of the LogL cdf:

$$P = \left[1+\left(\frac{\beta}{X_p}\right)^{\alpha}\right]^{-1} \quad (3)$$

Equation (3) was then solved for X_p to give equation (4)

$$X_p = \beta\left(\frac{1}{P}-1\right)^{-1/\alpha} \quad (4)$$

The parameters of the LogL distribution were derived from equation (4). Since this is the first attempt of applying this percentile-based estimator for the LogL, several sample percentiles were considered. The following sample percentiles were analysed:

- Method 1 (M1): 25th and 75th sample percentiles
- Method 2 (M2): 17th and 97th sample percentiles
- Method 3 (M3): 24th and 93rd sample percentiles
- Method 4 (M4): 40th and 80th sample percentiles
- Method 5 (M5): 30th and 70th sample percentiles
- Method 6 (M6): 35th and 65th sample percentiles
- Method 7 (M7): 40th and 60th sample percentiles
- Method 8 (M8): 33rd and 67th sample percentiles

From equation (4), method 1 was derived as:

$$X_{25} = \beta \left(\frac{1}{0.25} - 1 \right)^{-1/\alpha} \quad (5)$$

$$X_{75} = \beta \left(\frac{1}{0.75} - 1 \right)^{-1/\alpha} \quad (6)$$

where X_{25} and X_{75} are the diameters corresponding to the 25 and 75%, respectively. Equation (5) and (6) were solved iteratively using the Newton-Raphson algorithm of the 'lmfor' package (Mehtätalo, 2017) implemented in R (R Core Team, 2017). A similar procedure was used for M2 to M8 by substituting the various percentiles and solving the system of equation iteratively.

Table 1. Descriptive statistics of the inventoried data from natural forests and plantations.

Forest	Statistics	Dbh (cm)	Ht (m)	dg (cm)	Ho (m)	G (m^2ha^{-1})	N (trees ha^{-1})	A (years)
Ikrigon Natural Forest	Mean	30.0	24.1	33.8	31.9	28.0	310.4	
63 species	Min	10.0	5.7	29.4	27.6	16.9	184.0	na*
776 trees	Max	125.0	43.8	39.6	35.9	40.9	440.0	
	SD	15.9	7.5	3.5	2.5	8.3	72.8	
Oluwa Natural Forest	Mean	25.1	16.9	29.6	25.7	19.4	227.1	
58 species	Min	10.0	4.8	26.1	22.0	14.5	200.0	na*
486 trees	Max	118.5	63.7	34.5	33.0	31.7	352.0	
	SD	16.7	9.0	2.7	3.7	6.7	56.6	
G. arborea	Mean	21.2	13.9	22.3	21.6	30.5	710.3	26.6
486 trees	Min	3.0	1.6	14.7	8.8	6.2	300.0	11.0
	Max	54.5	42.7	31.0	34.5	72.9	1525.0	39.0
	SD	9.8	5.8	3.7	4.5	19.5	308.6	7.7
T. grandis	Mean	17.9	16.7	18.5	21.3	24.2	877.6	19.8
1,916 trees	Min	6.0	6.6	14.3	17.9	11.5	624.0	13.0
	Max	37.9	26.5	22.4	24.0	50.8	1744.0	27.0
	SD	5.3	3.5	2.2	1.9	8.9	198.6	5.2

*na = not applicable

To ascertain the suitability of the percentile method, it was compared with the well-known maximum likelihood estimator (MLE). The MLE involves taking the partial derivative of the log-likelihood function of LogL with respect to the parameters (α and β) and setting the expression equal to zero. The resulting function is solved by numerical algorithm to get the estimates of the parameters. The log-likelihood function (ℓ) of the LogL is expressed as:

$$log(\ell) = nlog(\alpha) - nlog(\beta) + (\alpha - 1)\sum_{i=1}^{n} log(x_i) - n(\alpha - 1)log(\beta) - 2\sum_{i=1}^{n}\left[1 + \left(\frac{x_i}{\beta}\right)^{\alpha}\right] \quad (7)$$

the partial derivative of equation (7) with respect to α and β will give equations (8) and (9), respectively:

