Advances in Forestry Science
Original Article

Univariate models to represent the diametric distribution of thinned stand of *Tectona grandis* Linn.F

João Paulo Sardo Madi[1*] Diogo Guido Streck Vendruscolo[2] Carlos Alberto Silva[3] Mariana Peres de Lima Chaves e Carvalho[2] Samuel de Pádua Chaves e Carvalho[1]

[1]Federal University of Mato Grosso, College of Forestry Engineering - Laboratory of Forest Management, Av. Fernando Corrêa da Costa, 2367, Cuiabá-MT, CEP: 78.060-900.
[2] Federal University of Mato Grosso, College of Forestry Engineering, Av. Fernando Corrêa da Costa, 2367, Cuiabá-MT, CEP: 78.060-900.
[3] University of Idaho, College of Natural Resources, 975 W 6th, Moscow, ID 83844, USA.

*Author for correspondence: joaosardomadi@gmail.com
Received: 23 March 2017 / Accepted: 29 May 2017 / Published: 30 June 2017

Abstract
The aim of this study was to evaluate the performance of probabilistic distribution models for predicting the number of trees in a teak plantation located in the Nossa Senhora do Livramento city, state of Mato Groso, central region in Brazil. In the field, the diameters at breast height (DBH) of 203 trees of seminal origin, at 16 years of age, were measured in 2015. A descriptive analysis of the DBH was performed. Five models were used to fit a diametric distribution of the teak trees at the stand level: Normal, Normal Log, Gamma and Weibull with two parameters (2P) and three parameters (3P). For the purpose of comparison and selection of the best model, the Akaike Information Criterion (AIC) was used. After fitting the models, a simulated dataset was used to compute the accuracy of the number of trees estimated at stand level in each model. Among the fitted models, Weibull 3P was the one that presented the best fit, followed by the Log Normal, Gamma, Normal and Weibull 2P according to the AIC values. For the simulated dataset, the best result was Weibull 2P. When evaluated the accuracy of the model we found a maximum deviance to the Normal Distribution (27.78%).
Key words: Probability model, Parametric Statistics, Teak, AIC.

Introduction

The planted forest segment in Brazil plays an important role in the generation wealth, goods and services. In 2015 planted forest areas in Brazil exceed 7.8 million hectares, with an area increase of 0.8 % compared to 2014, which represents 0.91% of the national territory (Ibá 2016). Out of this extensive planted area, 5.6 million hectares are occupied by eucalypt, 1.6 million hectares by *Pinus* spp. and the remaining area (0.6 million hectares) by other species like acacia, rubber tree, paricá, and teak (Ibá 2016).

Among the species with potential for noble wood production, the *Tectona grandis* Linn.F. is a promising alternative. Although this specie present less representativeness in terms of national scenario, Ibá (2016) reported an increase in planted area of 25% between the 2010 and 2015. This increase can be justified by the resultant product, a high quality wood with high commercial value, intended for the naval construction and fine furniture. Furthermore, the use of this specie has become popular because of its increasing use for the confection of popular furniture, with more accessible prices, made of wood panels using wood from the first thinning (Figueiredo et al. 2005; Pelissari et al. 2014; Cunha Neto et al. 2016).

According to Nogueira et al. (2006) teak stands must have their thinning regimes realized through support decision system, by determining the optimum intensity and the technical age of thinning, and not empirically. For example, Campos and Leite (2013) showed that in order to determine optimum regimes of thinning, it is necessary to utilize models that include effects of stand density and productive capacity.

Among the decision support tools are the diameter distribution models, which pursue to represent the forest diameter distribution by means of probabilistic models. According to Robinson (2004), the diameter distribution is the frequency histogram of diameter at breast height, which can present different forms, for instance, negative exponential, unimodal, bimodal or irregular, varying according with the forest structure. According Binoti et al. (2016), based on information acquired from the diameter distribution of forest stands, inference can be drawn about the dynamics of growing and about the interaction between trees.

For the prediction of production under thinning regime, the diameter distribution models are the most indicated among the models of growth and yield (Leite et al. 2006). The representation of the diameter distribution by this models, with a minor level of error, allows for a sensible planning of industrial activities and support the optimization of investments and profitability of forest stands (Téo et al. 2012).

