Mixed models for biomass prediction in the semiarid zone of Pernambuco State, Brazil

MODELOS MISTOS PARA PREVISÃO DE BIOMASSA NO SEMIÁRIDO DO ESTADO DE PERNAMBUCO, BRASIL

Jadson Coelho de Abreu1, Jose Antônio Aleixo da Silva2, Rinaldo Luiz Caraciolo Ferreira2, Carlos Pedro Boechat Soares3, Carlos Moreira Miquelino Eleto Torres3, Aline Araújo Farias3, Ivaldo da Silva Tavares Júnior3

1Universidade do Estado do Amapá – UEAP, Macapá, AP, Brasil
2Universidade Federal Rural de Pernambuco – UFRPE, Recife, PE, Brasil
3Universidade Federal de Viçosa – UFV, Viçosa, MG, Brasil

How to cite: Abreu, J. C., Silva, J. A. A., Ferreira, R. L. C., Soares, C. P. B., Torres, C. M. M. E., Farias, A. A., & Tavares Júnior, I. S. (2020). Mixed models for biomass prediction in the semiarid zone of Pernambuco State, Brazil. *Scientia Forestalis*, 48(128), e3246. https://doi.org/10.18671/scifor.v48n128.10

Abstract

Most of the datasets used for biomass modeling in native forests have high variability, containing several sample plots of different sites, trees of different species in the same plot, and several measurements of the same tree along the bole. Assuming that data from uneven-aged forests are highly diversified, the mixed-effects modeling analyzes hierarchically structured data more efficiently than any other approach and increases the prediction accuracy of the equations. The objective of this paper was to develop a mixed model to predict the above-ground biomass for Caatinga species in the municipality of Floresta – PE, Brazil. Biomass data of 100 trees belonging to five species were used for this purpose. The model of Schumacher and Hall was fitted considering the structure of a mixed linear model by the inclusion of intercepts and random coefficients, considering the species as the random effect. The best model was selected based on the root mean square error, bias, mean absolute error, Akaike information criterion, and residual graphic analysis. Species as a random effect contributed to increasing the accuracy of the estimates in a mixed model. The Schumacher and Hall model with random effects on dbh (diameter at breast height) and h (height) was the best way to predict the total biomass of the trees in the Pernambuco semiarid region.

Keywords: Hierarchical data; Indirect method; Random effects; Schumacher and Hall.

Resumo

A maioria dos conjuntos de dados utilizados para modelagem de biomassa em florestas nativas possui alta variabilidade e provêm de parcelas amostrais de diferentes sítios, árvores de diferentes espécies na mesma parcela e várias medidas da mesma árvore ao longo do fuste. Considerando que dados provenientes de florestas inequiâneas são altamente diversificados, a abordagem de modelagem de efeitos mistos permite analisar dados hierarquicamente estruturados de forma mais eficiente do que qualquer outra abordagem e aumenta a precisão da predição das equações. O objetivo deste trabalho foi ajustar um modelo misto para predizer a biomassa aérea para espécies da Caatinga no município de Floresta – PE. Dados de biomassa de 100 árvores, de cinco espécies, foram utilizadas para esta finalidade. O modelo de Schumacher e Hall (1933) foi ajustado considerando a estrutura de um modelo linear misto, pela inclusão de interceptos e coeficientes de inclinação aleatórios, considerando as espécies como efeito aleatório. O melhor modelo foi selecionado com base na raiz do erro médio quadrático, bias, distância média absoluta, critério de informação Akaike e distribuição gráfica dos resíduos. A inclusão das espécies como efeito aleatório contribuiu para aumentar a precisão das estimativas no modelo misto. O modelo...
INTRODUCTION

The quantification of above and below-ground biomass, at both individual tree and stand levels, is essential for estimating carbon stocks for forest management and climate-related issues (He et al., 2018). In Brazil, several studies have been developed using traditional regression models for predicting biomass (Barbeiro et al., 2009; Abreu et al., 2016; Oliveira et al., 2018).

However, most of the datasets for modeling the volume and biomass in native forests come from several sampling plots of different sites, trees of different species in the same plot, and several measurements of the same tree along the bole (de-Miguel et al., 2013). These factors result in traditional regression models showing a higher amplitude of error due to the diversity found in uneven-aged forests.