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} log(x_i) - nlog(\beta) - 2\sum_{i=1}^{n} \frac{\left(\frac{x_i}{\beta}\right)^{\alpha} log\left(\frac{x_i}{\beta}\right)}{1 + \left(\frac{x_i}{\beta}\right)^{\alpha}} = 0 \quad (8)$$

$$\frac{\partial \ell}{\partial \beta} = -\frac{n}{\beta} - \frac{n(\alpha - 1)}{\beta} + \frac{2\alpha}{\beta}\sum_{i=1}^{n} \frac{\left(\frac{x_i}{\beta}\right)^{\alpha}}{1 + \left(\frac{x_i}{\beta}\right)^{\alpha}} = 0 \quad (9)$$

where ∂ is the symbol of the partial derivative, ℓ represents the likelihood function, x_i is diameter at breast height and *i* ranged from 1 to n, other parameters are previously defined in equation (1). Since there are no explicit solutions to equations (8) and (9), the estimates were obtained numerically using the *mledist* function from the fitdistrplus package (Delignette-Muller & Dutang, 2015) implemented in R (R Core Team, 2017). The R code for the MLE is presented in Appendix 1.

Evaluation statistics

Three goodness-of-fit statistics were used to evaluate the suitability of the percentile estimator. For each method, the Kolmogorov-Smirnov (*KS*), Anderson-Darling (*AD*) and Cramer-von Mises (W^2) statistics were computed. The smaller the statistics are, the better the method.

KS statistic:

$$KS = max\{max_{1\le i\le n_i}[F_n(x_i) - F_0(x_j)], max_{1\le i\le n_i}[F_o(x_j) - F_n(x_{i-1})]\} \quad (10)$$

AD statistic:

$$AD = -n_i - \sum_{j=1}^{n_i} (2j-1)\left[ln\left(F_o(x_j)\right) + ln\left(1 - F_n(x_{i-1})\right)\right]/n_i \quad (11)$$

W^2 statistic:

$$W^2 = \sum_{i=1}^{n}\left\{\hat{F}(x_i) - \frac{(i-0.5)}{n}\right\}^2 + \frac{1}{12n} \quad (12)$$

where $F(x_i)$ = observed cumulative frequency distribution for x_i (i ranged from 1 to n), $F_0(x_i)$ = theoretical cumulative frequency distribution.

Ranking of methods

The relative rank introduced by Poudel & Cao (2013) and recently used by Sun *et al*. (2019) was used for this study. It is expressed as:

$$R_i = 1 + \frac{(m-1)(S_i - S_{min})}{S_{max} - S_{min}} \quad (13)$$

where R_i = relative rank of method i (i = 1, 2, …, m); m = number of methods assessed (8 percentiles plus the MLE, i.e., 9 altogether), S_i = evaluation statistics value of method i; S_{min} and S_{max} = minimum and maximum value of S_i, respectively. The relative rank is a real number between 1 (best) and 9 (worst). For each method, the relative ranks were summed across the three goodness-of-fit statistics to ascertain the best percentile point for the estimator.

Modelling the diameter distribution

A scatterplot was used to establish the form of association between the dependent (LogL parameters and diameter percentiles) and independent variables (stand variables, e.g., quadratic mean diameter (*dg*), dominant height (H_o), stand density (*N*), basal area per ha (*G*), stand age (*A*), etc.). Thereafter, a stepwise linear regression technique was used in the modelling process. This ensured that only the best predictor variables were included in the models. Relating the parameters directly with stand variables is known as the parameter prediction method (PPM); while modelling the diameter percentiles with stand variables from which the LogL parameters could be derived is referred to as the parameter recovery model (PRM). Both PPM and PRM were evaluated to determine the best option for the LogL distribution. The models were of this form:

$$Y_i = f(X_i, \beta) + \varepsilon_i \quad (14)$$

where Y_i is a vector of the dependent variables (parameters of the LogL distribution and various diameter percentiles), X_i is the vector of the predictor variables (quadratic mean diameter, dominant height, stand density, basal area per ha, stand age, etc.) and is the error term.

