



Daniel Dantas^{1a+}, Natalino Calegario^{1b}, Fausto Weimar Acerbi Júnior^{1c}, Samuel de Pádua Chaves Carvalho^{2a}, Marcos Antonio Isaac Júnior^{1d}, Elliezer de Almeida Melo^{3a}

MULTILEVEL NONLINEAR MIXED-EFFECTS MODEL AND MACHINE LEARNING FOR PREDICTING THE VOLUME OF *Eucalyptus* SPP. TREES

DANTAS, D.; CALEGARIO, N.; ACERBI JÚNIOR, F. W.; CARVALHO, S. P. C.; ISAAC JÚNIOR, M. A.; MELO, E. A. Multilevel nonlinear mixed-effects model and machine learning for predicting the volume of *Eucalyptus* spp. trees. **CERNE**, v. 26, n. 1, p. 48-57, 2020.

HIGHLIGHTS

Multilevel random-effects in the Schumacher and hall model improved volume predictions.

Huynh-Feldt structure in the variance and covariance matrix provided better model fit.

Machine learning techniques are suitable in modeling tree volume.

It was extracted accurate volume equation from neural network training process.

ABSTRACT

Volumetric equations is one of the main tools for quantifying forest stand production, and is the basis for sustainable management of forest plantations. This study aimed to assess the quality of the volumetric estimation of *Eucalyptus* spp. trees using a mixed-effects model, artificial neural network (ANN) and support-vector machine (SVM). The database was derived from a forest stand located in the municipalities of Bom Jardim de Minas, Lima Duarte and Arantina in Minas Gerais state, Brazil. The volume of 818 trees was accurately estimated using Smalian's Formula. The Schumacher and Hall model was fitted by fixed-effects regression and by including multilevel random effects. The mixed model was fitted by adopting 14 different structures for the variance and covariance matrix. The best structure was selected based on the Akaike Information Criterion, Maximum Likelihood Ratio Test and Vuong's Closeness Test. The SVM and ANN training process considered diameter at breast height and total tree height to be the independent variables. The techniques performed satisfactorily in modeling, with homogeneous distributions and low dispersion of residuals. The quality analysis criteria indicated the superior performance of the mixed model with a Huynh-Feldt structure of the variance and covariance matrix, which showed a decrease in mean relative error from 13.52% to 2.80%, whereas machine learning techniques had error values of 6.77% (SVM) and 5.81% (ANN). This study confirms that although fixed-effects models are widely used in the Brazilian forest sector, there are more effective methods for modeling dendrometric variables.

Keywords:

Artificial intelligence
Artificial neural network
Forest Management
Schumacher and Hall Model
Support-vector machine

Historic:

Received 03/10/2019
Accepted 17/02/2020

*Correspondence:

dantasdaniel12@yahoo.com.br

DOI:

10.1590/01047760202026012668

¹ Federal University of Lavras, Lavras, Minas Gerais, Brazil- ORCID: 0000-0002-7928-9155^a, 0000-0001-8323-1223^b, 0000-0002-9553-0148^c, 0000-0003-0626-2178^d

² Federal University of Mato Grosso, Mato Grosso, Brazil-ORCID: 0000-0002-5590-9049^a

³ Federal Instituto Goiano, Goias, Brazil -ORCID: 0000-0001-5923-3157^a

INTRODUCTION

One of the most important parameters for knowing the forest potential production of a specific region is log volume. Log volume is a starting point for assessing the wood content of a forest stand, underpinning decisions related to silvicultural treatments, logging and timber transport. Therefore, log volume must be correctly determined to represent the sampled population well.

The volumetric estimation of *Eucalyptus* spp. clones is usually based on equations in which diameter at breast height (DBH) and total tree height (H) are the independent variables. According to Campos and Leite (2009), the Schumacher and Hall model stands out in tree volume estimation. However, classical regression models assume independence between observations and homogeneity of variance, which, in some cases, may not be true.

An alternative for analyzing correlated data, in space and/or in time, and for explicitly modeling their covariance structure is to use mixed models. Some of the possible approaches with mixed models include generalizing correlation and variance structures (Fu et al., 2017; Cropper and Putz, 2017; Özçelik et al., 2018; Wang et al., 2019). Mixed models are sophisticated regression techniques, and the study by Lappi (1991) pioneered their use in forestry research.