The mixed models are an alternative to modeling heterogeneous environments. These statistical models contain both fixed and random effects and are frequently used to analyze data in settings with pooled data, such as longitudinal data, repeated measurements, and multilevel data (Resende et al., 2014; Özkale & Kuran, 2018).

The modeling using mixed models should include the estimated fixed and random parameters that explain the observed values because these models are available for non-normal data analysis collected in natural conditions involving random effects (Sakai et al., 2013). When fitting a mixed model, it is possible to predict the biomass for individual trees, which allows better verifying the variability of the model with different structures, such as random effect in all the parameters or only in some of the parameters involved.

There are several studies involving mixed models in forest measurement, for instance, prediction of dominant height growth of *Eucalyptus* sp. (Calegario et al., 2005a), height of *Eucalyptus* sp. (Mendonça et al., 2015); diameter annual increment in the Atlantic Forest (Soares et al., 2015), and volume of *Eucalyptus* sp. (Gouveia et al., 2015). Although well applied for the above purposes, studies with mixed models for predicting biomass are scarce in Brazil. In other countries, researches in this regard have been conducted, including the ones from Zeng (2015); Ou et al. (2016); Chen et al. (2017) and Poudel et al. (2018). Thus, the objective of this paper was to fit a mixed model to predict the above-ground biomass for Caatinga species in the municipality of Floresta – PE, Brazil.

MATERIAL AND METHODS

Study area

The study was carried out in an area of 50 ha (8°30’37”S and 37°59’07”W, altitude 323 m) with Caatinga vegetation, in the municipality of Floresta, in the São Francisco middle region of Pernambuco State. The climate of the area is semi-arid hot and corresponds to the Köppen class BSh. The mean annual precipitation is 503 mm, with a rainy period from January to April, and the mean annual temperature is 26.1 ºC (Araujo Filho et al., 2001).

Database

The species were selected based on the information obtained from a forest inventory carried out in the area, developed by Alves et al. (2017). The five species of the highest value of importance were selected for the study. These species were *Poincianella bracteosa* (Tul.) L. P. Queiroz, *Mimosa ophtalmoncentra* Mart. ex Benth, *Aspidosperma pyrifolium* Mart, *Cnidoscolus quercifolius* (Mull. Arg.) Pax. & Hoffm, *Anadenanthera colubrina* var. cebil (Griseb.) Altschul. Twenty individuals per species were selected for biomass quantification.
Quantification of biomass

The individuals belonging to the five above mentioned species were distributed in circumference classes and cubed. A total of 100 individuals were harvested for analyzing the above-ground biomass; four trees representative of each mean circumference class for each species. The individuals which were partially cut; burned or fallen were not considered in order to cover the predicted diameter class.

The circumference at breast height (cbh) of the selected trees was measured at 1.30 m and converted to the diameter at breast height (dbh). Then, the total height (h) and height of the commercial bole (ch) of the trees were measured. After measuring the dendrometric variables, the trees were cut down, and the bole, branches, and leaves were rigorously separated. In order to obtain the biomass, the tree compartments (leaves, branches, and bole) were weighed separately using a precision scale.

Model fitting

The model of Schumacher & Hall (1933) was initially fitted in its linearized form with the data of green biomass (B), diameter at breast height (dbh), and height (h), using:

\[ \ln B = \beta_0 + \beta_1 \ln \text{dbh} + \beta_2 \ln h + \varepsilon_i \]  

Where: \( \ln = \) natural logarithm; \( B = \) green biomass (in kg); \( \text{dbh} = \) diameter at breast height (in cm); \( h = \) total height (in m); \( \beta_0, \beta_1, \) and \( \beta_2 = \) model parameters; and \( \varepsilon = \) random error.

The logarithmic transformation of the data allowed the assumption of homoscedasticity, and the covariates were taken to be zero due to the non-use of longitudinal data (Gujarati & Porter, 2011).

The model was fitted by the maximum likelihood method using the package \texttt{glm2} from \texttt{R} (R Core Team, 2014). The goodness of the fitted model was assessed by fitted coefficient of determination (\( r^2 \)), root mean square error (RMSE%), bias, mean absolute error (MAE), and graphical analysis of the residues (Silva et al., 2009; Binoti et al., 2015).

