Forests and Forest Products Journal 19:73-79

Effects of Plot Dimension on Weibull Parameters

[1]Abwage W. D. and [2]Ogana, F. N.
[1]Department of Forestry and Wildlife, Taraba State University, Jalingo, Nigeria
[2]Department of Social and Environmental Forestry, University of Ibadan, Nigeria
**Corresponding author: E-mail: willcom08@yahoo.com*

Abstract
The main aim of this study was to assess the effect of plot dimension on Weibull parameters estimation in describing diameter distribution of *Eucalyptus camaldulensis* Dehn plantation, Afaka Forest Reserve, Kaduna State, Nigeria. Information on diameter distribution should be obtained from the rightful data analytical tool which will not mislead the manager in decision making. As such, plot dimension should be varied to accommodate a minimum of 60 trees which proves to have a better fit and prediction with Weibull parameters estimated through Maximum Likelihood. The parameters of Weibull were estimated from three dimensions (20 x 20m, 30 x 30m and 40 x 40m) using Maximum likelihood (MLEW), Traditional method of moment (MMW), Moment incorporating skewness (MISW) and percentile. (PW). The data used for this study came from temporary sample plots (TSP). The Weibull parameters were estimated from the three plots dimensions independently through the above parameters estimation method. The three plots dimensions were compared vis-a-vis the parameter estimation method. The comparison was based on Kolmogorov-Sminov (K-S), mean square error (MSE), mean absolute error (MAE) and bias. Based on plots dimension and parameter estimation method, plot size 40 x 40m and MLE proves to have better fitting and prediction. Hence, a larger sample size with MLE appears to be the most suitable for Weibull parameters fitting and diameter prediction study.

Keywords: Diameter distribution prediction, Weibull distribution function, parameter estimation methods, plot dimension

Introduction

In sustainable forest management, diameter distribution model plays a vital role in forecasting the stand structure which gives the manager an overview of what the forest looks like and how is likely going to be in the future. With this structure, the forest manager can make a sound decision in terms of what quantity to cut, when to cut, what amount of stocking is required to replace the amount needed to be cut, and what management practice to adopt etc. Ideal management of forests is based on having accurate information about the status of standing forest inventory. One of this basic information is the distribution of trees in diameter classes, which allows the tree marker to interfere in stands more confidently to preserve the stand structure for sustainable purposes (Harter, and Moore, 1965). This information should be collected from a reasonable sample size of not less than 40 stands (Sghaier *et al.*, 2016).

DeLiocourt (1898) pioneered the first study on tree diameter distribution. The author observed the inverse J-shaped when the number of trees was plotted against equal diameter classes as a frequency histogram. The inverse J-shaped is a typical nature of natural forests that is, a large number of trees with a decreasing frequency as the diameter increases. The nature of a forest plantation is a Gauss distribution, that is, most trees cluster near the average diameter with decreasing frequency at smaller and larger diameters. Different models of probability density functions have been applied to describe the structure of different forest stands including the Johnson's S_B (Gorgoso-Varela and Rojo-Alboreca, 2014), beta (Loetsch *et al.*, 1973; Gorgoso *et al.*, 2012; Ogana *et al.*, 2015), gamma (Zheng and Zhou, 2010; Eslami *et al.*, 2011; Ogana *et al.*, 2015) and Weibull distribution (Palahi *et al.*, 2007; Gorgoso-Varela and Rojo-Alboreca, 2014; Ogana and Gorgoso-Varela, 2015). Among these distributions, the Weibull distribution is the most widely used in forestry. This is because of its relative flexibility in describing varieties of shapes, ease of parameter estimation, and simplicity of estimating proportions in different size classes.

To date, different plots dimension is use for diameter distribution modelling e.g. 0.04 ha, 0.05 ha, 0.0625 ha, 0.09 ha, 0.5 ha etc. The choice of plot dimension is based on the intuition of the inventory expert. Most published forestry literatures varied the plot dimensions to achieve a sizeable number of trees in plots, that is, a minimum of 30 sample trees (e.g. Gorgoso-Varela *et al.*, 2015). The estimate of the parameters of distributions will vary with plot dimension; consequently, the overall fitting performance of the models. The question that comes to mind is, is there an optimum plot dimension for an effective distribution study? To address this question, we evaluated the effects of plot dimension on the

Weibull parameter using different fitting methods.

