Allometric models to biomass in restoration areas in the Atlantic rain forest

Emanuel José Gomes de Araújo[1], Gabrielle Hambrecht Loureiro[2],
Carlos Roberto Sanquetta[3], Mateus Niroh Inoue Sanquetta[3],
Ana Paula Dalla Corte[3], Sylvio Péllico Netto[3], Alexandre Behling[3]

[1]Departamento de Silvicultura, Universidade Federal Rural do Rio de Janeiro – UFRRJ, Seropédica/RJ, Brazil
[2]Suzano Papel e Celulose, Limeira, SP, Brazil
[3]Departamento de Ciências Florestais, Universidade Federal do Paraná – UFPR, Curitiba/PR, Brazil

ABSTRACT

The objective of the study was to present mathematical models and strategies for fitting equations to estimate dry biomass for tree species in forest restoration areas. The presence of outliers was analyzed in each fitted equation using values of the matrix H, leverage points, means of standard and studentized residuals, and of influential points through DFFITS, DFBETAS and COOK distance values. Furthermore, the normality, homoscedasticity and independence of residuals were checked. The accuracy of the fitted equations was evaluated by means of the $R^2_{adj.}$, S_{yx}, analysis of residuals, and AIC and BIC criteria. The results showed that the model for estimating dry biomass as a function of the variables Dc^2, DBH^2, Hc^2 and DBH provides the more accurate solution, with $S_{yx} = 40.91\%$ and $R^2_{adj.} = 0.92$. We concluded that the performance of this equation improves when adjusted to data stratified by classes of height-diameter ratio, which reduces the value of the estimated error.

Keywords: modeling, dry weight, carbon, allometry.

Araújo EJG, Loureiro GH, Sanquetta CR, Sanquetta MNI, Corte APD,
2/13 Péllico Netto S et al.

Floresta e Ambiente 2018; 25(1): e20160193

1. INTRODUCTION

Forests are important for fixing carbon, absorbing it from the atmosphere in the form of carbon dioxide and converting it into carbohydrates to form wood tissues, leaves, seeds and fruits (Rochadelli, 2001). This process promotes the reduction of the concentration of greenhouse gases (GHG) and, consequently, of the risk of global warming caused by their high levels in the atmosphere.

Forests play an important role in the global carbon cycle, being responsible for about one-third of emissions of anthropogenic carbon dioxide (CO_2) to the atmosphere. However, human activities in forests have also been a source of emission of carbon dioxide into the atmosphere, with deforestation contributing with about a fifth of annual anthropic emissions of GHG, especially in the tropics. Planting forests for the recovery of degraded areas has become one of the main measures to mitigate global warming, because carbon is stored permanently.

Measuring how well forests store carbon is essential because it allows not only knowing the accumulated biomass but also highlighting the importance of it within the climate change scenario. Besides providing information on the quantity of CO_2 released in the case of slash-and-burn, biomass estimates allow the monitoring and evaluation of export of nutrients during forest exploitation (Higuchi et al., 1998). The term biomass corresponds to the organic matter stored in a particular ecosystem, composed mainly of carbon structures and mineral elements (Larcher, 1986).

Forest biomass can be quantified by direct and indirect methods. The direct method is destructive, because the tree is chopped down and its components are weighted. The indirect method is performed by means of estimates, which can be made with satellite images (Goetz et al., 1999; Chen et al., 2003; Drolet et al., 2005; Watzlawick et al., 2009; Corte & Sanquetta, 2007; Frankenberg et al., 2011; Song et al., 2013) or via mathematical models (Rezende et al., 2006; Urbano et al., 2008; Morais et al., 2013b; Melo et al., 2014). In the case of mathematical models, the biomass is estimated as a function of variables easily measured in forest inventories. Linear and non-linear mathematical models can be used in such estimations, but non-linear models are preferred (Baskerville, 1972; Ketterings et al.,

2001; Gehring et al., 2004; Sanquetta et al., 2004; Segura & Kanninen, 2005; Wang, 2006; Soares et al., 2006; Kenzo et al., 2009; Moore, 2010; Addo-Fordjour & Rahmad, 2013), because their application enables the estimation in more extensive forests areas.

However, modeling should be done with caution because various factors affect biomass production. Some examples are the species composition, stage of forest development, nutritional status and edaphological characteristics (Larcher, 1986), in addition to factors related to respiration and photosynthesis (Kramer & Koslowski, 1972) and the ecological groups to which the species belong to (Luo et al., 2014; Barbosa et al., 2014). Vogel et al. (2006) still mention that biomass estimation of a forest depends on its vegetation type and its location.

The objective of this work was to present alternative mathematical models and strategies for fitting equations to estimate the dry biomass of species in forest restoration areas.

2. MATERIAL AND METHODS

2.1. Data collection

The data for this study were collected in a restoration forest area in the municipality of Seropédica, RJ, in the coordinates 22° 43' 34" S and 43° 38' 34" W. The climate of the region, according to the Köppen classification is Aw (Brasil, 1992), is tropical with rainy summer. According to data collected in the last 20 years by the PESAGRO-RJ meteorological station, the closest to the place of study, the average annual rainfall is 1,245 mm and the average annual temperature is 23.7 °C, with a relative humidity of 69%.

The data used in this study were obtained from 111 trees, distributed in 50 different species and contemplating all the diametric structure of individuals measured in the inventory. The species sampled were: *Acacia polyphylla* DC., *Aegiphila sellowiana* Cham, *Albizia polycephala*, *Alibertia concolor* K. Schum, *Anadenanthera macrocarpa* (Benth.) Brenan, *Anadenanthera falcata* (Benth.), *Bauhinia forficata* Link., *Bixa orelana* L, *Caesalpinia echinata* Lam, *Caesalpinia ferrea* Mart, *Caesalpinia peltophoroides* (Benth.), *Cecropia pachystachya* Trec, *Cedrela fissilis* (Vell.), *Centrolobium robustum* (Vell) Mart. ex Benth., *Cordia superba* Cham, *Cordia*

trichotoma (Vell.) Arrab. ex Steud., *Croton urucurana* Baill., *Cytharexyllum myrianthum* Cham. , *Dalbergia nigra* (Vell), *Enterolobium contortisiliquum* (Vell.) Morong, *Erythrina mulungu* Mart., *Eugenia uniflora* L., *Genipa americana* L., *Gochnatia polymorpha* (Less) Cabr., *Guarea guidonia* (L.) Sleumer, *Guazuma ulmifolia* Lam., *Helietta apiculata* Benth., *Inga uruguensis* Hooker at Arnott, *Jacaranda mimosifolia* D. Don, *Lafoensia glyptocarpa* Kohne, *Lithraea molleoides* (Vell.) Engl., *Luehea divaricata* Mart., *Mabea fistulifera* Mart., *Machaerium stipitatum* (DC.) Vog, *Melia azedarach* L., *Mimosa caesalpiniaefolia* Benth., *Myroxylon peruiferum* L. f, *Peltophorum dubium* (Spreng), *Piptadenia gonoacantha* (Mart.) J.F. Macbr., *Psidium guajava*, *Pterocarpus violaceus* Vogel, *Samanea saman* (Jacq.) Merr., *Schinus terebinthifolius* Raddi, *Schizolobium parahyba* (Vell.) Blake, *Syzygium malaccense* (L.), *Handroanthus chrysotricha* (Mart. ex DC.), *Handroanthus impetiginosa* (Mart ex, DC.), *Handroanthus tuberculosis* (Mart. ex DC.), *Trema micrantha* (L.) Blume and *Triplaris americana* L.