Models were only developed for the *G. arborea* and *T. grandis* stands because a sufficient number of plots is required for effective modelling. There were more than 30 plots from these stands compared to the few plots in the natural stands (10 and 7 in Ikrigon and Oluwa, respectively). Since sufficient data are not available for modelling, a k-fold cross-validation was used to ascertain the predictive ability of the models (Sileshi, 2014). Different fit indices, such as root mean square error (*RMSE*), coefficient of variation (*CV*), corrected Akaike Information Criterion (*AICc*), Bayesian Information Criterion (*BIC*) and adjusted coefficient of determination ($\bar{R}^2$) were used to evaluate the cross-validation results.

Results and Discussion

Fit of the LogL from percentile and MLE

The estimated parameters of the LogL from the different percentiles and those from the MLE are presented in Table 2 and Table 3 for the natural forests and plantations, respectively. Little variation was observed in the values of the parameters of the LogL estimated from the MLE and percentiles methods. Evaluation of the performance of the methods based on the fit statistics showed that MLE had the smallest relative rank sum (R) of 3.79 and 3.32 for Ikrigon and Oluwa natural forests, respectively. This was followed by method 4 (M4: 40th and 80th diameter percentiles), method 1 (M1: 25th and 75th) and method 7 (M7: 40th and 60th) with respect to Ikrigon forest. In Olwua natural forest, M4 had the largest R. The Cramer-von Mises (W^2) statistics of M1, M4 and M7 were smaller relative to the value of the MLE in Ikrigon natural forest. Just as with KS and AD, W^2 statistic is an important index that gives the overall squared difference between empirical distribution and the theoretical cdf (Wang, 2005).

Table 2. Estimated parameters, goodness-of-fits and the relative rank sum (R) of the percentiles and MLE in natural forest.

Natural Forest	Percentiles	Parameters		Goodness-of-fit			R
		Shape	Scale	KS	AD	W^2	
Ikrigon	M1	2.6293	26.2730	0.0736	8.4744	0.6452	7.29
	M2	3.4970	23.7620	0.1263	21.5400	3.4572	27.00
	M3	3.2008	24.5129	0.0889	10.3310	1.5687	12.34
	M4	2.7855	26.1412	0.0648	5.9910	0.4154	4.49
	M5	2.6278	26.3674	0.0731	8.5303	0.6549	7.29
	M6	2.5619	26.2305	0.0785	9.9085	0.8018	8.97
	M7	2.7658	26.1683	0.0658	6.2362	0.4338	4.77
	M8	2.5572	26.3818	0.0778	10.0510	0.8231	9.00
	MLE	3.2957	26.2105	0.0547	5.2472	0.7167	3.79
Oluwa	M1	3.0771	20.2929	0.1061	6.5088	0.6674	7.29
	M2	3.0150	21.6576	0.0988	8.5493	1.3996	21.48
	M3	2.9375	20.7274	0.1089	7.0406	0.8385	11.56
	M4	2.7995	20.1118	0.1299	8.6210	0.9602	22.20
	M5	3.0966	19.8522	0.1125	7.3039	0.7098	11.73
	M6	3.0229	19.8817	0.1172	7.5092	0.7532	13.74
	M7	3.2683	19.6982	0.1037	7.5823	0.6984	11.12
	M8	3.0186	19.8512	0.1180	7.6141	0.7661	14.41
	MLE	3.3928	20.4790	0.0827	6.4187	0.6964	3.32

Similarly, in *G. arborea* stand, the MLE had the smallest relative rank sum of 3.05 which was closely followed by M4 with a value of 3.47. The performances of M1 and M7 were equally good. In the case of *T. grandis* stand, M4 had the smallest relative rank sum of 3.00 whereas MLE had 3.52. This was followed by M1 and M5 (30th and 70th diameter percentiles). Method 2 (17th and 97th diameter percentiles) with the largest relative rank sum had the worst performance in three out of the four stands. The 17th and 97th sample percentiles were recommended for the Weibull distribution by Shiver (1988). Thus, the study shows that these diameter percentiles are unsuitable for the LogL distribution. Gorgoso *et al.* (2007) also observed a poor fit with the percentiles for the Weibull distribution in the birch-dominated stands of Northwest Spain. Also, the most efficient sample percentiles – 24th and 93rd recommended for the Weibull function by Dubey (1967) performed poorly (M3) for the LogL function.