Methodology
Data Collection
The data for this study were collected from a total of 48 temporary sample plots (TSPs) in the *Eucalyptus camaldulensis* Dehn plantation, Afaka Forest Reserve, Kaduna State, Nigeria. Three plot dimensions were considered in this study: 20m x 20m (0.04 ha), 30m x 30m (0.09 ha) and 40m x 40m (0.16 ha). These plots were established across different age series to capture site and age variations. All live trees within the sampled plots were measured for diameter at breast height and total height. The following stand variables were computed from the data: density, quadratic mean diameter, dominant height, basal area and volume etc. The descriptive statistics as presented in Table 1 below

Table 1: Descriptive Statistics of Stand Variables

Stand variable	Mean	Maximum	Minimum	Standard deviation
Dbh (cm)	10.3	47.2	2.0	6.2
Quadratic mean	11.8	23.9	5.9	3.8
Dominant Height (m)	21.0	30.6	9.0	5.4
Density (N/ha)	753.0	1328	448.0	202.8
Basal area (m^2/ha)	8.52	27.38	1.73	4.92
Volume (m^3/ha)	151.22	792.50	9.70	112.57
No. of Plots = 48				

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

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

The Weibull cumulative distribution function (CDF) is obt ained by the integration of equation 1. It is expressed as:

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

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

Where: *F(x)* is the Weibull cumulative distribution function; *x* is tree diameter to be measured, a, b and c are the location, scale and shape parameters of the distribution respectively.

Fitting Methods
In this study, four methods were used to estimate the parameters of the Weibull distribution, these include: Maximum Likelihood (MLW), Traditional Method of Moments (MMW), Moments Incorporating Skewness (MISW) and Percentiles (PW). These methods were used to fit the Weibull distributions to the data from the three plot dimensions independently.

Maximum Likelihood (MLW): this involves taking the partial derivatives of the

$$L(f) = \prod_{i=1}^{n} \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} exp\left[-\left(\frac{x-a}{b}\right)^{c}\right] \quad (4)$$

Where: *n* the number of sample observations and x_i (cm) is the diameter of each tree. Other parameters are previously defined in equation 1.

Traditional Method of Moment (MMW): is based on the relationship between the parameters of the Weibull distribution and the first and second moments of the diameter distribution (i.e. the mean and variance, respectively). This method used by Stankova and Zlatanov, (2010); Gorgoso *et al.* (2012) and Ogana *et al.* (2015). It is expressed as:

$$b = \frac{\bar{d}-a}{\Gamma\left(1+\frac{1}{c}\right)} \quad (5)$$

$$\sigma^2 = \frac{(\bar{d}-a)^2}{\Gamma^2\left(1+\frac{1}{c}\right)}\left[\Gamma\left(1+\frac{2}{c}\right) - \Gamma^2\left(1+\frac{1}{c}\right)\right] \quad (6)$$

Where: *a* is the location parameter, *d* is the arithmetic mean diameter of the distribution, σ^2 is the variance and Γ*(i)* is the gamma function. Equation 6 was resolved by a bisection iterative procedure in SAS (SAS Institute Inc., 2003).

et al. (2003). It is expressed as:

$$\alpha_3 = \frac{\Gamma\left(1+\frac{3}{c}\right)-3\Gamma\left(1+\frac{1}{c}\right)\Gamma\left(1+\frac{2}{c}\right)+2\Gamma^3\left(1+\frac{1}{c}\right)}{\left[\Gamma\left(1+\frac{2}{c}\right)-\Gamma^2\left(1+\frac{1}{c}\right)\right]^{1.5}} \quad (7)$$

$$b = \frac{\sigma^2}{\sqrt{\left(\Gamma\left(1+\frac{2}{c}\right)-\Gamma^2\left(1+\frac{1}{c}\right)\right)}} \quad (8)$$

Where: α_3 is the sample skewness, other variables and parameters are previously defined in equation 5 and 6. Similarly log-likelihood function concerning the distribution's parameters and setting the expression to zero, then solve by the numerical iterative algorithm to give the estimates. Thus, the computation was done by minimizing the negative log-likelihood function of equation 1 using the 'optimum function' in R (R Core Team, 2016). The log-likelihood of Weibull function is expressed as: equation 7 was resolved by bisection iteration in SAS.