Subsequently, the model of Schumacher and Hall was fitted considering the structure of a linear mixed model, by including random intercept and coefficients, using the species as random effect variables:

\[ \ln B_i = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i}) \times \ln \text{dbh}_i + (\beta_2 + b_{2i}) \times \ln h_i + \varepsilon_i \]  

Where: \( \beta_0, \beta_1, \) and \( \beta_2 = \) fixed parameters of the model; \( b_{0i} = \) random intercept for the \( i \)th species; \( b_{1i} \) to \( b_{2i} = \) random coefficients for the \( i \)th species.
The linear mixed model was in the following form (Wu, 2009):

\[ y_i = X_i \beta + Z_i b_i + e_i, \quad i = 1, 2, \ldots, n. \]  

(9)

\[ b_i \sim N(0, G), \quad e_i \sim N(0, R_i) \]  

(10)

Where \( \beta \) are fixed effects, \( b_i \) are random effects, \( X_i \) is a design matrix containing covariates of individual \( i \), \( Z_i \) is a design matrix, \( e_i \) are random errors, \( R_i \) is a \( n \times n \) variance-covariance matrix within individual measurements \((R_i = I \times \sigma^2_i)\), where \( I \) is an identity matrix), and \( G \) is the variance-covariance matrix of random effects.

The fittings of the models related to the mixed form were performed by the Restricted Maximum Likelihood Method, using the \textit{nlme} package from \textit{R}. The Akaike information criterion (AIC) was used instead of the coefficient of determination to select the best functional form for the mixed model.

The result of the inclusion of the random effect on the intercept and coefficients was verified by the maximum likelihood ratio test (Resende et al., 2014). Thus, the significance of the difference (D) between the deviations \([-2\log(L)]\) for the models was verified by the \(\chi^2\) test at the 5% significance level. After selecting the best equation, the White and Shapiro-Wilk tests were applied to check the assumptions of homoscedasticity and normality of the residuals.

At the end of the modeling process, the model selected as the best mixed linear model could be the complete model (or random effect model, with random effect on all parameters), a partial model (i.e., with random effect associated only with some parameters of the model), or the fixed model (due to the non-significance of the random effect by the maximum likelihood ratio test).

RESULTS AND DISCUSSION

The biomass and height data showed higher variability compared with the diameter distribution for the species, despite the presence of some outliers for this variable (Figure 1).
All parameter estimates of the Schumacher and Hall equations fitted with fixed and mixed effects were significant for the fixed effects (Table 1). The random coefficients of each equation are presented in Table 2. The variability of the data was captured in all the structures.

### Table 1. Fixed parameters of the Schumacher and Hall equation

| Coefficients | Fixed Schumacher | Random Effect on $\beta_0$ | Random Effect on $\beta_0$, $\beta_1$ | Random Effect on $\beta_0$, $\beta_1$ e $\beta_2$ | Random Effect on $\beta_1$ | Random Effect on $\beta_1$, $\beta_2$ | Random Effect on $\beta_1$, e $\beta_2$ |
|--------------|------------------|---------------------------|----------------------------------------|---------------------------------|-------------------|---------------------------------|---------------------------------|
|              | $\beta_j$        | Std. Error                | p-value                                | Std. Error                      | Std. Error        | p-value                          | Std. Error                      | p-value                          |
| $\hat{\beta}_0$ | 0.6870169        | 0.11791279                | <0.0001                                | 0.8799123                       | 0.1429455        | <0.0001                          | 0.8917796                       | 0.11440684                      | <0.0001                          |
| $\hat{\beta}_1$ | 0.5854446        | 0.08278914                | <0.0001                                | 0.8348354                       | 0.0803142        | <0.0001                          | 0.8214599                       | 0.0944473                       | <0.0001                          |
| $\hat{\beta}_2$ | 0.6510529        | 0.09728160                | <0.0001                                | 0.2660971                       | 0.1084436        | <0.0160                          | 0.2740579                       | 0.10435063                      | <0.0101                          |
| $\hat{\beta}_0$ | 0.2660971        | 0.1084436                 | <0.0160                                | 0.8799123                       | 0.1429455        | <0.0001                          | 0.8917796                       | 0.11440684                      | <0.0001                          |
| $\hat{\beta}_1$ | 0.8348354        | 0.0803142                 | <0.0001                                | 0.2660971                       | 0.1084436        | <0.0160                          | 0.8214599                       | 0.0944473                       | <0.0001                          |
| $\hat{\beta}_2$ | 0.8799123        | 0.1429455                 | <0.0001                                | 0.8348354                       | 0.0803142        | <0.0001                          | 0.8214599                       | 0.0944473                       | <0.0001                          |
| $\hat{\beta}_0$ | 0.8348354        | 0.0803142                 | <0.0001                                | 0.2660971                       | 0.1084436        | <0.0160                          | 0.8214599                       | 0.0944473                       | <0.0001                          |
| $\hat{\beta}_1$ | 0.8799123        | 0.1429455                 | <0.0001                                | 0.2660971                       | 0.1084436        | <0.0160                          | 0.8214599                       | 0.0944473                       | <0.0001                          |
| $\hat{\beta}_2$ | 0.8348354        | 0.0803142                 | <0.0001                                | 0.2660971                       | 0.1084436        | <0.0160                          | 0.8214599                       | 0.0944473                       | <0.0001                          |
Table 2. Random effects of Schumacher and Hall equation