Each tree was classified in their respective ecological group and measured for canopy diameter (CD), diameter at 1.30 m above ground (DBH) and total height (Ht). Each tree was felled and its commercial height (Hc) up to a stem diameter equal to 5 cm was measured.

Different fractions of forest biomass from each tree were separated using a complete dissection technique and weighted. Each tree was separated into the following compartments: bole, live branches, dried branches, leaves and roots. All components of the tree were separated and weighed on a scale in the field. For the root sampling, a 0.50 m deep trench was opened around each felled tree and the roots with a diameter equal or greater than 1 cm were pulled out. The removed roots were cleaned and then weighed. Finally, the total green weight of each felled tree was obtained by the sum of all compartment weights.

After the weighing (total green weight) of each compartment in the field, samples were collected and stored in plastic bags to prevent moisture loss and sent to the laboratory for analysis. In the laboratory, the samples were weighed in a precision scale to obtain the green weight of the sample. After drying, the dry weight of the material was estimated and, consequently, the moisture content according to the following Equation 1:

$$MC\% = \left(\frac{GW - DW}{DW} \right).100 \tag{1}$$

Where:

MC = moisture content (%); DW= dry weight (kg); GW= green weight (kg).

The dry weight of each compartment was estimated by multiplying the total green weight obtained in the field by its moisture content (%). The total dry biomass of each tree was obtained by the sum of the dry weights of all compartments.

During the felling of 111 trees in the field, wood disks were removed at DBH level in each tree to quantify the wood basic density (d), which was determined according to the Equation 2:

$$d = \frac{M}{V} \tag{2}$$

where:

d = wood basic density (g.cm⁻³); M = dry mass (g); V = saturated volume (cm³).

2.2. Mathematical modeling

Sixteen mathematical models were fitted by means of 5 different strategies in order to estimate the total dry biomass as a function of the variables: diameter at 1.30 m above ground (DBH), total height (Ht), commercial height (Hc), average canopy diameter (CD) and wood basic density (d) (Table 1). These parameters were chosen because they are the basic properties of a tree, besides being easy to measure. Density was included as an auxiliary variable attempting to improve the performance of the models to estimate tree biomass. The estimated descriptive statistics for each variable analyzed were mean, standard deviation, coefficient of variation, mean standard deviation, absolute and percentage sampling error, and confidence interval (Péllico & Brena, 1997).

In the strategy I, the single-entry models 1 to 6 were fitted only in function of the DBH. In the double-entry models 7 to 12, the adjustment was done for total dry biomass as a function of the independent variables DBH and Ht.

In the strategy II, the models 1 to 6 were fitted only in function of the basic density of the wood. The models 7 to 12 were fitted using the basic wood density and the DBH as independent variables. In the third strategy, the models 13 and 14 proposed by Chave et al. (2005) were considered and denominated Chave 1 and 2, respectively. These models are appropriate for inclusion of wood

Araújo EJG, Loureiro GH, Sanquetta CR, Sanquetta MNI, Corte APD,
4/13 Péllico Netto S et al.

Floresta e Ambiente 2018; 25(1): e20160193

Table 1. Adjusted mathematical models to estimate total biomass of trees in forest restoration areas.

Nº	Model	Mathematical Structure
1	Linear	$y = \beta_0 + \beta_1 x + e$
2	Trorey	$y = \beta_0 + \beta_1 x + \beta_2 x^2 + e$
3	Assmann	$y = \beta_0 + \beta_1 (1/x) + e$
4	Henricksen	$y = \beta_0 + \beta_1 \ln(DBH) + e$
5	Stofells	$\ln(y) = \beta_0 + \beta_1 \ln(DBH) + e$
6	Curtis	$\ln(y) = \beta_0 + \beta_1 (1/DBH) + e$
7	Schumacher-Hall	$\ln(y) = \beta_0 + \beta_1 \ln(DBH) + \beta_2 \ln(Ht) + e$
8	Spurr	$y = \beta_0 + \beta_1 \left(DBH^2 . Ht_2\right) + e$
9	Meyer	$y = \beta_0 + \beta_1 DBH + \beta_2 DBH^2 + \beta_3 DBHHt + \beta_4 DBH^2 Ht + e$
10	Stoate	$y = \beta_0 + \beta_1 DBH^2 + \beta_2 Ht^2 + \beta_3 \left(DBH^2 Ht\right) + e$
11	Sanquetta	$y = \beta_0 + \beta_1 DBH^2 + \beta_2 DBH^2 Ht + e$
12	Naslund	$y = \beta_0 + \beta_1 DBH^2 + \beta_2 DBH^2 Ht + \beta_3 \left(DBH . Ht^2\right) + \beta_4 Ht^2 + e$
13	Chave 1	$\ln(y) = \beta_0 + \beta_1 \ln(DBH) + \beta_2 \ln(Ht) + \beta_3 \ln(d) + e$
14	Chave 2	$\ln(y) = \beta_0 + \beta_1 \ln(DBH) + \beta_2 \left(\ln(DBH)\right)^2 + \beta_3 \left(\ln(DBH)\right)^3 + \beta_4 \ln(d) + e$
15	Type 1Model	$y = \beta_0 + \beta_1 d + \beta_2 DBH + \beta_3 Hc + \beta_4 Dc + e$
16	Type 2 Model	$y = \beta_0 + \beta_1 DBH + \beta_2 DBH^2 + \beta_3 Hc^2 + \beta_4 Dc^2 + e$

Y = total dry biomass; x_i = independent variable, which can assume the values of DBH = diameter at 1.30 m from the ground (cm); Ht = total height (m); Hc = commercial height (m); CD = average canopy diameter (cm); d = basic density (g.cm^{-3}) and their transformations, namely, inverse, squared, logarithmic, square of the inverse, logarithm of the inverse and all combinations of them; ln = natural logarithm; e = estimation error. The models 1 to 12 were adapted from Scolforo (2005) and the models 13 and 14 from Chave et al. (2005).

basic density of total dry biomass as an independent variable, according to Nogueira et al. (2008).