Percentiles Method (PW): 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:

$$\ln\hat{b} = \frac{\ln P_r - \frac{\ln P_t \cdot \ln(-\ln(1-r))}{\ln(-\ln(1-t))}}{1-\frac{\ln(-\ln(1-r))}{\ln(-\ln(1-t))}} \quad (9)$$

$$\hat{c} = \frac{\ln\left[\frac{\ln(1-r)}{\ln(1-t)}\right]}{\ln\left(\frac{P_r-a}{P_t-a}\right)} \quad (10)$$

Where: P_r and P_t are the sample percentiles with $0 < r < 1$ and $0 < t < 0$. The proposed values of $r = 0.97$ and $t = 0.17$ by Dubey (1967) was used for this study.

The estimation methods of the Weibull distribution across the different plot dimensions were assessed based on Kolmogorov-Smirnov (K-S), mean square error (MSE), mean absolute error (MAE) and bias. The smaller the values are, the better the estimation method.

Result

The summary statistics of the estimated parameters of the Weibull distribution fitted with MLE, percentile, traditional moments (MOM) and moments incorporating skewness (MIS) for the dimensions are presented in Table 2. The Weibull location parameter was constrained to a minimum diameter minus 0.5; as such, the four fitting methods had the same value for the location parameter for the three dimensions. The mean, maximum, minimum and standard deviation of the location parameter were 7.70, 12.10, 1.60 and 4.47 for 20m x 20m, 6.27, 10.40, 1.60 and 3.95 for 30m x 30m and 6.97, 11.10, 1.60 and 4.03 for 40 x 40m, respectively. The scale parameter had the highest mean, maximum and standard deviation values of 8.26, 13.83 and 3.43, respectively in 20m x 20m under the MIS method. MOM had the least value for the scale parameter under the same plot dimension. The Weibull shape parameter fitted with MIS had the largest mean, maximum, minimum and standard deviation of 2.73, 5.93, 1.18 and 1.31, respectively in 20m x 20m plot dimension. The method of MLE in 40m x 40m gave the smaller shape parameter value.**Table 2**: Descriptive statistics of the Weibull parameters across three plot dimension

Table 2 : Descriptive statistics of the Weibull parameters across three plot dimension

Plot dimension (m)	Method	Parameters	Mean	Maximum	Minimum	Std. Dev
20 x 20	Percentile	Location	7.70	12.10	1.60	4.47
		Scale	6.27	9.70	4.00	1.61
		Shape	2.07	3.27	1.30	0.54
30 x 30		Location	6.97	10.40	1.60	3.95
		Scale	7.39	11.60	5.20	2.39
		Shape	1.97	2.47	1.59	0.30
40 x 40		Location	6.97	11.10	1.60	4.03
		Scale	6.50	9.70	5.90	1.28
		Shape	1.88	2.29	1.71	0.21
20 x 20	MLE	Location	7.70	12.10	1.60	4.47
		Scale	6.20	8.60	3.95	1.28
		Shape	1.96	2.73	1.51	0.41
30 x 30		Location	6.97	10.40	1.60	3.95
		Scale	7.22	10.54	5.19	2.09
		Shape	1.95	2.48	1.62	0.29
40 x 40		Location	6.97	11.10	1.60	4.03
		Scale	6.70	10.42	5.68	1.61
		Shape	1.90	2.26	1.74	0.21
20 x 20	MIS	Location	7.70	12.10	1.60	1.29
		Scale	8.26	13.83	3.84	3.43
		Shape	2.73	5.93	1.18	1.31
30 x 30		Location	6.97	10.40	1.60	3.95
		Scale	8.22	11.44	5.47	1.94
		Shape	2.30	3.16	1.46	0.54
40 x 40		Location	6.97	11.10	1.60	4.03
		Scale	6.94	10.49	5.57	1.63
		Shape	1.98	2.56	1.55	0.31
20 x 20	MOM	Location	7.70	12.10	1.60	4.47
		Scale	6.17	8.60	3.83	1.33
		Shape	1.75	2.49	1.27	0.40
30 x 30		Location	6.97	10.40	1.60	3.95
		Scale	7.20	10.54	5.10	2.12
		Shape	1.76	2.24	1.50	0.28
40 x 40		Location	6.97	11.10	1.60	4.03
		Scale	6.66	10.41	5.64	1.61
		Shape	1.71	2.05	1.48	0.21