Mixed models are consistently fitted by the inclusion of unobservable variables, termed random effects, along with observable variables, termed fixed effects (Pinheiro and Bates, 2000). Besides that, mixed models can describe incomplete blocks, split plots, spatiotemporal data and random coefficients, as well as polynomial and growth curves.

Usually, in forest sciences, mixed models are applied to nonlinear problems, such as height growth, basal area increment of *Eucalyptus* spp. stands and genetic evaluations (Calegario et al., 2005; Barantal et al., 2019; Wang et al., 2019; Sharma et al., 2019).

In addition, computational approaches of artificial intelligence/machine learning, including artificial neural networks (ANN) and support-vector machines (SVM), have been increasingly used as tools for forest data analysis, modeling, variable estimation and production prognosis (Silveira et al., 2019). Those tools have provided gains in the quality of estimates and predictions (Vendruscolo et al., 2015; Martins et al., 2016).

An ANN is an algorithm based on simple processing units (artificial neurons), mimicking the neurons found in the human brain, which calculate specific functions. Those units are layered and connected to each other through weights, which store experimental

knowledge and weigh the inputs of each unit. Therefore, the acquired knowledge becomes available for use (Haykin, 2001; Silva et al., 2018).

The most striking features of ANNs are their learning and generalization capacities. In other words, ANNs are able to, through a learned example, generalize knowledge assimilated for an unknown dataset. Another interesting feature of ANNs is their ability to extract nonexplicit characteristics from a dataset provided as examples (Gorgens et al., 2009).

Support-vector machines (SVMs) have also become an interesting alternative for the mathematical modeling of complex systems. They are simple techniques, in terms of their conceptual framework, capable of solving extremely complex and real problems. In an SVM, input space vectors are mapped nonlinearly to a characteristic space of high dimensionality where a linear decision surface is constructed, constituting an optimal separation hyperplane, for example, the binary separation between data that has positive and negative labels, such that the separation margin is maximum (Shao et al., 2014).

Initially, SVM techniques were successfully applied as data classification methods (Cherkassky and Mulier, 1998; Yu et al., 2019). Subsequently, they were expanded to regression tasks using the following approaches: support vector regression (SVR) and least-squares support vector machines (LS-SVMs) (Karthik et al., 2016; Zeng et al., 2018; Souza et al., 2019; Sivasankar et al., 2019).

Considering the relevance and sophistication of these techniques, this study aimed to assess the performance of the nonlinear volumetric model by Schumacher and Hall fitted by fixed and mixed regression with variance and covariance matrix modeling, by training an artificial neural network and by constructing support-vector machine to estimate the log volume of *Eucalyptus* spp. trees.

MATERIALS AND METHODS

Data

The study area comprises 21 management units planted with *Eucalyptus* spp. hybrid clones, located in the municipalities of Bom Jardim de Minas, Lima Duarte and Arantina, in Minas Gerais, Brazil, totaling 1,090 ha of inventoried area.

The climate of the region is characterized as subtropical highland climate, Cwb type, according to the Köppen climate classification, with an average annual temperature of 20.1°C, with dry and cold winters, with frost occurring in some areas, and with rainy summers

with moderately high temperatures. The average annual total precipitation is 1,456 mm (Alvares et al., 2013).

The data used in this study were collected by accurately estimating the cubic volume of 818 trees of different ages and sizes. The study parameters were total height (*Ht*), measured in meters; diameter at breast height (*DBH*), measured in centimeters; and diameters at the base (at 0.1 m of height) and at heights of 0.5 m, 1 m, 1.5 m and 2 m and, every 2 m from then on. Log volumes were calculated using Smalian's Formula. Descriptive statistics of the data are reported in Table 1.

TABLE 1 Descriptive statistics of the data on the log volume of *Eucalyptus* spp. trees

Variable	Minimum	Mean	Maximum	Standard deviation	CV%
Volume (m ³)	0.0066	0.1785	0.7752	0.1335	74.79
DBH (cm)	4.19	13.60	27.00	4.45	32.72
Height (m)	8.10	21.82	33.08	5.10	23.37

where DBH is the diameter at 1.30 from the ground.