The graph of the relative frequency of trees against the equal diameter class of the best diameter percentiles and MLE for the four stands are shown in Figure 1a, b, c and d. The graph showed that the estimation methods produce a good representation of the four stands, that is, the reverse J-shaped and the Gaussian distribution typical of natural forests (uneven-aged) and plantations (even-aged), respectively.

Table 3. Estimated parameters, goodness-of-fits and the relative rank sum (R) of the percentiles and MLE in plantations.

Plantation	Percentiles	Parameters		Goodness-of-fit			R
		Shape	Scale	*KS*	*AD*	W^2	
G. arborea	M1	2.9081	19.2593	0.0739	19.0410	2.0426	4.13
	M2	3.7818	16.2732	0.2120	182.6300	33.5450	27.00
	M3	3.5858	17.7908	0.1325	60.1980	11.6370	11.50
	M4	3.3184	19.8874	0.0598	19.9870	2.0377	3.47
	M5	2.8103	19.6023	0.0853	22.5520	2.1931	4.91
	M6	2.8513	19.8796	0.0862	22.9980	2.1995	4.98
	M7	3.0304	20.1196	0.0773	21.7450	2.0812	4.44
	M8	2.8908	19.8028	0.0824	21.2250	1.9992	4.66
	MLE	3.3267	19.4248	0.0526	18.0290	2.1989	3.05
T. grandis	M1	5.0963	17.4920	0.0593	8.5247	0.7829	6.42
	M2	6.0548	16.5020	0.1084	33.8000	6.4582	27.00
	M3	6.1169	16.9031	0.0840	16.3320	2.8739	14.37
	M4	5.3251	17.2658	0.0415	5.2366	0.5085	3.00
	M5	5.1543	17.5622	0.0594	8.9145	0.8782	6.67
	M6	4.9264	17.4620	0.0657	10.6700	1.0045	8.08
	M7	4.8683	17.3897	0.0657	11.1180	1.0443	8.26
	M8	4.9804	17.5226	0.0655	10.5200	1.0175	8.03
	MLE	5.6892	17.2884	0.0436	5.4299	0.6729	3.52

However, the relative frequency of trees was underestimated in the diameter class of < 20 cm in Oluwa natural forest (Figure 1b) and overestimated in the middle diameter class of *G. arborea* stand.

The results obtained with the LogL function in this study were better than those reported by Ogana & Wali (2018) for similar natural forest stands. Studies on LogL in forestry have used either the MLE (Ogana & Wali, 2018) or the least square regression method (Ogana & Dau, 2019). The least square method may seem handy, it is however, time-consuming. One important advantage of the MLE is the ability to evaluate the reliability of the estimates of the parameters using the standard error which can be derived from the hessian matrix. However, the MLE requires specifying the log-likelihood function for the distribution which may not converge due to poor initial guess for the parameters. The MLE is also affected by sample size (Shiver, 1988). Contrary to the MLE and least square method, the percentile estimator is more simplified, handier and does not require a complex iterative procedure. Therefore, when considering simplicity, vis-à-vis the ease of estimating the LogL parameters, the percentile estimator can be considered as a good alternative.