The overall fitting performance of the estimation methods across the different plot dimensions are presented in Table 3. The result showed that the performance of the estimation methods improved as the plot dimension increased i.e., from 20 x 20m to 40 x 40m. The best result was observed in 40 x 40m plot dimension for the four estimation methods considered. The method of MLE had the smallest values for the fit indices for the three plot dimensions (20 x 20m, 30 x 30m and 40 x 40m); as such, ranked best. The K-S, MSE, MAE and bias values for MLE were 0.08979, 0.00410, 0.04449 and 0.00525 in 20 x 20m, 0.06973, 0.00172, 0.03079 and 0.00100 in 30 x 30m, and 0.06997, 0.00126, 0.02618 and 0.00068 in 40 x 40m, respectively. This was followed by percentiles and MOM. The method of MIS had the worst performance for the three plot dimensions. The fit indices values for the method MIS were usually

Table 3: Comparison of Weibull parameters estimation methods across the three plot Dimension

Plot Dim.	Fit Indices	MLE	Percentile	MOM	MIS
20 x 20	K-S	0.08979	0.10045	0.09961	0.30329
	MSE	0.00410	0.00411	0.00453	0.00568
	MAE	0.04449	0.04456	0.04548	0.05051
	Bias	0.00525	0.00481	0.00655	0.01166
30 x 30	K-S	0.06973	0.09504	0.08409	0.18308
	MSE	0.00172	0.00173	0.00182	0.00195
	MAE	0.03079	0.03060	0.03136	0.03255
	Bias	0.00100	0.00113	0.00179	0.00089
40 x 40	K-S	0.06997	0.08353	0.07955	0.13724
	MSE	0.00126	0.00129	0.00141	0.00150
	MAE	0.02618	0.02683	0.02747	0.02837
	Bias	0.00068	0.00071	0.00128	0.00083

The graph of the diameter distributions for the three plot dimensions are presented in Figure 1-3. Plot 1 was used as the representative sample plot. The graph showed the observed and predicted relative frequency of tree using the four estimation methods per 1cm diameter class. All methods gave good prediction except the method of MIS. The MIS method had a different shape that is negatively skewed in 20 x 20m plot dimension (Figure 1) and a Gauss shape distribution in 30 x 30m plot dimension.

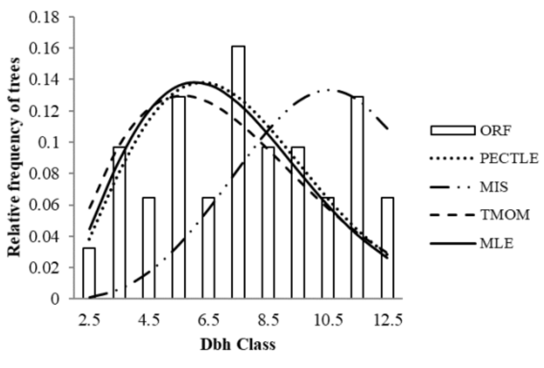

Figure 1: Observed relative frequency (ORF) and predicted relative frequency with MLE, MIS, traditional moment (TMOM) and percentile (PECTLE) methods in 20 x 20m plot dimension.