In the strategy IV, a modeling via stepwise procedure was performed using the variables DBH, Ht, Hc, CD and d in their normal and operational forms. In the Type 1 model, variables were used only in their normal form. In the Type 2 model, the transformations made were: inverse, squared, logarithmic, square of the inverse, logarithm of the inverse and all combinations of them, totaling 66 possible variables. In both cases, after analyzing the correlation matrix between variables, it was possible selecting the best mathematical combination by the stepwise method.

Finally, the model with the best performance used in the previously described strategies was applied in

the strategy V. For this reason, the fit of the model was carried out with the data stratified into classes of diameter, height-diameter ratio, and ecological group of species. The diameter classes were: DBH < 5 cm (Group A), with 67 trees; 5 ≤ DBH < 9 cm (Group B), with 37 trees; and DBH ≥ 9 cm, with 7 trees. The Ht/DBH ratio classes were: Ht/DBH < 1 (Group C), with 39 trees; and Ht/DBH > 1 (Group D), with 72 trees. Finally, the ecological group classes were: Pioneer (Group F), with 80 trees; and Secondary (Group G) species, with 31 trees. As there was no combination of stratification criteria, the number of trees used for adjustment in each class was considered sufficient, ensuring the control of tree variability in the diameter classes.

Significant differences between the equations fitted with the entire data set and appropriately stratified, considering the different strategies previously mentioned, were examined by means of an identity test between the models (Regazzi, 1992).

The accuracy of the equations, adjusted in all cases, was evaluated by means of the adjusted coefficient of determination ($R^2_{adj.}$) (Equation 3), standard error of the estimate in percentage (S_{yx}%) (Equation 4), graphical analysis of residuals and significance of the coefficients by means of a t-test, at 95% probability level. In addition, the Akaike information criterion (Akaike, 1998) (Equation 5) and the Bayesian information criterion (Schwarz, 1978) (Equation 6) were evaluated.

$$R^2_{aj.} = 1 - \left[\left(\frac{SQres.}{SQtotal} \right) \left(\frac{(n-1)}{(n-p)} \right) \right] \qquad (3)$$

$$s_{yx} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-p-1}} \therefore s_{yx}(\%) = \left(\frac{s_{yx}}{\overline{y}} \right) \cdot 100 \qquad (4)$$

$$AIC = -2n \left(\frac{-n}{2} \ln \left(\frac{SQres}{n} \right) \right) + 2p \qquad (5)$$

$$BIC = -2n \left(\frac{-n}{2} \ln \left(\frac{SQres}{n} \right) \right) + \ln(n) p \qquad (6)$$

Where: $R^2_{adj.}$ = adjusted coefficient of determination; S_{yx} = absolute standard error of the estimate; S_{yx}% = standard error of the estimate in percentage; y_i = observed dry biomass (kg); $\hat{y}$ = estimated dry biomass (kg); $\overline{y}$ = arithmetic mean of observed dry biomass (kg); n = number of observed data; p = number of coefficients of the model; $SQ_{res.}$ = sum of squares of residuals; SQ_{total} = total sum of squares; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; e_i = error estimate and ln = natural logarithm.

In addition, a diagnosis of the models was done for each adjusted situation, as described in Scolforo (2005). The projection matrix H was analyzed for leverage points to check the occurrence of outliers in the data base, and the standard residuals (rp) and Studentized residuals (rs) were used to detect outliers in the total dry biomass variable. The significant presence of discrepant observations was verified by the Bonferroni's test at 95% probability level. The detection of influential points on model fitness was performed by means of DFFITS, DFBETAS and COOK distance values. The Shapiro-Wilk's test, White test, and Durbin-Watson statistics were applied to check normality, homoscedasticity and independence of the residuals, respectively, at 95% probability level, applying the box-cox transformation when needed. When appropriated, multicollinearity was also examined to check the existence of correlation between the independent variables. The adjustments of the models as well as the statistical analysis applied in this research were done with the softwares R (R Development Core Team, 2001) and SAS (*Statistical Analysis System*).

3. RESULTS AND DISCUSSION

The statistics in Table 2 indicate that the variable total dry biomass (DW) had one of the greatest variability among tree volumes in the forest restoration area. The coefficient of variation for this variable was 140% and the sampling error approximately 20%. An error less than 10% indicates that the sample is representative of the population to estimate the variable in question (Péllico Netto & Brena, 1997). However, the sampling error obtained in this research can be considered low in view of the high variability of dry biomass among

Table 2. Descriptive statistics for the selected variables.

Statistics	DBH (cm)	Ht (m)	Hc (m)	Dc (cm)	PS (kg)
$\overline{x}$	6.17	5.85	3.08	3.33	16.98
s^2	19.62	5.38	3.84	2.63	572.78
s	4.43	2.32	1.96	1.62	23.93
CV	71.85	39.66	63.70	48.72	140.99
$s_{\overline{x}}$	0.33	0.17	0.15	0.12	1.78
e	0.65	0.34	0.29	0.24	3.52
$e(\%)$	10.57	5.83	9.37	7.17	20.74
IC	$5.51 \leq M \leq 6.82$	$5.51 \leq M \leq 6.19$	$2.79 \leq M \leq 3.37$	$3.09 \leq M \leq 3.57$	$13.45 \leq M \leq 20.50$

Araújo EJG, Loureiro GH, Sanquetta CR, Sanquetta MNI, Corte APD,
Péllico Netto S et al.

6/13

Floresta e Ambiente 2018; 25(1): e20160193

trees in this forest community (Melo et al., 2014; Morais et al., 2013a).

The error was less than 10% for the variables DBH, Ht and Hc, confirming the representativeness of the sample. The low mean values of these variables characterize young stands in the development phase. Furthermore, all trees scaled in the study area belong to the ecological group of pioneer or secondary species.

The presence of outliers and influential points was analyzed in all the adjustment strategies, in addition to normality, homoscedasticity and independence of residues. Only in the case of the Type 2 model adjusted according to the strategy V, the values of the matrix H, Standardized (rp) and Studentize (rs) residuals indicated three observations as possible leverage points or outliers in relation to the dependent variable Dry Weight. This is owing to values of the matrix H above the critical value of 0.1351 (3p/n) and of residues within the acceptable limit ($-3 \leq rp$ or $rs \leq 3$). However, the non-significant (p-value > 0.05) result of the Bonferroni test at 5% probability indicated that these observations are not considered as discrepant values. The DFFITS, DFBETAS and COOK distance values indicated some influential points in the fit of the Type 2 model. An alternative to improve the fit of the model in the presence of these points is to use robust regression (Cunha et al., 2002).