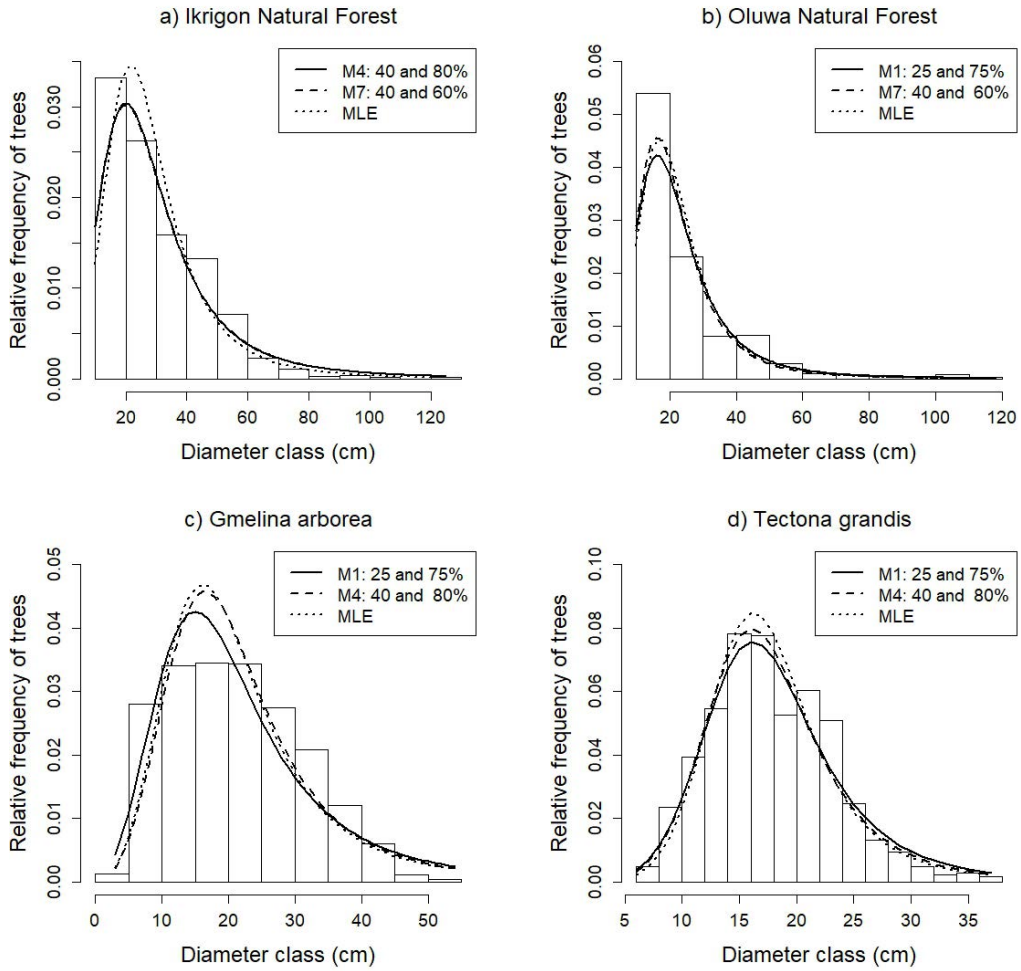

Figure 1. Observed and fitted LogL diameter distributions of the natural forests and plantations.

Modelling the diameter distribution

Two sets of models for predicting the diameter distributions of *G. arborea* and *T. grandis* stands were developed. The first set of models relate the best diameter percentiles with stand variables (parameter recovery model, PRM) in a simple linear relationship wherein the shape and scale parameters of the LogL were recovered. The quadratic mean diameter (*dg*) was the only suitable predictor variable for the dependent variables (diameter percentiles). The second set of models relate the LogL parameters directly with stand variables (parameter prediction model, PPM). The scale (β) parameter was predicted from the quadratic mean diameter (*dg*) in a simple linear relationship. The shape (α) parameter was predicted from the inverse of the natural logarithm of the ratio of the 25th diameter percentile and *dg* i.e.,

$$\alpha = a + b\frac{1}{ln\left(P25/dg\right)};$$

where *a* and *b* are the estimated parameters. The parameters and standard errors of the models are presented in Table 4. The results of the 5-fold cross-validation showed that both model forms had a relatively low *RMSE*, *CV*, *AICc*, *BIC* and a high $\bar{R}^2$. Thus, they can be relied on for prediction purpose. Sun *et al*. (2019) also found *dg* as one of the most important predictors of the Weibull parameters in the mixed stand of oak and pine.

The graphs of the application of the PRM and PPM to independent stands of *G. arborea* (0.08 ha) and *T. grandis* (0.125 ha) are shown in Figure 2a and b. The predicted diameter distributions were comparable to the observed diameter distributions of the stands. The first PRM (models from the 25th and 75th percentiles) had *KS*, *AD* and W^2 of 0.0670, 0.7583 and 0.0751, respectively in the 0.08 ha *G. arborea* stand. These values were respectively, 0.0778, 0.6879 and 0.0782 in the second PRM (models from the 40th and 80th percentiles). The PPM had 0.0727, 0.5968 and 0.0669, respectively. In the *T. grandis*, the *KS*, *AD* and W^2 of the PRM were 0.0993, 0.9301 and 0.1780, respectively. Whereas the PPM had 0.1178, 2.0522 and 0.3479, respectively. The second PRM could not predict the diameter distribution of the stand; as such, it was not represented in the graph.