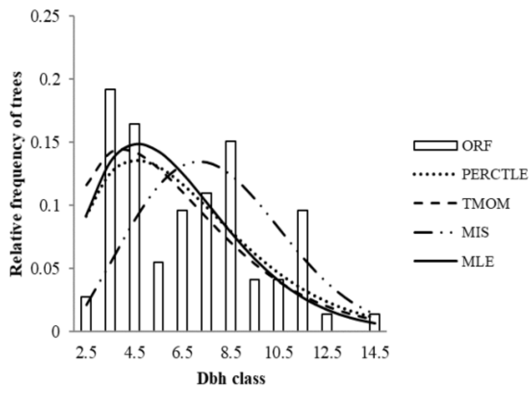

Figure 2: Observed relative frequency (ORF) and predicted relative frequency with MLE, MIS, traditional moment (TMOM) and percentile (PECTLE) methods in 30 x 30m plot dimension.

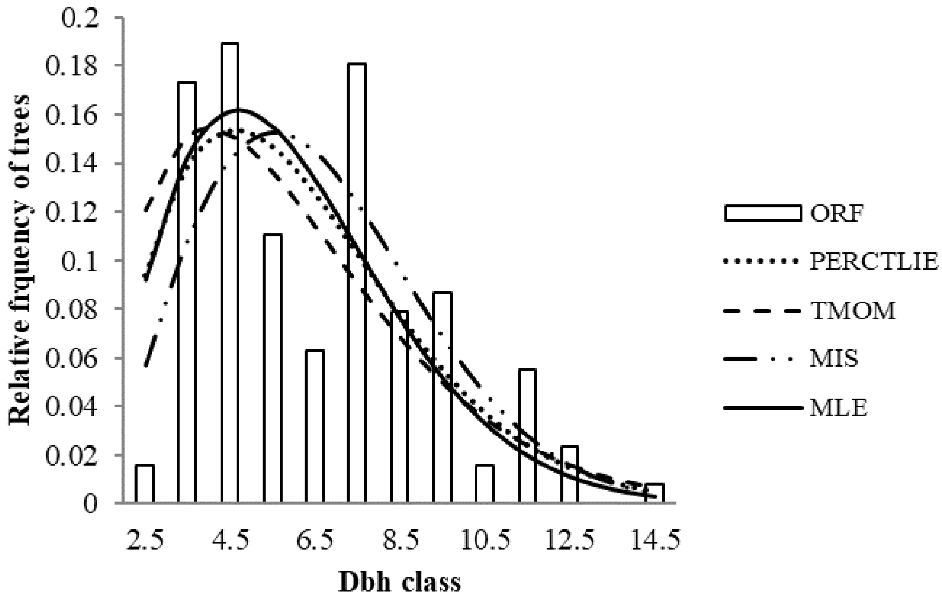

Figure 3: Observed relative frequency (ORF) and predicted relative frequency with MLE, MIS, traditional moment (TMOM) and percentile (PECTLE) methods in 40 x 40m plot dimension

Discussion

The effect of plot size on Weibull parameters has been investigated. The study reveals that the performance of the estimation methods improves as the size of the plot dimension increase. Minimum values for the fit indices were observed in 40 x 40m plot dimension. This is not surprising as there are more trees in the large plot dimension. In distribution modeling, the plot dimension is usually adjusted to have a minimum of 30 trees. This is in line with the central limit theory; which states that the distribution of the sum (or average) of the sample (*n*) will be approximately normal as the sample size increases, regardless of the underlying distribution. This study is in tandem with Saborowski (1994) who investigated the minimum sample size required to estimate the three parameters of the Weibull distributions. The author's simulation study showed that a sample size of 80 would produce satisfactory results. Furthermore, Shiver (1988) advocated a sample size of 50 for the estimated distribution to have less than 10% error in any diameter class. Sghaier et al. (2016) also varied plot dimension from 88 to 835m^2 to accommodate at least 40 trees. The authors reported better performance in a larger plot dimension.