| Structure of the model | \(a_0\)       | \(b_1\)       | \(b_2\)       |
|------------------------|---------------|---------------|---------------|
| Random Effect on \(\beta_0\) | 0.299837762  | -0.002976402  | 0.0584743468  |
|                        | -0.290564712  | -0.0012394655 | -0.01201064177|
|                        | 0.06098006    | 0.137151110   | 0.07376847    |
| Random Effect on \(\beta_0\) and \(\beta_1\) | 0.0005878151  | -0.16748280   | 0.0305180637  |
|                        | 0.0305261921  | -0.02744014   | 0.006098006   |
|                        | 0.137151110   | -0.0012394655 | 0.07376847    |
|                        | 0.0005878151  | 0.0305180637  | 0.006098006   |
|                        | 0.0305261921  | -0.02744014   | 0.07376847    |
| Random Effect on \(\beta_1\) | 0.0584743468  | 0.137151110   | -0.1201064177 |
| Random Effect on \(\beta_2\) | 2.148548e-10  | 0.182245873   | 0.012394655   |
|                        | -1.235430e-09 | 0.013999095   | -0.0012394655 |
|                        | -2.165385e-09 | -0.203528323  | 0.0905878151  |
|                        | 2.217069e-09  | 0.004520327   | 0.0305261921  |
|                        | 9.688910e-10  | 0.11456207    | 0.0305261921  |
|                        | -2.226138e-10 | 0.14944391    | 0.11456207    |
|                        | -2.315203e-09 | 0.011803683   | 0.14944391    |
|                        | 7.655274e-11  | -0.129586739  | 0.14944391    |
|                        | 1.712570e-09  | -0.05223089   | 0.14944391    |
|                        | 7.486943e-10  | 0.03679983    | 0.14944391    |
| Random Effect on \(\beta_0\), \(\beta_1\) and \(\beta_2\) | -0.1677194232 | -0.0124538819 | 0.0142910276 |
|                        | -0.1677194232 | 0.0142910276  | 0.0142910276  |
|                        | 0.031074509   | 0.14944391    | 0.14944391    |
|                        | 0.119696187   | 0.14944391    | 0.14944391    |
| Random Effect on \(\beta_1\) | 0.1663645281  | 0.14944391    | 0.14944391    |
| Random Effect on \(\beta_2\) | 0.011803683   | 0.14944391    | 0.14944391    |
|                        | 0.006536734   | 0.14944391    | 0.14944391    |
|                        | 0.03679983    | 0.14944391    | 0.14944391    |
| Random Effect on \(\beta_1\) and \(\beta_2\) | 0.031074509   | 0.14944391    | 0.14944391    |
|                        | 0.119696187   | 0.14944391    | 0.14944391    |
| Random Effect on \(\beta_1\) | -0.129586739  | -0.05223089   | -0.129586739  |
| Random Effect on \(\beta_2\) | 0.031074509   | 0.14944391    | 0.14944391    |
|                        | 0.119696187   | -0.11456207   | 0.14944391    |
| Random Effect on \(\beta_1\) and \(\beta_2\) | 0.031074509   | 0.14944391    | 0.14944391    |
| Random Effect on \(\beta_1\) | 0.006536734   | -0.01945078   | 0.006536734   |
| Random Effect on \(\beta_2\) | 0.03679983    | -0.01945078   | 0.03679983    |

The Schumacher and Hall equation, containing fixed effects, presented a good fit: the adjusted coefficient of determination was 73.72, correlation (84.45%), mean absolute error (4.59), bias (0.66), root mean square error (6.84) and graphical distribution of residues tending to homogeneity (Figure 2). However, when random effects were added, there was an improvement in the accuracy of the equation (Table 3). All equations with random effects were superior compared to the equation in the fixed form. The equation with the lowest AIC uses the random effect only in the coefficient with the height variable, and the best statistics were observed for the equations with random effects in dbh and h.