Studies about the diametric distribution in teak plantations, using probabilistic density functions, in similar manner as for other species, are found in Nogueira et al. (2006), Leite et al. (2006), Binoti et al. (2011), Binoti et al. (2012) and Medeiros (2016).

The aim of this study was to evaluate the performance of probabilistic distribution models to represent the diametric structure in a thinned teak stand in Mato Grosso, Brazil. In addition, we evaluated the performance of each model in predicting the number of large diameter trees submitting in a hypothetical scenario.

Material and methods

This study was carried out in stands of *T. grandis* located in the Nossa Senhora do Livramento city, state of Mato Groso, central region in Brazil. The local climate is of the type Aw, according to the Köppen classification, presenting well defined dry and rainy seasons (Peel et al. 2007), total and mean precipitation of 1,300 mm.year^{-1}, mean annual temperature of 25ºC (Alvares et al. 2013).

The teak seedlings were produced via seminal propagation and planted with an initial spacing of 3 m × 3 m. At the time of this study, the stand was 16 years old. In total, we applied four thinning treatments, the first at 4 years, the second at 8, the third at 10 and the last at 15 years, remaining 170 trees per hectare for the final cut.

In the field, we randomly measured the diameters at 1.30 m in height (DBH) of 203 trees, that represents 8% the total of trees in the stand. Posteriorly the trees were grouped in six diameter classes with 5 cm of intervals (Table 1).

Table 1 – Trees randomly measured in a thinned stand of *Tectona grandis* L.f.

Class	Inferior limit [cm]	Superior limit [cm]	Nº of observations	Relative frequency (%)
1	20	25	3	1.48
2	25	30	54	26.60
3	30	35	89	43.84
4	35	40	37	18.23
5	40	45	16	7.88
6	45	50	4	1.97
Total	-	-	203	100

In the first step a descriptive analysis of field data was performed, where the metrics mean, median and mode were obtained with the objective of interpreting a tendency of the interest variable, DBH. We also computed kurtosis and skewness for the DBH.

The second step was to fit a series of candidate probabilistic models to represent the diametric structure of the studied stand. The tested models were Normal, Log-Normal and Weibull with two parameters (Weibull 2P), Weibull with three parameters (Weibull 3P) and Gamma (Table 2) were the candidates to represent the diametric structure of stand in study.

Table 2 - Probabilistic density functions used in this study to model the diameter distribution of the DBH in a thinned stand of *Tectona grandis* L.f.

	Function	Conditions
Normal	$f(x) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right) e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma^2}\right)^2}$	$\sigma > 0$ $-\infty < x < +\infty$ $-\infty < \mu < +\infty$
Log Normal	$f(x) = \left(\frac{1}{x\left(\sqrt{2\pi\sigma^2}\right)}\right) e^{\left(-\frac{1}{2}\right)\frac{(\ln(x)-\mu)^2}{\sigma^2}}$	$x \geq a; \sigma > 0$ $-\infty < x < +\infty$ $-\infty < \mu < +\infty$
Weibull 2P	$f(x) = \frac{c}{b}\left(\frac{x}{b}\right)^{c-1} e^{-\left(\frac{x}{b}\right)^c}$	$x \geq 0$ $c, b > 0$
Weibull 3P	$f(x) = \frac{c}{b}\left(\frac{(x-a)}{b}\right)^{c-1} e^{-\left(\frac{x-a}{c}\right)^c}$	$x \geq 0$ $c, b, a > 0$
Gamma	$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}$	$x \geq a$ $-\infty < a < +\infty$ $\alpha, \beta > 0$

where: f (x) – density function of the variable x; x- Diameter of the center of class; μ- arithmetic mean; σ – standard deviation; σ^2 variance; π – constant "pi" (3.1416); a- minimum value for variable x; b – scale parameter, c – shape parameter.$\Gamma(\alpha) = \int_0^\infty u^{a-1}e^{-u}\delta u$.

The adherence of the fitted models was evaluating by the Akaike Information Criterion - AIC (Akaike 1974, 1981). According to Vismara (2009) this statistic has the advantaged to compare non-hierarchical models, considering them only as concurrent.

$$AIC = -2\ln(mv) + 2p$$

In which: ln = logarithmic to the napierian base; mv = value of likelihood function; p = number of parameters in the model.