The coefficients and statistics to adjust the models used in the strategy I are presented in the Table 3. The lower values of adjusted coefficients of determination were obtained when DBH alone was used to explain the total dry biomass of trees (models 1 to 6). The equations derived from these models also provide the largest

errors of estimate (S_{yx}), reaching values greater than 100%, characterizing low accuracy, in addition to high AIC and BIC values. When the variable total height was included in the models (7 and 12), there was an increase in accuracy of the estimated dependent variable, which was characterized by a decrease of AIC and BIC values and the standard error of the estimate, which varied between 60 and 80%. However, this accuracy is still low when compared with that obtained in other reaserchs in which double entry variables were applied in models for estimating tree biomass in native forests (Watzlawick et al., 2009; Rezende et al., 2006). An alternative would be to apply the models to the species separately, thus reducing the variability caused by the characteristics of different species. Good results can be obtained even when the models include only the diameter (Litton & Kauffman, 2008; Melo et al., 2014). In the work of Melo et al. (2014), the best models to estimate the biomass of "caixeta" (*Tabebuia cassinoides* (Lam.) DC.) were the ones that included only the diameter as independent variable.

For the strategy I, the Schumacher-Hall model was the one that provided the best estimates of dry biomass, even though it did not provide the best adjusted statistics and the largest adjusted coefficient of determination ($R^2_{adj.}$). This model was used by Scolforo et al. (2008) to estimate the amount of biomass and carbon in different forest physiognomies in the State of Minas Gerais, demonstrating its applicability. This model was selected by means of graphical analysis of residuals, which presents a distribution without trends when compared to other models (Figure 1). The Shapiro-Wilk's test (normality) and Durbin-Watson

Table 3. Coefficients and adjusted statistics of the single- and double-entry models adjusted according to the strategy I.

Nº	β_0	β_1	β_2	β_3	β_4	S_{yx} %	$R^2_{adj.}$	AIC	BIC
1*	-10.58*	4.60*				81.69	0.67	59014.68	59020.10
2*	-0.02	1.42	0.16*			75.80	0.71	57057.78	57065.68
3*	32.69*	-76.62*				123.84	0.23	69265.48	69270.90
4*	-23.18*	24.30*				106.60	0.43	65573.48	65578.89
5*	-0.47*	1.62*				79.59	0.68	58372.67	58378.09
6*	3.59*	-6.50*				121.47	0.26	68789.32	68794.74
7*	**-2.04***	**0.91***	**1.62***			**83.16**	**0.65**	**59344.97**	**59352.87**
8*	8.44*	0.02*				78.64	0.69	58077.48	58082.90
9*	-11.30*	-0.48	0.12	3.93*	0.00	64.33	0.79	52788.96	52801.94
10*	-1.97	0.14*	0.31*	0.00		62.72	0.80	52277.86	52288.32
11*	5.93*	0.16*	0.00			76.59	0.71	57312.86	57320.76
12*	-1.49	0.15*	-0.01	0.01	0.26*	62.95	0.80	52254.92	52267.90

*Significant at 95% probability level.

statistic were not significant (p-value > 0.05), indicating that the residuals are normally distributed and are not correlated among themselves. The White's test indicated the presence of heteroscedasticity (p-value < 0.05), resulting in inefficient and biased estimators for the variances, consequently invalidating the significance tests for the regression equation. A solution to this problem was to use weighted regression using the weight $w_i = \dfrac{1}{DBH.Ht_i^2}$. The weighted equation obtained was $\ln(y) = -2.0375 + 0.9137\ln(DBH) + 1.6240\ln(Ht)$ with the adjusted statistics ($R^2_{adj.} = 0.65$) and ($S_{yx}\% = 83.16\%$).

When DBH and the wood basic density were used as independent variables in the modeling (strategy II), the performance of the models 7 to 12 was similar to that observed in the strategy I, when evaluating the

obtained statistics for the adjustments. Strong emphasis should be put on single-entry models (models 1 to 6), revealing that the variable wood basic density should not be used alone to explain the total dry biomass. This was demonstrated by the substantial increase of AIC and BIC values, errors greater than 140% and the $R^2_{adj.}$ values tending to zero (Table 4). This implies that the changes in the variable basic density explained little of the dry biomass variation. For this strategy, the best performance was also achieved with the adjusted model of Schumacher-Hall, as seen in the graphical distribution of residuals (Figure 2).

The adjustment in strategy III (Table 5), the models proposed by Chave et al. (2005), which also include wood basic density as an independent variable, presented
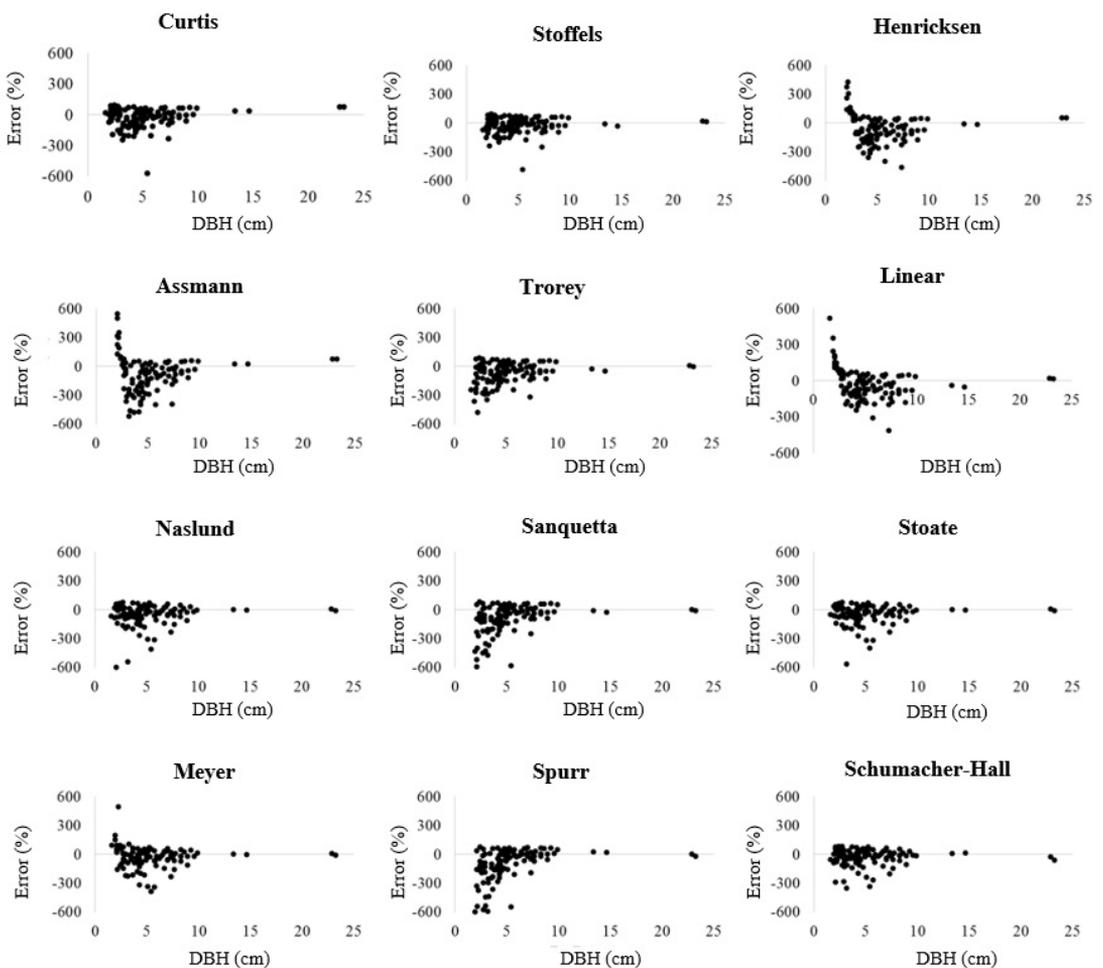
Figure 1. Residual graphic distribution of the single- and double-entry models adjusted for estimating total biomass in the strategy I.