Table 3. Evaluation statistics of the Schumacher and Hall model in the mixed form

| Structure of the model | AIC   | MAE   | BIAS  | RMSE (%) |
|------------------------|-------|-------|-------|----------|
| Fixed Effect           | 40.32 | 4.59  | 0.66  | 6.84     |
| Random Effect on \(\beta_0\) | 12.62 | 3.27  | 0.43  | 5.03     |
| Random Effect on \(\beta_0\) and \(\beta_1\) | 6.31  | 3.10  | 0.30  | 4.73     |
Mixed models for biomass prediction in the semiarid zone of Pernambuco State, Brazil

Table 3. Continued...

| Structure of the model                  | AIC  | MAE  | BIAS | RMSE (%) |
|----------------------------------------|------|------|------|----------|
| Random Effect on $\beta_0$ and $\beta_2$ | 6.47 | 3.16 | 0.29 | 4.95     |
| Random Effect on $\beta_0$, $\beta_1$ and $\beta_2$ | 6.21 | 3.06 | 0.28 | 4.74     |
| Random Effect on $\beta_1$              | 4.74 | 3.13 | 0.30 | 4.86     |
| Random Effect on $\beta_2$              | 4.47 | 3.16 | 0.29 | 4.95     |
| Random Effect on $\beta_1$ and $\beta_2$ | 4.21 | 3.06 | 0.28 | 4.74     |

Where: AIC = Akaike information criterion, MAE = mean absolute error, BIAS = computes the average amount by which actual is greater than predicted and RMSE = root mean square error.

Figure 2. Graphical distribution of the residuals (%) as a function of biomass (B) for the equations in the fixed and mixed form.
We observed that the residues were well distributed, with mean around zero and constant variance, the assumption of homogeneity for the equation with random effects on dbh and h were accepted (White's test = 62.5182, \( p = 0.7798 \)). Only three observations were discrepant, i.e., outside the confidence limits of -50 to 50% (Figure 2).

In addition to rejecting the hypothesis of variance heterogeneity, the residuals follow the normal distribution confirmed by the Shapiro Wilk test (\( W = 0.97937, p = 0.1187 \)).

The species as a random effect was significant (\( p > 0.05 \)) by the maximum likelihood ratio test in all equations, which means that the random effects improve the accuracy of a model with only fixed parameters. Thus, the final equation resulted in an equation with fixed and random effects (Table 4).

### Table 4. Maximum Likelihood Ratio Test

| Model | Restricted Maximum Likelihood | Test | Maximum Likelihood Ratio | p-value |
|-------|-------------------------------|------|--------------------------|---------|
| M2    | -1.31145                      | F vs M2 | 30.19799                | <0.0001 |
| M3    | 2.84592                       | F vs M3 | 38.51273                | <0.0001 |
| M4    | 2.76140                       | F vs M4 | 38.34368                | <0.0001 |
| M5    | 3.89454                       | F vs M5 | 40.60997                | <0.0001 |
| M6    | 2.62748                       | F vs M6 | 38.07584                | <0.0001 |
| M7    | 2.76140                       | F vs M7 | 38.34368                | <0.0001 |
| M8    | 3.89454                       | F vs M8 | 40.60997                | <0.0001 |
| Fixed | -16.41044                     |       |                         |         |

Where: F = fixed model, M2 to M8 = Model 2 to model 8.

The random-effects by species of the random equation in dbh and h (M8) are shown in Figure 3. The prediction of the best model is on the left, and we can observe the presence of considerable variability. The graph of the predicted values versus the observed values is displayed on the right side and followed the 45\(^\circ\) line.