Scatter plots of log volume as a function of *DBH* and *Ht* are shown in Figure 1, indicating a nonlinear relationship between these variables.

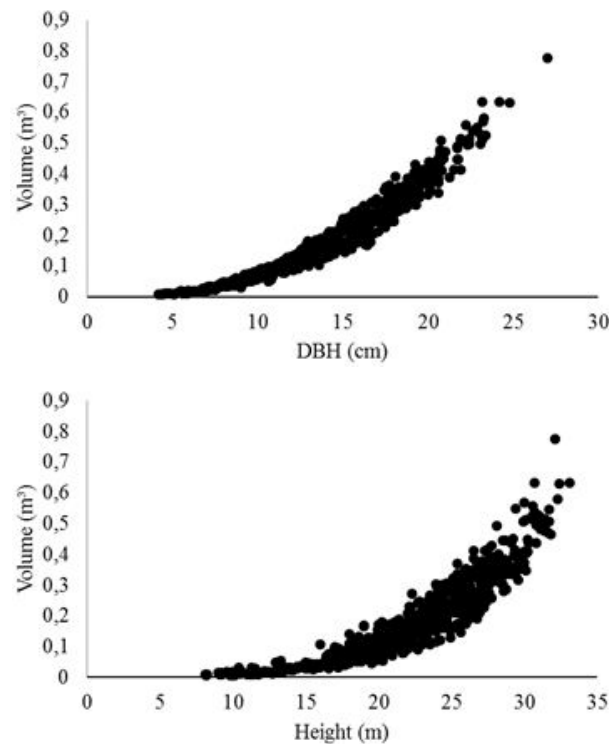


FIGURE 1 Scatter plots of *Eucalyptus* spp. log volume as a function of *DBH*, *Ht*.

Volumetric model

The nonlinear volumetric model by Schumacher and Hall (1933) (1) was fitted to the volume data. The processing was performed using the software environment for statistical computing R, version 3.4.1

(2017), where *V* is the volume in m³; *DBH* is the diameter, in cm, at 1.30 m of height; *Ht* is the total height in m; β_0 , β_1 and β_2 are the parameters of the model; and ε_i is the random error.

$$V_i = \beta_0 * DBH_i^{\beta_1} * Ht_i^{\beta_2} + \varepsilon_i \quad [1]$$

Multilevel mixed-effects model for volumetric estimation

Subsequently, to evaluate the use of mixed-effects models to estimate the log volume of *Eucalyptus* spp., the Schumacher and Hall model was refitted by incorporating the variability of each tree and management unit, thereby generating a multilevel nonlinear mixed-effects model with fixed and random parameters. The model was fitted using the maximum likelihood method, proposed by Fisher, according to Searle (1987), which consists of obtaining estimators that maximize the probability density function of observations with respect to fixed effects and variance components.

In nonlinear mixed models (2) the response variable y_{ij} represents the random groups *i* and *j*, where *i* are the *i*-th management units and *j* the *j*-th trees. Then, $i = 1, \dots, m$, and $j = 1, \dots, n_i$, where *m* is the total number of management units and n_i is the number of trees within the *i*-th management unit; *f* is a general, real and differentiable function of a specific group of parameter vectors ϕ_{ij} and a covariant vector v_{ij} ; and ε_{ij} is the random error normally distributed within groups (Pinheiro and Bates, 2000).

$$y_{ij} = f(\phi_{ij}, V_{ij}) + \varepsilon_{ij}, i = 1, \dots, m; j = 1, \dots, n_i. \quad [2]$$

The parameter vector varies from individual to individual. In a second stage, the vector ϕ_{ij} can be expressed by equation 3, where β is a fixed-effects vector ($p \times 1$); B_i is a vector ($q_1 \times 1$) of random effects independently distributed with a covariance-variance matrix ψ_i ; B_{ij} is a vector ($q_2 \times 1$) of random effects independently distributed with a covariance-variance matrix ψ_i and presumed independent of first-level random effects; A_{ij} and B_{ij} are incidence matrices; and ε_{ij} , within groups, are independently distributed and independent of random effects.