Araújo EJG, Loureiro GH, Sanquetta CR, Sanquetta MNI, Corte APD,
8/13 Péllico Netto S et al.

Floresta e Ambiente 2018; 25(1): e20160193

Table 4. Coefficients and adjusted statistics of the single- and double-entry models adjusted according to the Strategy II.

Nº	β_0	β_1	β_2	β_3	β_4	S_{yx} %	$R^2_{adj.}$	AIC	BIC
1	5.42	15.29				140.95	0.00	72454.77	72460.19
2	-21.93	124.08	-100.14			140.39	0.01	72246.56	72254.46
3	22.34*	-4.29				140.42	0.01	72362.53	72367.95
4	19.51*	8.85				140.62	0.00	72397.79	72403.21
5	2.33*	0.54				148.67	0.00	73769.78	73775.20
6	2.44*	-0.23				148.65	0.00	73765.34	73770.76
7*	**0.07**	**1.71***	**1.03***			**76.82**	**0.70**	**57389.41**	**57397.31**
8*	5.51*	0.40*				75.71	0.71	57140.47	57145.89
9*	-13.68*	1.99*	0.12	21.86*	0.03	72.41	0.73	55706.40	55719.38
10*	0.11	0.20*	16.22*	0.03		74.46	0.72	56505.78	56516.25
11*	5.32*	0.07	0.27			75.77	0.71	57049.83	57057.74
12*	-1.06	0.29*	-0.24	4.83*	2.94	73.28	0.73	55999.43	56012.41

*Significant at 95% probability level.
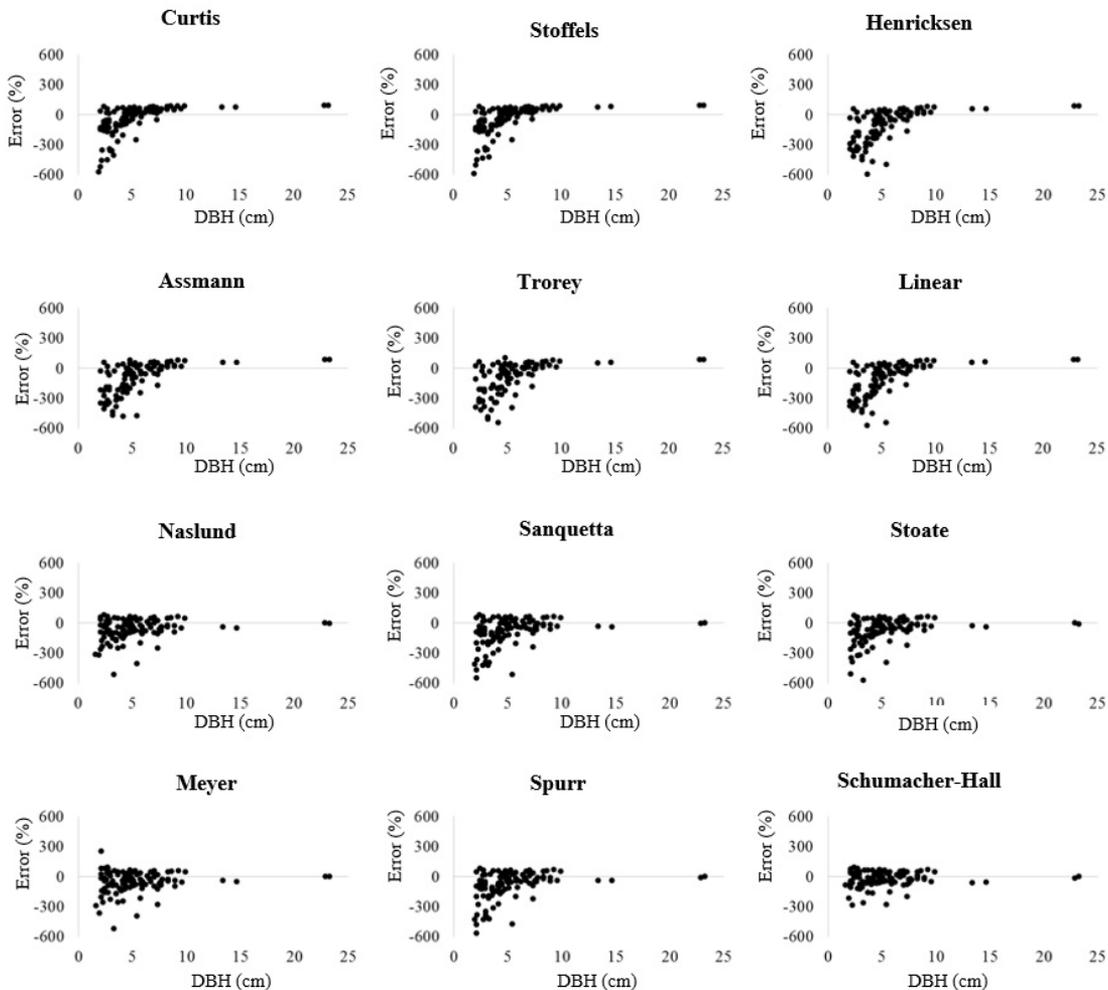
Figure 2. Residual graphic distribution of traditional models adjusted for estimating total dry biomass in the strategy II.

adjusted statistics similar to the Schumacher-Hall model used in strategies I and II. As a comparison, it was clear that the Schumacher-Hall model fitted as a function of DBH and Ht (Strategy I) stood out for providing the best results for the standard error of the estimate (67.42%) and the adjusted coefficient of determination (0.77). Thus, the inclusion of wood basic density of the species, regardless of the model adopted, is not an technically feasible option when the intention is to estimate the dry biomass in forest restoration areas. This is why the model that included this variable did not generate significant gains in accuracy, as showed in the graphical analysis of residuals (Figure 3). Furthermore, this variable is difficult to

be measured, as highlighted by Chave et al. (2005), justifying the use of DBH and Ht alone as they are more accessible to be measured.