The performances of the PRM and PPM were relatively the same for the LogL distribution. One major argument with the PPM is the difficulty of developing models that would account for high variations in the parameters of distributions (Borders & Patterson, 1990). While this statement may be valid for the shape parameters of some distributions, such as the well-known Weibull, it is not application to the LogL function. The models predicting the shape and scale parameters of the LogL explained more than 0.75 (> 75%) of the variation in the parameters for the stands. Thus, both the PRM and PPM could be used for the LogL in the *G. arborea* and *T. grandis* stands.

Table 4. Models for the percentiles, scale and shape parameters and their cross-validation indices.

Species	DV*	Parameter	Estimate	*SE*	5-fold Cross-validation				
					RMSE	*CV*	*AICc*	*BIC*	$\bar{R}^2$
G. arborea	P25	b_0	-0.590	1.570	4.830	4.801	116.937	123.460	0.533
		b_1	0.632	0.069					
	P40	b_0	-2.975	1.676	5.539	5.441	126.491	133.015	0.677
		b_1	0.912	0.074					
	P60	b_0	-0.768	1.159	2.519	2.615	72.663	79.186	0.854
		b_1	1.050	0.051					

Species	DV*	Parameter	Estimate	*SE*	5-fold Cross-validation				
					RMSE	*CV*	*AICc*	*BIC*	$\bar{R}^2$
	P75	b_0	-0.677	1.116	2.446	2.418	67.149	73.672	0.897
		b_1	1.232	0.049					
	P80	b_0	-0.468	1.148	2.567	2.561	71.250	77.774	0.900
		b_1	1.288	0.051					
	Shape	b_0	1.952	0.124	0.138	0.147	-146.720	-140.196	0.764
		b_1	-0.826	0.054					
	Scale	b_0	-0.753	0.955	1.803	1.768	44.331	50.855	0.860
		b_1	0.888	0.042					
T. grandis	P25	b_0	-3.885	1.374	0.917	0.874	-1.105	2.556	0.844
		b_1	0.975	0.074					
	P40	b_0	1.562	1.189	0.668	0.652	-10.687	-7.025	0.874
		b_1	0.951	0.064					
	P60	b_0	0.263	1.016	0.495	0.475	-21.063	-17.401	0.912
		b_1	0.995	0.055					
	P75	b_0	2.704	1.232	0.671	0.704	-8.334	-4.672	0.873
		b_1	0.982	0.066					
	P80	b_0	4.946	1.331	0.786	0.829	-3.200	0.461	0.833
		b_1	0.905	0.071					
	Shape	b_0	2.514	0.397	0.368	0.378	-30.438	-26.776	0.773
		b_1	-1.028	0.098					
	Scale	b_0	-1.306	0.625	0.186	0.181	-53.142	-49.480	0.965
		b_1	1.002	0.033					

*DV = dependent variables

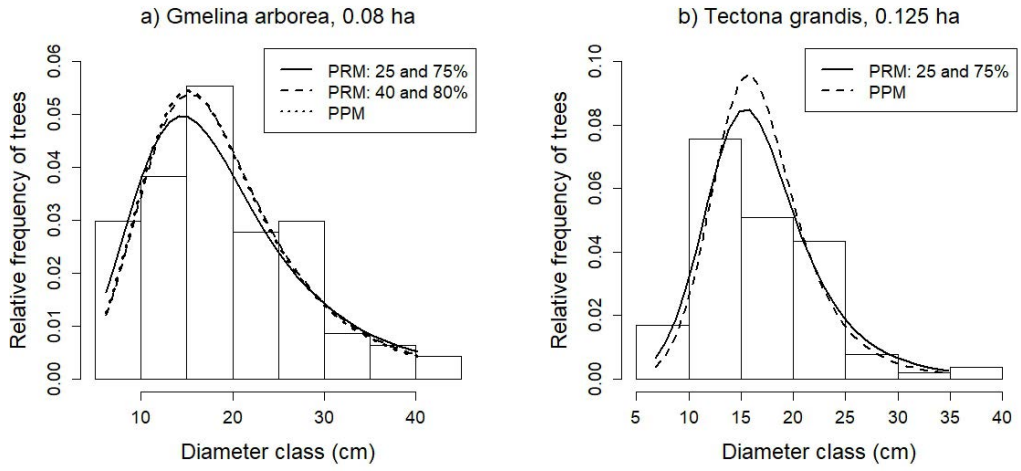

Figure 2. The observed and predicted diameter distributions from PRM and PPM models.