All the procedures were implemented in R language (R Development Core Team 2015), using the function *fitdistr* from the package *MASS* (2002) and the package rLIDAR (Silva et al. 2015).

After the model selection, we created a hypothetic forest scenery from field data, with the goal of predicting the number of trees greater than 40 cm in DBH. In this case we considered a forest stand area of 500 hectares and tree density of 250 trees ha^{-1}, totalizing, therefore, 125,000 plants. Afterwards, the probabilities, for the different concurrent models in the study, were obtained in order to evaluate the difference between estimated proportions of trees for each model in each diameter distribution. The number of trees in each class in the simulation was then obtained by extrapolating the same proportion from field data. We tested this scenario only to evaluate the accuracy of the models when applied at large scale.

Results and discussion

The observed DBH of teak trees in this study ranged from 24.6 cm to 49.0 cm. The distribution of number of trees by class had a tendency towards normal distribution with values of mean, median and mode close to each other, namely, 33 cm, 32.1cm e 31.5 cm. Normal distribution of tree diameters are usually associated to even-aged stands. The distribution of trees in stand can be observed in Figure 1, which evidences the predominance of trees in the diametric classes 2, 3 and 4 (Table 1).

The value of kurtosis was 0.77, which infers that the DBH distribution has a sharper peak and a more compact shape than the normal distribution, similar to Téo et al. (2012). The skewness was characterized as to the right, presenting the value of 0.83, thus the tail of distribution is bigger to the right side than to left side, different from observed by Araújo Júnior et al. (2013).

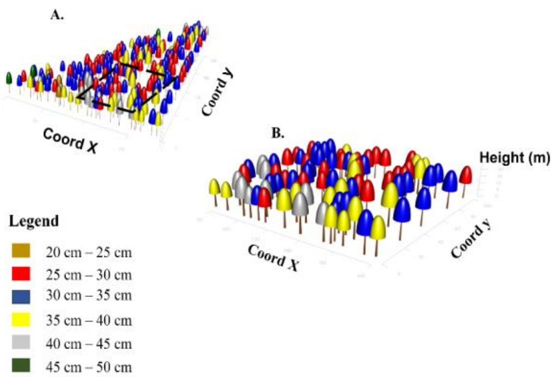

Figure 1. Tree distribution in the studied *Tectona grandis* stand, in Nossa Senhora do Livramento, Mao Grosso state, Brazil**,** A = Total number of trees in the stand; B = Enlarged view of the area delimited by the dotted polygon.

The estimated parameters for all the probabilistic density functions tested in this study and their respective AIC are shown in table 3. The best fit was found for the Weibull 3P function with AIC of 1,186.94, followed by the Log Normal function with AIC of 1,194.42, therefore presenting a difference of 7.5 units. When the difference in AIC is greater than 2 units, the models are considered different and the model with the smallest AIC is considered the most accurate (Burnham and Anderson 2002; Arnold 2010).

The good performance of the Weibull 3P model, was also observed in the study of Ribeiro et al.(2014), Binoti et al. (2012), Téo et al.(2012), working with the species *Eremanthus erythropappus DC.*, teak and *Pinus taeda* L. respectively. In these studies, the flexibility of the Weibull 3P model to represent the diametric structure among different species can be observed. The flexibility of the Weibull 3P function is an advantage that makes it suitable for the management of thinned stands (Leite et al. 2006), also allowing for the determination of the technical age of subsequent thinning as demonstrated by Nogueira et al. (2006).

Lana et al. (2013), obtained accurate results with Log Normal model, analyzing the diametric distribution of *Escheweilera ovata* in a fragment of dense umbrophilus forest, featuring among the models tested. Noting the flexibility of Log Normal model for even aged forests, present study, and uneven-aged forest.

The Weibull 2P was the diametric distribution model that obtained the poorest performance among the tested models. The difference in AIC values between this model and the Weibull 3P was 62.6. This lower quality in the fit is suggested by the absence of the location parameter, which was a significant parameter in the Weibull 3P function.

Table 3 - Parameters of the probabilistic density functions used in the study to model the diameter distribution in a Tectona grandis stand in Mato Grosso state, Brazil.