The best equations were derived from modeling by the stepwise method for variable selection, in particular in the Type 2 model. Of the 66 evaluated variables, DBH, DBH^2, $(Hc)^2$ and $(CD)^2$ were selected to compose the model and resulted in an adjusted equation with coefficient of determination $(R^2_{adj.})= 0.92$ and the lowest standard error of the estimate (40.91%). The residuals of this equation presented homogeneous distribution. However, this equation still presented a clear tendency to

Table 5. Coefficients and adjusted statistics of the models used in the adjustment strategies III (Models 13 and 14) and IV (Models 15 and 16).

Nº	β_0	β_1	β_2	β_3	β_4	S_{yx} %	$R^2_{adj.}$	AIC	BIC
13*	-1.10*	1.07*	1.27*	0.94*		71.17	0.75	55393.48	55403.94
14*	-0.18	2.19	-0.24	0.04	1.04*	76.62	0.71	57096.77	57109.74
15*	-27.66*	14.00*	2.64*	3.44*	3.32*	58.72	0.83	50540.10	50553.08
16*	**-0.97**	**-0.13***	**0.12***	**0.28***	**0.55**	**40.91**	**0.92**	**60753.60**	**60766.58**

*Significant at 95% probability level.
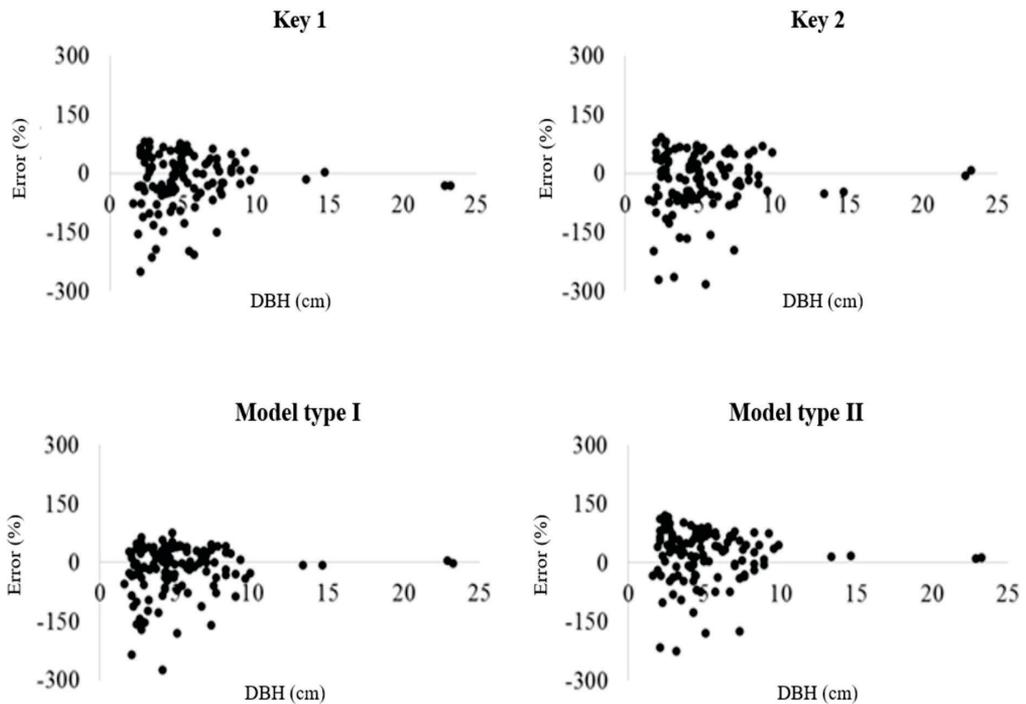
Figure 3. Residual graphic distribution of models adjusted according to the strategies III (Keys 1 and 2) and IV (type 1 and 2).

Araújo EJG, Loureiro GH, Sanquetta CR, Sanquetta MNI, Corte APD,
10/13 Péllico Netto S et al.

Floresta e Ambiente 2018; 25(1): e20160193

overestimate dry biomass results, mainly in the case of trees with smaller diameters (Figure 3).

The Shapiro-Wilk's test and Durbin-Watson statistics were not significant (p-value > 0.05), indicating that the residuals were normally distributed and were not correlated among themselves, while the White test was significant (p-value < 0.05), showing the presence of heteroscedasticity of residuals. To correct this problem, two alternatives were used: the first consisted in the transformation of the dependent variable, and the second in the use of weighted regression. The transformation of the dependent variable was identified from a family of data transformations $Y' = Y^\lambda$ suggested by Box-Cox. In the present case, λ ranged from -1 to 1, with an interval of 0.1. Homogeneity of residuals was not

achieved even after Box-Cox transformation. Because of that, we decided to use weighted regression, applying the weight $w_i = \dfrac{1}{DBH.Ht_i^2}$. The resulting equation was:

$y = -0.97316 - 0.13167\,DBH + 0.121527\,DBH^2 + 0.284539\,Hc^2 + 0.555392\,Dc^2$, with

R^2_{adj} = 10.60 and $S_{yx}\%$ = 88.19%.

The best adjusted statistics were obtained when the Type 2 model was applied to adjust the set of equations with stratified data. The adjusted coefficients of determination (R^2_{adj}.) were higher than those obtained with other adjustment procedures, reaching values greater than 0.90. Moreover, there was a clear reduction in the standard error of the estimate ($S_{yx}\%$), reaching errors close to 16% when the number of diameter classes was increased (Figure 4). Also, both statistical
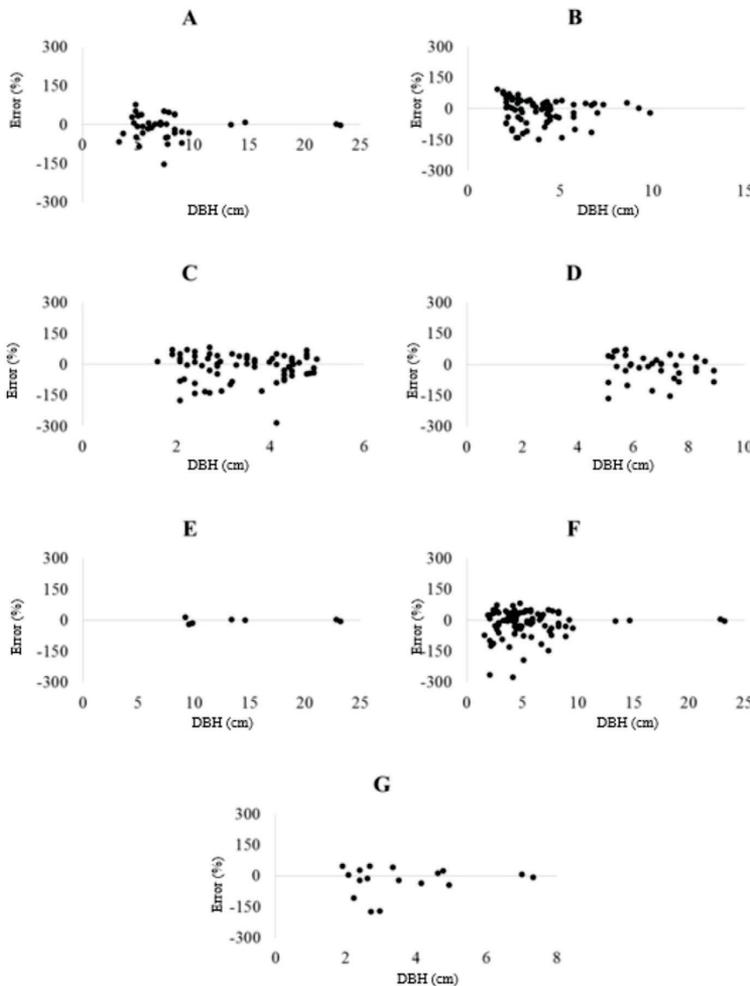
Figure 4. Residual graphic distribution of the Type 2 model adjusted to the stratified data of Ht/DBH ratio (A and B), diameter classes (C, D and E) and ecological categories (F and G).