The method of MLE had the overall best performance in the different plot dimensions. Thus, it is the most suitable method for describing the diameter distribution of the forest stand. The MLE is an accurate, efficient and unbiased estimator. The quality of the estimated parameters from MLE can be ascertain by the computation of the standard errors of the estimates. In fact, it is the only estimation method that gives a means of assessing the quality of the estimate. Lei (2008) asserted that "the MLE is a commonly used procedure for the Weibull distribution in forestry because of its desirable properties". Furthermore, Cao and McCarty (2006) reported that estimation of the parameters using maximum likelihood has been found to produce consistently better goodness-of-fit statistics compared to other methods, but it also puts the greatest demands on the computational resources. Although, this is not a major problem with the development of high-speed computers and open-source statistical tools. Ogana and Gorgoso-Varela (2015) also reported better performance with MLE compare to moment and percentile methods in their study on comparison of estimation methods for fitting Weibull distribution to the natural stand in Oluwa forest reserve.

The method of moments and percentile performed relatively well across the different plot dimension. However, these methods lack the desirable properties of the MLE method, except for the simplicity of computation. The method of MIS had the worst fits, especially for a smaller sample size.

Conclusion

This study has shown that plot size affects Weibull parameter estimates. The larger the size of the plot, the better the performance of the distribution. And in consequence, the better the overall prediction of the diameter distribution irrespective of the estimation method used. Thus, in any diameter distribution study, the plot dimension should be adjusted to accommodate a sizeable number of trees, especially in the stand with low density i.e., the number of tree per ha.

References

Balakrishanan, N. and Kateri, M. (2008). Maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. *Statistics & Probability Letters, 78*(17): 2971–2975.

Burk T. E, Newberry JD (1984). A simple algorithm for moment-based recovery of Weibull distribution parameters. Forest Science 30 (2): 329-332.

Cao QV, McCarty SM (2006). News methods for estimating parameters of Weibull functions to characterize future diameter distributions in forest stands. In: Proceedings of the "13th Biennial Southern Silvicultural Research Conference". General Technical Report SRS-92, Southern Research Station, USDA Forest Service, Asheville, NC, USA, pp. 338-340.

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

George F. (2014). Comparison of Shape and Scale Estimators of the Two-Parameter Weibull Distribution, *Journal of Modern Applied Statistical Methods* May 2014, Vol. 13, No. 1, 23-35.

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

Harter, H. L., and Moore A. H. (1965). Maximum likelihood estimation of the parameters of the gamma and Weibull populations from complete and censored samples. Technometrics 8: 135-175.

Jeffrey H. G. (2003). *Moment and maximum likelihood estimators for Weibull distributions under length- and area-biased sampling. Environmental and Ecological Statistics 10*:455-467

Jonhson N. L, Kotz S (1970). Distribution in statistics: continuous univariate distributions. Houghton Mifflin Co, Boston, MA, USA, vol 2, pp. 306.

Lei Y. (2008). Evaluation of three methods for estimating the Weibull distribution parameters of Chinese pine (*Pinus tabulaeformis*). Journal of Forest Science 54 (12): 566-571.

Ogana, F.N. and Gorgoso-Varela, J.J. (2015). Comparison of estimation methods for fitting Weibull distribution to the natural stand of Oluwa forest reserve, Ondo State, Nigeria. *Journal of Research in Forestry, Wildlife and Environment*, 7(2): 81-90.

Ogana, F.N., Osho, J.S.A. and Gorgoso-Varela, J.J. (2015). Comparison of beta, gamma and Weibull distributions for characterisng tree diameter in Oluwa forest reserve, Ondo State, Nigeria. *Journal of Natural Science Research* 5(4): 28-36.

Palahı M, Pukkala T, Blasco E, and 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

Saborowski, J. (1995). Minimum sample size for estimating the Weibull parameters of diameter distributions. Proc. IUFRO Symp., Stellenbosch, pp 1-7

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

Sghaier T, Cañellas I, Calama R, and Sánchez-González M. (2016). Modelling diameter distribution of *Tetraclinis articulata* in Tunisia using normal and Weibull distributions with parameters depending on stand variables. iForest 9: 702-709. – doi: 10.3832/ifor1688-008

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. and Zlatanov T.M. (2010). Modeling diameter distribution of Austrian black pine (Pinus nigra Arn) 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., Chuanmin 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.