Conclusion

Developing a simplified estimation method without compromising the performance of the distribution is germane to forest modelling. The percentile estimator introduced for the Log-Logistic did not compromise the quality of fits of the model across the four forest stands in Nigeria and are comparable to the maximum likelihood estimator. The 25th and 75th, and 40th and 80th were the best sample percentiles for the estimator in the forest stands. The predicted diameter distributions of *G. arborea* and *T. grandis* stands from the PRM and PPM were reasonable and compared well with the observed diameter distribution. This implies that both the PRM and PPM can be incorporated into the growth and yield system of forest stand management.

References

Ashkar, F., Mahdi, S. 2006. Fitting the log-logistic distribution by generalized moments. – Journal of Hydrology, 328, 694–703.

Bennett, S. 1983. Log-Logistic regression models for survival data. – Applied Statistics, 32(2), 165–171. https://doi. org/10.2307/2347295.

Borders, B.E., Patterson, W.D. 1990. Projecting stand tables: a comparison of the Weibull diameter distribution method, a percentile-based projection method, and a basal area growth projection method. – Forest Science, 36, 413–424.

Burkhart, H.E., Tomé, M. 2012. Modelling Forest Trees and Stands (2nd ed). Dordrecht, Springer. 457 pp. https://doi. org/10.1007/978-90-481-3170-9.

Cao, Q.V. 2004. Predicting parameters of a Weibull function for modelling diameter distribution. – Forest Science, 50, 682–685.

Chukwu, O., Osho, J.S.A. 2018. Basal area-stump diameter models for *Tectona grandis* Linn. F. stands in Omo Forest Reserve, Nigeria. – Journal of Forest and Environmental Science, 34, 119–125.

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., Bailey, R.L. 1983. Timber Management: A Quantitative Approach. New York, John Wiley & Sons. 333 pp.

Delignette-Muller, M.L., Dutang, C. 2015. Fitdistrplus: An R package for fitting distributions. – Journal of Statistical Software, 64(4), 1–34. http://www.jstatsoft.org/v64/i04/.

Dubey, S.D. 1967. Some percentile estimators for Weibull parameters. – Technometrics, 9, 119–129.

Gorgoso-Varela, J.J., Rojo-Alboreca, A. 2014. Short communication. A comparison of estimation methods for fitting Weibull and Johnson's S_B functions to pedunculate oak (*Quercus robur*) and birch (*Betula pubescens*) stands in northwest Spain. – Forest Systems, 23(3), 500–505.

Gorgoso, J.J., Rojo, A., Camara-Obregon, A., Dieguez-Aranda, U. 2012. A comparison of estimation methods for fitting Weibull, Johnson's S_B and beta functions to *Pinus pinaster, Pinus radiata* and *Pinus sylvestris* stands in northwest Spain. – Forest Systems, 21(3), 446–459.

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

Gorgoso, J.J., Alvárez-González, J.G., Rojo, A., Grandas-Arias, J.A. 2007. Modelling diameter distributions of *Betula alba* L. stands in northwest Spain with the two parameter Weibull function. – Investigación Agraria: Sistemas y Recursos Forestales, 16(2), 113–123.

Kleiber, C., Kotz, S. 2003. Statistical Size Distributions in Economics and Actuarial Sciences. New Jersey, John Wiley & Sons. 352 pp.

Mehtätalo, L. 2017. Lmfor: functions for Forest Biometrics. R package version 1.2. [WWW document] – URL https://CRAN.R-project.org/package=lmfor. [Accessed 25 October 2018].

Nord-Larsen, T., Cao, Q.V. 2006. A diameter distribution model for even-aged beech in Denmark. – Forest Ecology and Management, 231, 218–225.