![Figure 3](image)

**Figure 3.** Biomass prediction of the best random effect model vs. dbh and graph of predicted (Predicted B) and observed (Observed B) biomass for the five studied species: (1) *Poincianella bracteosa*, (2) *Mimosa ophthalmocesta*, (3) *Aspidosperma pyrifolium*, (4) *Cnidoscolus quercifolius*, (5) *Anadenanthera colubrina*.

The significance of the coefficients of the equations in both fixed and mixed forms indicates that the use of dbh and h to obtain the biomass was adequate. Since the \( p \)-values were significant for all parameters, they were essential for the model (Calegario et al., 2005b).
The mixed-effects models were superior to the conventional model in the fixed form. Comparing the equation in the fixed form with the best equation in the mixed form, the correlation coefficient increased (from 85.68% to 92.86%) and the root mean square error decreased (from 6.84 to 4.79), evidencing the improvement of the models with the insertion of the random effects, which corroborates the results obtained by Adame et al. (2008) and Crecente-Campo et al. (2010).

All the equations presented proper fitting, especially those with random effect in all the parameters and the equation with random effect in dbh and h. These results corroborate those observed by Gouveia et al. (2015), who used the Schumacher and Hall model to predict Eucalyptus volume on the Araripe plateau and obtained a better fit with the random-effects model on all parameters.

The model with random effect in dbh and h can be considered more parsimonious than the model with random effect in all the parameters, due to the lower number of parameters, lower AIC, and higher maximum ratio likelihood ratio (Littell et al., 2006). Regarding the residual graphic analysis, the mixed model with random effects in the parameters of dbh and h presents an amplitude of error similar to the model with random effect in all parameters and smaller error amplitude compared to the other model structures.

The residues showed proper distribution, even with some data outside the confidence interval. However, few data outside a defined range can be considered negligible, since it represents a small amount when compared to the sample, and does not hugely interfere in the model prediction (Ebling & Péllico Netto, 2015). The predictive performance of the mixed model is substantially better than the fixed model when a sample is available to predict random effects (Temesgen et al., 2008). All equations with random effects show residuals with smaller amplitudes when compared to equations in the fixed form, as observed in the present study.

The improvement of the biomass prediction noticed in this study, considering the species as a random effect, reinforces the affirmation of Huff et al. (2018) that including species as random effect improves the mixed model compared to the fixed model for total biomass prediction. Other variables can be included as random effects such as forest formation, regional and local quality classes (Ou et al., 2016), precipitation, soil, elevation, and other geographic characteristics (Meng et al., 2007).

All variations of the Schumacher and Hall model resulted in significant equations by the maximum likelihood ratio test, showing that a mixed model is more efficient than adjusting an equation for each species, although biomass equations fitted by species predominate. One disadvantage of the latter approach is the lack of sufficient samples for each species (Huff et al., 2018). In a mixed model, if there is information for the variable response for a new species, then the random coefficients can be obtained and estimated explicitly for each species rather than obtained from the mean population response. In the mean population response, the vector of the random coefficients for a new individual has an expected value of zero (Burkhart & Tomé, 2012).

CONCLUSION

All the equations showed improvement when random variables were inserted in the Schumacher and Hall model, and the best functional form was with the inclusion of random effects in diameter at breast height (dbh) and height (h).

A single fixed-effect model for the database may present a good fit, but a mixed model further improves these results by capturing the effect of each observed species.

ACKNOWLEDGEMENTS

This work was carried out during the master’s degree of the first author at the Federal University of Pernambuco, with funding from the National Council for Scientific and Technological Development (CNPq).
REFERENCES
Abreu, J. C., Silva, J. A. A., Ferreira, R. L. C., & Alves Junior, F. T. (2016). Ajuste de modelos matemáticos lineares e não lineares para estimativa de biomassa e nutrientes de *Anadenanthera colubrina* var. *cebil* no semiárido pernambucano. *Ciência Florestal*, 44(111), 739-750. http://dx.doi.org/10.18671/scifor.v44n111.20.

Adame, P., Del Rio, M., & Cañellas, I. (2008). A mixed nonlinear height-diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). *Forest Ecology and Management*, 256(1-2), 88-98. http://dx.doi.org/10.1016/j.foreco.2008.04.006.

Alves, A. R., Ferreira, R. L. C., Da Silva, J. A. A., Dubex Júnior, J. C. B., Osajima, J. A., & De Holanda, A. C. (2017). Conteúdo de nutrientes na biomassa e eficiência nutricional em espécies da caatinga. *Ciência Florestal*, 27(2), 377-390. http://dx.doi.org/10.5902/1980509827686.