Function	Parameters	AIC	Dif.AIC
Weibull 3P	c =1.91 b=9.81 a =24.32	1,186.94	0.0
Log-Normal	$\ln(\mu) = 3.48$ $\ln(\sigma) = 0.14$	1,194.42	7.5
Gamma	$\alpha = 50.99$ $\beta = 1.54$	1,199.27	12.3
Normal	$\mu = 33.03$ $\sigma = 4.75$	1,212.56	25.6
Weibull 2P	$c = 6.72$ $b = 35.16$	1,249.53	62.6

Legend: μ= arithmetic mean; σ = standard deviation; b, c, a, α and β = estimated parameters; Dif.AIC = AIC rank where the difference between models were obtained.

In general, the Log Normal and Gamma functions (Figure 2) estimated accurately the first diametric class, overestimated intermediary classes and underestimated the last classes. The Weibull 3P, in relation of Log Normal and Gamma, presented inferior results in the first three classes, but it was more accurate in the other classes, although presents tendencies along of distribution as the other tested functions. Similar comportment of the Gamma and Log Normal functions can be noted in the study of Lana et al. (2013), by graphic analysis was observed the analogy performance in estimation of initial classes.

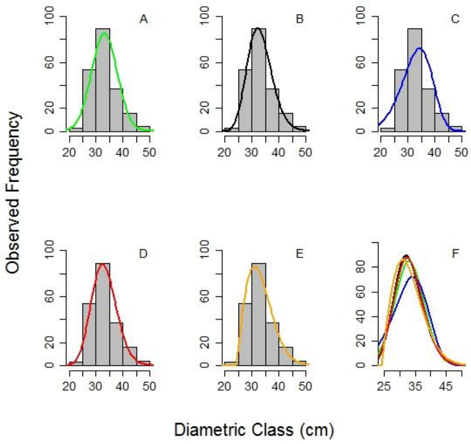

Figure 2. Diametric distribution histograms, with the curves from the models fitted to *Tectona grandis* stands, in Nossa Senhora do Livramento, Mato Grosso, Brazil. Each histogram represents one probability density function namely: A= Normal; B= Log-Normal; C = Weibull 2P; D = Gamma; E = Weibull 3P; F= overlapped estimated curves for all tested models.

The Normal and Weibull 2P functions presented the worst results according to AIC. Proves the minor efficiency of they in the adjust when observing the adjusted lines of models, comparing to Weibull 3P, Log Normal and Gamma presented variance high relatively, but Weibull 2P, in the last two classes, show itself more accurate. Nogueira et al. (2006) tested the Weibull 3P model also in a teak stand and obtain an efficient description of the stand diametric structure, suggesting, thus, that the presence of the location parameter generates results better fitted to the real data. In other words, this parameter improved directly the quality of the model fitting. Similarly, Téo et al (2012) and Ribeiro et al (2014), studying *P. taeda* and *E. erythropappus* stands, respectively, also found that the Weibull 3P was the best function for modeling the diameter distribution.

In the hypothetic scenery, all models underestimated the number of trees compared to the number of trees obtained with the proportion observed with the field sample (Table 4).

Table 4 - Estimated and observed number of trees with DBH greater than 40 cm in a hypothetical scenery of 125,000 trees.

Function	Nº of estimated trees	Nº of observed trees	Dif.	Dif. (%)
Weibull 3P	10,607	12,315	-3,000	-13.87
Log-Normal	9,201		-3,114	-25.29
Gamma	9,059		-3,256	-26.44
Normal	8,894		-3,421	-27.78
Weibull 2P	11,570		-741	-6.05

To consider the Weibull 2P model as the most accurate to situation of study, the model that obtain the worst performance was the normal, tending sub estimate 27.78%. Although the Weibull 2P model had the poorest performance in terms of AIC, it underestimated the number of trees above 40 cm of DBH by only 6.02%. As shown in Figure 2, this result can be explained by the fact this model performed badly for the smallest diameter classes but had significant improvement in the bigger diameter classes.

In the Figure 3, is showing the performance of the models to estimate the number of trees greater than 40 cm of DBH. This figure proves the better result of the Weibull's 2P model, where the curve is nearest of center than compared with the Weibull 3P for the values between 40 and 45 cm. Although Weibull's 3P model was superior than Weibull 2P in the class 45 – 50 cm, nevertheless the number of trees is smallest than the class 40 cm – 45 cm, resulting the advantage to Weibull 2P. So, in this case has shown, that the best model to the stand not always better when analyzed specifics cases like that.