Ogana, F.N., Dau, J.H. 2019. Deriving tree crown distributions from diameter at breast height – Journal of Tropical Forestry and Environment, 9(1), 80–88. https://doi.org/10.31357/jtfe.v9i1.3954.

Ogana, F.N., Osho, J.S.A., Gorgoso-Varela, J.J. 2018. Application of extreme value distribution for assigning optimum fractions to distributions with boundary parameters: an eucalyptus plantations case study. – Siberian Journal of Forest Science, 4, 39–48.

Ogana, F.N., Wali, M.D. 2018. Characterising truncated tree diameter data from tropical forest stands in Nigeria. – Journal of Research in Forestry, Wildlife and Environment, 10(4), 81–91.

Ogana, F.N., Itam, E.S., Osho, J.S.A. 2017. Modelling diameter distributions of *Gmelina arborea* plantation in Omo Forest Reserve, Nigeria with Johnson's S_B. – Journal of Sustainable Forestry, 36(2), 121–133. https://doi.org/10.1080/10549811.2016.1263575.

Ogana, F.N., Osho, J.S.A., Gorgoso-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.

Onyekwelu, J.C. 2001. Growth characteristics and management scenarios for plantation-grown *Gmelina arborea* and *Nauclea diderrichii* in south-western Nigeria. – Doctoral thesis. Munich, Hieronymus Publishers. 204 pp.

Palahí, M., Pukkala, T., Blasco, E., Trasobares, 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). – European Journal of Forest Research, 126, 563–571.

Poudel, K.P., Cao, Q.V. 2013. Evaluation of methods to predict Weibull parameters for characterising diameter distributions. – Forest Science, 59(2), 243–252.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [WWW document] – URL http://www.R-project.org/. [Accessed 30 June 2017].

Shiver, B.D. 1988. Sample sizes and estimation methods for the Weibull distribution for unthinned Slash pine plantations diameter distributions. – Forest Science, 34(3), 809–814.

Sileshi, G.W. 2014. A critical review of forest biomass estimation models, common mistakes and corrective measures. – Forest Ecology and Management, 329, 237–254.

Sun, S., Cao, Q.V., Cao, T. 2019. Characterizing diameter distributions for uneven-aged pine-oak mixed forests in the Qinling Mountains of China. – Forests, 10, 596. https://doi.org/10.3390/f10070596.

Wang, M. 2005. Distributional modelling in forestry and remote sensing. – Doctoral thesis. London, University of Greenwich. 187 pp.

Wang, M., Rennolls, K. 2005. Tree diameter distribution modelling: introducing the Logit-Logistic distribution. – Canadian Journal of Forest Research, 35, 1305–1313.

Zhang, L., Packard, K.C., Liu, C. 2003. A comparison of estimation methods for fitting Weibull and Johnson's S_B distributions to mixed spruce-fir stands in north-eastern North America. – Canadian Journal of Forest Research, 33, 1340–1347.

Zheng, L., Zhou, X. 2010. Diameter distribution of trees in natural stands managed on polycyclic cutting system. – Forestry Studies in China, 12(1), 21–25.

Received March 26, 2020, revised June 11, 2020, accepted June 19, 2020

Appendix 1. The R (R Core Team, 2017) code for maximum likelihood estimation.

```
#### Maximum Likelihood Estimation ####
rm(list=ls(all.names=TRUE))
library(fitdistrplus)
library(readxl)
dat <- read_excel("Data.xlsx", sheet="gmelina")
dat1 <- split(dat$dbh, dat$Plot)
dLogL <- function(x, alpha, beta)(alpha/beta)*(x/beta)^(alpha-1)*((1+(x/beta)^al-
pha)^-2)  ##  LogL pdf
pLogL <- function(x, alpha, beta) (1 + (beta/x)^alpha)^-1                    ##  LogL cdf
result <- c()
for (i in 1:75) {
        d <- dat1[[i]]
fitLogL <- mledist(d, "LogL", start=list(alpha=3.1, beta=20.2))
result <- rbind(result, data.frame(Plot = i,
        shape = fitLogL$estimate[1],
        scale = fitLogL$estimate[2]))
}
result
```