Araujo Filho, J. C., Silva, A. B., Silva, F. B. R., & Leite, A. P. (2001). Equações para estimativa da biomassa individual de *Ficus spp.* plantations in a plantation system. *Forest Ecology and Management*, 256(1-2), 88-98. http://dx.doi.org/10.1016/j.foreco.2008.04.006.

Barbeiro, L. S. S., Vieira, G., & Sanquetta, C. R. (2009). Equações para estimativa da biomassa individual de *Nectandra grandiflora* Ness (canela-amarela). *Floresta*, 39(4), 833-843. http://dx.doi.org/10.5380/rf.v39i4.16318.

Binoti, M. L. M. S., Leite, H. G., Binoti, D. H. B., & Gleriani, J. M. (2015). Prognose em nível de povoamento baseada na teoria dos modelos não lineares em multinível de efeito misto. *Cerne*, 21(1), 97-105. http://dx.doi.org/10.1590/01047760201521011153.

Burkhart, H. E., & Tomé, M. (2012). *Modeling forest trees and stands*. Dordrecht: Springer.

Calegario, N., Daniels, R. F., Maestri, R., & Neiva, R. (2005a). Modeling dominant height growth based on nonlinear mixed-effects model: a clonal *Eucalyptus* plantation case study. *Forest Ecology and Management*, 204(1), 11-20. http://dx.doi.org/10.1016/j.foreco.2004.07.051.

Calegario, N., Maestri, R., Leal, C. L., & Daniels, R. F. (2005b). Estimativa do crescimento de povoamentos de *Eucalyptus* baseada na teoria dos modelos não lineares em multinível de efeito misto. *Ciência Florestal*, 15(3), 285-292. http://dx.doi.org/10.1590/S0104-776020152101113.

Chen, D., Huang, X., Zhang, S., & Sun, X. (2017). Biomass modeling of larch (*Larix* spp.) plantations in China based on the mixed model, dummy variable model, and Bayesian hierarchical model. *Forests*, 8(8), 5-8. http://dx.doi.org/10.3390/f8080268.

Crecente-Campo, F., Tomé, M., Soares, P., & Diéguez-Aranda, U. (2010). A generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain. *Forest Ecology and Management*, 259(5), 943-952. http://dx.doi.org/10.1016/j.foreco.2009.11.036.

de-Miguel, S., Guzmán, G., & Pukkala, T. (2013). A comparison of fixed- and mixed-effects modeling in tree growth and yield prediction of an indigenous neotropical species (*Centrolobium tomentosum*) in a plantation system. *Forest Ecology and Management*, 291, 249-258. http://dx.doi.org/10.1016/j.foreco.2012.11.026.

Ebling, Â. A., & Péllico Netto, S. (2015). Modelagem de ocorrência de coortes na estrutura diamétrica da *Araucaria angustifolia* (Bertol.) Kuntze. *Cerne*, 21(2), 251-257. http://dx.doi.org/10.1590/01047760201521111667.

Gouveia, J. F., Da Silva, J. A. A., Ferreira, R. L. C., Gadelha, F. H. L., & Lima Filho, L. M. A. (2015). Modelos volumétricos mistos em clones de *Eucalyptus* no polo gesseiro do Araripe, Pernambuco. *Floresta*, 45(3), 587-598. http://dx.doi.org/10.5380/rf.v45i3.36844.

Gujarati, D. M., & Porter, D. C. (2011). *Econometria básica*. Porto Alegre: Editora Buckman.

He, H., Zhang, C., Zhao, X., Fousseni, F., Wang, J., Dai, H., Yang, S., & Zuo, Q. (2018). Allometric biomass equations for 12 tree species in coniferous and broadleaved mixed forests, Northeastern China. *PLoS One*, 13(1), e0186226. PMid:29351291. http://dx.doi.org/10.1371/journal.pone.0186226.

Huff, S., Poudel, K. P., Ritchie, M., & Temesgen, H. (2018). Quantifying aboveground biomass for common shrubs in northeastern California using nonlinear mixed effect models. *Forest Ecology and Management*, 424, 154-163. http://dx.doi.org/10.1016/j.foreco.2018.04.043.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS system for mixed models*. Cary: SAS Institute.