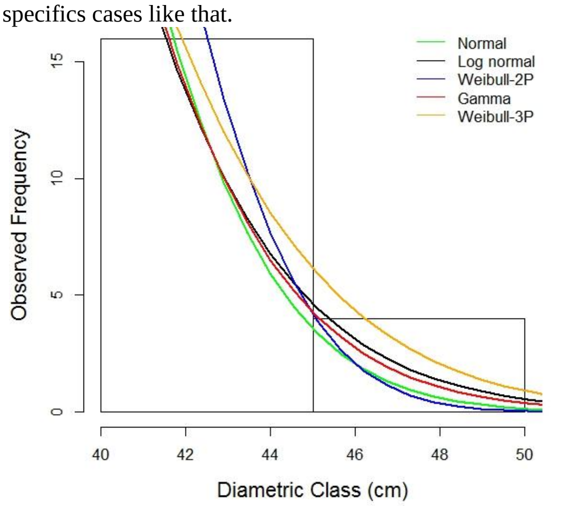

Figure 3. Diametric distribution greater than 40 cm DBH histograms, with the curves from the models fitted to *Tectona grandis* stands, in Nossa Senhora do Livramento, Mato Grosso, Brazil.

The estimation of other models followed the same tendency in relation to the AIC (Table 3), where the better model was Weibull 3P followed by Log Normal , Gamma and worst result for the hypothetical scenery was Normal.

This hypothetical scenery showed that not always the better model for the stand is the best model to estimate the number of trees of certain class.

Conclusions

The Weibull 3P model had the best performance in modeling the diameter distribution of the studied thinned stand of *T. grandis*. The selection of a sensible function for modeling stand diameter distribution is fundamental to quantify the tree assortments. As shown in this study, an error of up to 27.78% was obtained in estimating the assortment greater than 40 cm in DBH, which certainly would influence the decision making in the stand management. Based on the model accuracy obtained in the simulated scenario, this study suggested to use Weibull 2P in order to model diameter distribution of *T. grandis* trees with DBH greater than 40 cm. Furthermore, herein it was evident the need of selecting the most accurate model to each analyzed situation, whereas this is an important step to take decision of forest manager.

References

Alvares CA, Stape JL, Sentelhas PC, Gonçalves JLM, Sparovek G (2013) Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22 (6):711-728.

Akaike HA (1974) A new look at statistical model identification. *IEEE Transactions automatic control*, Tokio, 19 (6):717-723.

Akaike H (1981) Likelihood of a model and information criteria. *Journal of econometrics*, Amsterdan, 16 (1):3-14.

Araújo Júnior CA, Leite HG, Castro RVO, Binoti, DHB, Alcântara AEM, Binoti MLM da S (2013) Modelling the diameter distribution of eucalyptus stands using the gamma function. *Cerne*, 19(2):307-314. doi: http://dx.doi.org/10.1590/S0104-77602013000200015

Arnold TW (2010) Uninformative Parameters and Model Selection Using Akaike's Information Criterion. *Journal of Wildlife Management*, 74(6): 1175-1178. doi: 10.2193/2009-367

Binoti DHB, Leite HL, Silva MLM da S (2016) Computer System for Adjusting the Probability Density Functions. *Floram*, In press. doi: http://dx.doi.org/10.1590/2179-8087.081514

Binoti DHB, Binoti MLM da S, Leite HG, Fardin L, Oliveira JC (2012) Probability density function for description of diameter distribution in thinned stands of *Tectona grandis*. *Cerne*, 18(2):185-196. doi: http://dx.doi.org/10.1590/S0104-77602012000200002

Binoti DBH, Leite HG, Guimarães DP, Silva MLM da, Garcia SLR, Fardin LP (2011) Efficiency of the weibull and hyperbolic functions for describing the diametric distributions of *Tectona grandis* stands. *Árvore*, 35(2): 299-306. doi: http://dx.doi.org/10.1590/S0100-67622011000200014.

Burnham KP, Anderson DR (2002) *Model selection and multi model inference*: a practical information-theoretic approach. 2th edition. New York: Springer. 511p.