Table 6. Coefficients and adjusted statistics of the Type II model applied to different stratification methods.

Stratification	Group	Range	β_0	β_1	β_2	β_3	β_4	S_{yx} %	$R^2_{adj.}$	AIC	BIC
Ratio Ht/DBH	A*	< 1.0	3.21	0.71*	0.15*	0.44*	-1.55	31.88	0.94	5358.33	5364.83
	B*	> 1.0	-1.16	0.64*	0.20	0.17	-0.29	49.51	0.86	16560.79	16571.26
Diameter Classes	C*	< 5 cm	0.57	0.64*	0.30	0.31*	-1.43	63.60	0.70	12316.08	12326.12
	D*	5 to 9 cm	-18.67	0.81*	-0.24	0.35*	4.38	43.14	0.66	5248.87	5254.99
	E*	> 9 cm	-1.05	0.62	0.19	0.28	-1.43	16.71	0.93	240.79	180.52
Ecological Groups	F*	Pioneer	0.97	0.69*	0.16*	0.36*	-1.15*	39.50	0.92	23240.17	23251.42
	G	Secondary	-7.90	0.16	-0.37	0.09	4.86	44.88	0.57	423.35	422.06

*Significant at 95% probability level.

values of AIC and BIC presented a significant reduction. However, these results occurred only in the case of trees of larger size, while the error remained high in trees of smaller diameter (Table 6). Data stratification using the ratio Ht/DBH indicates the best strategy. In this situation, the errors were smaller than 50% and the graphical distribution of residuals homogeneous when compared to other strategies. Additionally, the adjusted coefficients of determination were acceptable for biomass estimation with the obtained equations.

4. CONCLUSION

Schumacher-Hall model should be used to estimate the total dry biomass in forest restoration areas when only the variables diameter at 1.30 m above ground and total height are used as independent variables.

The inclusion of the variable wood basic density does not provide significant gains in accuracy of the equation, making its use impractical in mathematical models, besides the difficulty to be obtained. The average canopy diameter, commercial height and diameter at 1.30 m above ground, which are easily measured in the field, are the variables that best explain the variation in total dry biomass. They should be used in the Type 2 model, as independent variables, because they contribute to a more accurate equation when estimating the dry weight of native species in forest restoration areas.

Finally, to promote gains in accuracy of the estimation of the total dry biomass in the Type 2 model, the model should be adjusted to data stratified into classes for the height-diameter ratio (Ht/DBH).

SUBMISSION STATUS

CORRESPONDENCE TO

Emanuel José Gomes de Araújo

Departamento de Silvicultura, Universidade Federal Rural do Rio de Janeiro – UFRRJ, BR 465, km 7, CEP 23890-000, Seropédica, RJ, Brazil
e-mail: ejgaraujo@gmail.com

REFERENCES

Addo-Fordjour P, Rahmad ZB. Development of allometric equations for estimating above-ground liana biomass in tropical primary and secondary forests, Malaysia. *International Journal of Ecology* 2013; 2013: 1-9.

Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In: Parzen E, Tanabe K, Kitagawa G, editors. *Selected Papers of Hirotugu Akaike*. New York: Springer; 1998. (Perspectives in Statistics).

Barbosa ERM, Langevelde F, Tomlinson KW, Carvalheiro LG, Kirkman K, Bie S et al. Tree species from different functional groups respond differently to environmental changes during establishment. *Oecologia* 2014; 174(4): 1345-1357. PMid:24337711. http://dx.doi.org/10.1007/s00442-013-2853-y.

Baskerville GL. Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research* 1972; 2(1): 49-53. http://dx.doi.org/10.1139/x72-009.

Brasil. Ministério da Agricultura e Reforma Agrária. *Normais climatológicas 1961-1990*. Brasília: Mara; 1992.

Chave J, Andalo C, Brown S, Cairns MA, Chambers JQ, Eamus D et al. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* 2005; 145(1): 87-99. PMid:15971085. http://dx.doi.org/10.1007/s00442-005-0100-x.

Chen JM, Liu J, Leblanc SG, Lacaze R, Roujean J-L. Multi-angular optical remote sensing for assessing vegetation structure and carbon absorption. *Remote Sensing of Environment* 2003; 84(4): 516-525. http://dx.doi.org/10.1016/S0034-4257(02)00150-5.

Araújo EJG, Loureiro GH, Sanquetta CR, Sanquetta MNI, Corte APD,
Péllico Netto S et al.

12/13

Corte APD, Sanquetta CR. Quantificação do estoque de carbono fixado em reflorestamento de Pinus na área de domínio de da floresta Ombrófila Mista no Paraná. *Cerne* 2007; 13(1): 32-39.

Cunha US, Machado AS, Figueiredo Filho A. Uso de análise exploratória de dados e de regressão robusta na avaliação do crescimento de espécies comerciais de terra firme da Amazônia. *Árvore* 2002; 26(4): 391-402.

Drolet GG, Huemmrich KF, Hall FG, Middleton EM, Black TA, Barr AG et al. A Modis-derived photochemical reflectance index to detect inter-annual variations in the photosynthetic light-use efficiency of a boreal deciduous forest. *Remote Sensing of Environment* 2005; 98(2-3): 212-224. http://dx.doi.org/10.1016/j.rse.2005.07.006.

Frankenberg C, Fisher JB, Worden J. New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters* 2011; 38(17): L17706.

Gehring C, Park S, Denich M. Liana allometric biomass equations for Amazonian primary and secondary forest. *Forest Ecology and Management* 2004; 195(1-2): 69-83. http://dx.doi.org/10.1016/j.foreco.2004.02.054.

Goetz SJ, Prince SD, Goward SN, Thawley MM, Small J. Satellite remote sensing of primary production: An improved production efficiency modeling approach. *Ecological Modelling* 1999; 122(3): 239-255. http://dx.doi.org/10.1016/S0304-3800(99)00140-4.