Mendonça, A. R., Carvalho, S. P. C., & Calegario, N. (2015). Modelos hipsométricos generalizados mistos na predição da altura de *Eucalyptus* sp. *Cerne*, 21(1), 107-115. http://dx.doi.org/10.1590/01047760201521011191.
Meng, Q., Cieszewski, C. J., Madden, M., & Borders, B. (2007). A linear mixed-effects model of biomass and volume of trees using Landsat ETM+ images. *Forest Ecology and Management*, 244(1-3), 93-101. http://dx.doi.org/10.1016/j.foreco.2007.03.056.

Oliveira, T. W. G., Lima, E. M., Sanquetta, C. R., & Dalla Corte, A. P. (2018). Ajuste simultâneo na predição de biomassa aérea em *Araucaria angustifolia* (Bertol.) Kuntze. *BIOFIX Scientific Journal*, 3(1), 137-144. http://dx.doi.org/10.5380/biofix.v3i1.58119.

Ou, G., Wang, J., Xu, H., Chen, K., Zheng, H., Zhang, B., Sun, X., Xu, T., & Xiao, Y. (2016). Incorporating topographic factors in nonlinear mixed-effects models for aboveground biomass of natural Simao pine in Yunnan, China. *Journal of Forestry Research*, 27(1), 119-131. http://dx.doi.org/10.1007/s11676-015-0143-8.

Özkale, M. R., & Kuran, Ö. (2018). Principal components regression and r-k class predictions in linear mixed models. *Linear Algebra and Its Applications*, 543, 173-204. http://dx.doi.org/10.1016/j.laa.2018.01.001.

Poudel, K. P., Flewelling, J. W., & Temesgen, H. (2018). Predicting volume and biomass change from multi-temporal Lidar sampling and remeasured field inventory data in Panther Creek watershed, Oregon, USA. *Forests*, 9(1), 1-14. http://dx.doi.org/10.3390/f9010028.

R Core Team. (2014). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved in 2018, June 10, from http://www.R-project.org/

Resende, M. D. V., Silva, F. F., & Azevedo, C. F. (2014). Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão aleatória, seleção genômica, QTL-GWAS, estatística espacial e temporal, competição, sobrevivência. Viçosa: Suprema gráfica e Editora Ltda.

Sakai, Y., Ishizuka, S., & Takenaka, C. (2013). Predicting deadwood densities of *Cryptomeria japonica* and *Chamaecyparis obtusa* forests using a generalized linear mixed model with a national-scale dataset. *Forest Ecology and Management*, 295, 228-238. http://dx.doi.org/10.1016/j.foreco.2013.01.030.

Schumacher, F. X., & Hall, F. D. S. (1933). Logarithmic expression of timber-tree volume. *Journal of Agricultural Research*, 47, 719-734.

Silva, M. L. M., Binoti, D. H. B., Gleriani, J. M., & Leite, H. G. (2009). Ajuste do modelo de Schumacher e Hall e aplicação de redes neurais artificiais para estimar volume de árvores de eucalipto. *Revista Árvore*, 33(6), 1133-1139. http://dx.doi.org/10.1590/S0100-67622009000600015.

Soares, C. P. B., Gezan, S. A., Silva, G. F., & Castro, R. V. O. (2015). Individual-tree growth and mortality models for a semideciduous Atlantic forest in Brazil. *Australian Journal of Basic and Applied Sciences*, 9, 542-552.

Temesgen, H., Monleon, V. J., & Hann, D. W. (2008). Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Canadian Journal of Forest Research*, 38(3), 553-565. http://dx.doi.org/10.1139/X07-104.

Wu, L. (2009). *Mixed effects models for complex data*. New York: Chapman and Hall/CRC.

Zeng, W. S. (2015). Using nonlinear mixed model and dummy variable model approaches to develop origin-based individual tree biomass equations. *Trees (Berlin)*, 29(1), 275-283. http://dx.doi.org/10.1007/s00468-014-1112-0.

Authors’ contributions: JCA: formal analysis, investigation, methodology; JAAS: supervision, writing – review & editing; RLCF: project administration, data curation; CPBS: writing – review & editing, visualization, supervision; CMMET: supervision, writing – review & editing; AAF: writing – original draft; ISTJ: formal analysis; validation.