Campos JCC, Leite HG (2013) *Mensuração florestal*: perguntas e respostas. 4th edition. Viçosa: UFV. 605p.

Cunha Neto FV da, Vendruscolo DGS, Drescher, R (2016) Artificial form factor equations for *Tectona grandis* in different spacings. *African Journal of Agricultural Research*, 11(37): 3554-3561. doi: 10.5897/AJAR2016.11379

Embrapa - Empresa Brasileira de Pesquisa Agropecuária. Centro Nacional de Pesquisa de Solos (2006) *Sistema Brasileiro de Classificação de Solos*. 2th edition. Rio de Janeiro: Embrapa. 306p.

Figueiredo EO, Oliveira AD, Scolforo JRS (2005) economic analysis of unthinned stands of Tectona grandis L.f in the baixo rio acre microregion, Acre. *Cerne*, 11(4): 342-353. doi: http://www.redalyc.org/pdf/744/74411404.pdf>. 01 Dez. 2016.

Indústria Brasileira de árvores – IBÁ (2016) O setor brasileiro de árvores plantadas. 100p.

Lana MD, Brandão CFL e S, Péllico Netto S, Marangon LC, Retslaff FA de S (2013) Distribuição diamétrica de Escheweilera ovata em um fragmento de floresta ombrófila densa - Igarassu, PE. *Floresta*, 43(1):59-68. doi: http://dx.doi.org/10.5380/rf.v43i1.25252

Leite HG, Nogueira GS, Campos JCC, Takizawa FH, Rodrigues FL (2006) Um modelo de distribuição diamétrica para povoamentos de *Tectona grandis* submetidos a desbaste. *Árvore*, 30(1): 89-98. doi: http://dx.doi.org/10.1590/S0100-67622006000100011

Medeiros RA (2016) *Productive potential, management and experimentation in stands of Tectona grandis L.f. in the State of Mato Grosso*. Thesis, Universidade Federal de Viçosa. 182p.

Nogueira GS, Leite HG, Campos JCC, Takizawa FH, Couto L (2006) Evaluation of a diametric distribution model adjusted for thinned Tectona grandis stands. *Revista Árvore*, 30(3): 377-387. doi: http://dx.doi.org/10.1590/S0100-67622006000300008

Peel MC, Finlayson BL, McMahon TA (2007) Updated world map of the Köppen–Geiger climate classification. *Hydrology Earth System Sciences*, 11: 1633–1644.

Pelissari AL, Guimarães PP, Behling A, Ebling A (2014) Teak cultivation: specie characteristics for the formation and conduction of forest stands. *Agrarian Academy*, 1(1): 127-145.

R development core team. R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: <http://www.R-project.org/>. Access: Sep. 20 2015.

Ribeiro A, Ferraz Filho AC, Scolforo, JRS, Péllico Netto S, Machado AS (2014) Estrutura da distribuição diamétrica em plantio experimental de candeia (*Eremanthus erythropappus* (DC.) MacLeish). *Ciência Florestal*, 24 (4): 1055-1065. doi: http://dx.doi.org/10.1590/1980-509820142404024.

Robinson A (2004) Preserving correlation while modelling diameter distributions. *Canadian Journal of Forest Research*, 34(1): 221-232.

Silva CA, Crookston NL, Hudak AT, Vierling LA (2015) *rLiDAR*: An R package for reading, processing and visualizing LiDAR (Light Detection and Ranging) data, version 0.1. 2015. Accessed Oct. 15 2015, < http://cran.rproject.org/web/packages/rLiDAR/index.html>.

Téo SJ, Bianchi JC, Peloso A, Nava PR, Marcon A, Ehlers T, Costa RH da (2012) Desempenho de funções de densidade probabilísticas para descrever a distribuição diamétrica de *Pinus taeda*, na região de Caçador, SC. *Floresta*, 42(4):741-754.

Venables WN, Ripley BD (2002). *Modern Applied Statistics with S*. 4th Edition. New York: Springer .ISBN 0-387-95457-0.

Vismara ES (2009) *Biomass measurement and models selection for biomass equations*. Dissertation, Escola superior de Agricultura "Luiz de Queiróz", Universidade de São Paulo. 102p.