Higuchi N, Santos J, Ribeiro RJ, Minette L, Biot Y. Biomassa da parte aérea da vegetação da floresta tropical úmida de terra-firme da Amazônia brasileira. *Acta Amazonica* 1998; 28(2): 153-166. http://dx.doi.org/10.1590/1809-43921998282166.

Kenzo T, Ichie T, Hattori D, Itioka T, Handa C, Ohkubo T et al. Development of allometric relationships for accurate estimation of above- and below-ground biomass in tropical secondary forests in Sarawak, Malaysia. *Journal of Tropical Ecology* 2009; 25(04): 371-386. http://dx.doi.org/10.1017/S0266467409006129.

Ketterings QM, Coe R, Noordwijk M, Ambagau Y, Palm CA. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. *Forest Ecology and Management* 2001; 146(1-3): 199-209. http://dx.doi.org/10.1016/S0378-1127(00)00460-6.

Kramer PJ, Koslowski T. *Fisiologia das árvores*. Lisboa: Fundação Calouste Gulbenkian; 1972.

Larcher W. *Ecofisiologia vegetal*. São Paulo: EPU; 1986.

Litton CM, Kauffman JB. Allometric models for predicting aboveground biomass in two widespread woody plants in Hawaii. *Biotropica* 2008; 40(3): 313-320. http://dx.doi.org/10.1111/j.1744-7429.2007.00383.x.

Luo Y, Zhang X, Wang X, Ren Y. Dissecting variation in biomass conversion factors across China's Forests: implications for biomass and carbon accounting. *PLoS One* 2014; 9(4): e94777. PMid:24728222. http://dx.doi.org/10.1371/journal.pone.0094777.

Melo LC, Sanquetta CR, Corte APD, Hentz AK. Estimativa de biomassa e carbono total para árvores de caixeta no Paraná. *Pesquisa Florestal Brasileira* 2014; 77(77): 21-29. http://dx.doi.org/10.4336/2014.pfb.34.77.592.

Moore JR. Allometric equations to predict the total above-ground biomass of radiata pine trees. *Annals of Forest Science* 2010; 67(8): 1-11. http://dx.doi.org/10.1051/forest/2010042.

Morais VA, Scolforo JRS, Silva CA, Mello JM, Gomide LR, Oliveira AD. Carbon and biomass stocks in a fragment of cerradão in Minas Gerais state, Brazil. *Cerne* 2013a; 19(2): 237-245. http://dx.doi.org/10.1590/S0104-77602013000200007.

Morais VA, Silva CA, Scolforo JRS, Mello JM, Araújo EJG, Assis EA. Modelagem do teor de carbono orgânico em solos de fragmentos de cerrado de Januária e Bonito de Minas, Minas Gerais. *Pesquisa Florestal Brasileira* 2013b; 76(76): 343-354. http://dx.doi.org/10.4336/2013.pfb.33.76.507.

Nogueira EM, Fearnside PM, Nelson BW. Normalization of wood density in biomass estimates of Amazon forests. *Forest Ecology and Management* 2008; 256(5): 990-996. http://dx.doi.org/10.1016/j.foreco.2008.06.001.

Péllico Netto S, Brena DA. *Inventário florestal*. Curitiba: UFPR, UFSM; 1997.

R Development Core Team. *R: a language and environment for statistical computing* [online]. Vienna: R Foundation for Statistical Computing; 2001. [cited 2017 Nov 10]. Available from: www.R-project.org

Regazzi AJ. Teste para verificar a identidade de modelos de regressão e a igualdade de alguns parâmetros num modelo polinomial ortogonal. *Revista Ceres* 1992; 228: 176-195.

Rezende AV, Vale AT, Sanquetta CR, Figueiredo A Fo, Felfili JM. Comparação de modelos matemáticos para estimativa do volume, biomassa e estoque de carbono da vegetação lenhosa de um cerrado sensu stricto em Brasília, DF. *Scientia Forestalis* 2006; 71: 65-76.

Rochadelli R. *Estrutura de fixação dos átomos de carbono em reflorestamento (Estudo de caso: Mimosa scabrella Bentham, Bracatinga)* [Tese]. Curitiba: Setor de Ciências Agrárias, Universidade Federal do Paraná; 2001.

Sanquetta CR, Balbinot R, Ziliotto MAB. *Fixação de carbono: atualidades, projetos e pesquisas*. Curitiba: AM Impressos; 2004.

Schwarz G. Estimating the dimensional of a model. *Annals of Statistics* 1978; 6(2): 461-464. http://dx.doi.org/10.1214/aos/1176344136.

Scolforo JR, Rufini AL, Mello JM, Trugilho PF, Oliveira AD, Silva CPC. Equações para peso de matéria seca das fisionomias, em Minas Gerais. In: Scolforo JRS, Oliveira

AD, Acerbi FW Jr. *Inventário florestal de Minas Gerais: equações de volume, peso de matéria seca e carbono para diferentes fitofisionomias da flora nativa*. Lavras: UFLA; 2008.

Scolforo JRS. *Biometria florestal: parte I - modelos de regressão linear e não linear; parte II - modelos para relação hipsométrica, volume, afilamento e peso de matéria seca*. Lavras: UFLA, FAEPE; 2005. 352 p.

Segura M, Kanninen M. Allometric models for tree volume and total aboveground biomass in a tropical humid forest in Costa Rica. *Biotropica* 2005; 37(1): 2-8. http://dx.doi.org/10.1111/j.1744-7429.2005.02027.x.

Soares BS Fo, Nepstad DC, Curran LM, Cerqueira GC, Garcia RA, Ramos CA et al. Modeling conservation in the Amazon basin. *Nature* 2006; 440(7083): 520-523. PMid:16554817. http://dx.doi.org/10.1038/nature04389.

Song C, Dannenberg MP, Hwang T Optical remote sensing of terrestrial ecosystem primary productivity. *Progress in Physical Geography* 2013; 37(6): 834-854. http://dx.doi.org/10.1177/0309133313507944.

Urbano E, Machado AS, Figueiredo-Filho A, Koehler HS. Equações para estimar o peso de carbono fixado em árvores de *Mimosa scabrella* Bentham (bracatinga) em povoamentos nativos. *Cerne* 2008; 14(3): 194-203.

Vogel HLM, Schumacher MV, Trüby P. Quantificação da biomassa em uma Floresta Estacional Decidual em Itaara, RS, Brasil. *Ciência Florestal* 2006; 16(4): 419-425. http://dx.doi.org/10.5902/198050981923.

Wang C. Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests. *Forest Ecology and Management* 2006; 222(1-3): 9-16. http://dx.doi.org/10.1016/j.foreco.2005.10.074.

Watzlawick LF, Kirchner FF, Sanquetta CR. Estimativa de biomassa e carbono em floresta com Araucaria utilizando Imagens do satélite ikonos II. *Ciência Florestal* 2009; 19(2): 169-181. http://dx.doi.org/10.5902/19805098408.