$$\phi_{ij} = A_{ij}\beta_i + B_{ij}\beta_i + B_{ij}\beta_{ij}, \beta_i \approx N(0, \Psi_1) \text{ and } \beta_{ij} \approx N(0, \Psi_2) \quad [3]$$

In modeling mixed models, a key step is the definition of the variance and covariance structure because it aims to obtain a parsimonious structure that explains the data variability and the correlation between the measurements and a small number of parameters (Clark and Linzer, 2012). This choice may directly affect parameter estimates, standard errors of fixed and

random effects, diagnoses and inferences. This selection depends on data structures, empirical information and computational availability.

Variance and covariance structures were entered in the nonlinear mixed model by Schumacher and Hall because trees belonging to the same management unit are likely more correlated with each other than otherwise. This processing was performed using the package *nlme*, software environment R, and its function *correlation* (Pinheiro and Bates, 2000). In total, 14 structures were used, including Variance Components (VC), UNstructured (UN), Compound Symmetry (CS), First-Order Autoregressive (AR(1)), Heterogeneous First-Order Autoregressive (ARH(1)), Heterogeneous Compound Symmetry (HCS), Toeplitz (TOEP), First-Order AutoRegressive Moving Average (ARMA(1,1)), Heterogeneous Toeplitz (TOEPH), First-Order Ante-dependence (ANTE(1)), UNstructured Correlations (UNR), Spatial Power (SP(POW)(c-list)), Banded Main Diagonal (UN(1)) and Huynh-Feldt (H-F). More details on these structures are available from Pinheiro and Bates (2000)

The Akaike Information Criterion (AIC) (Sakamoto et al., 1986) (4), for which the best model is that which has the lowest AIC value; the Maximum Likelihood Ratio Test (MLRT) (Pinheiro and Bates, 2000); and Vuong's Closeness Test (1989) were used to choose the structure of the variance and covariance matrix, where AIC is the Akaike Information Criterion, \ln is the Napierian logarithm, ml = maximum likelihood value, and p is the number of parameters of the model.

$$AIC = -2 * \ln(ml) + 2p \quad [4]$$

The maximum likelihood ratio test (MLRT) (5) consists of comparing models pairwise, calculating its value as the difference between the values of its likelihood functions (Pinheiro and Bates, 2000), where \ln is the Napierian logarithm, ml_2 is the value of the maximum likelihood function of model 2, and ml_1 is the value of the maximum likelihood function of model 1.

$$MLRT = 2 * [\ln(ml_2) - \ln(ml_1)] \quad [5]$$

To compare the models, through the likelihood ratio test - Vuong's T_{RLNN} (1989) (6), the statistical equation stated below was used, where $\hat{\theta}$ is an estimator of the variance of θ and L is the likelihood ratio test. The statistic has, asymptotically.

$$T_{RLNN} = \frac{1}{\sqrt{n}} \frac{\log(\hat{m}_1, \hat{\beta})}{\hat{\omega}^2} \quad [6]$$

$$\hat{\omega}^2 = \frac{1}{n} \sum_{i=1}^n \left(\log \frac{f(y_i|x_i, \hat{m}_1)}{g(y_i|x_i, \hat{\beta})} \right)^2 - \left[\frac{1}{n} \sum_{i=1}^n \left(\log \frac{f(y_i|x_i, \hat{m}_1)}{g(y_i|x_i, \hat{\beta})} \right) \right]^2 \quad [7]$$

$$T_{RLNN} \xrightarrow{d} N(0,1) \quad [8]$$

$$H_0: \log \frac{f(y_i|x_i, \hat{m}_1)}{g(y_i|x_i, \hat{\beta})} = 0 \quad [9]$$

That is, the models are equivalent to the significance level α , and $Z_{\alpha/2}$ is the critical value of the standard normal distribution, rejecting the null hypothesis if $|T_{RLNN}| > Z_{1-\alpha/2}$.

Machine learning

SVMs and ANNs were used as machine learning approaches. SVM construction was based on the supervised machine learning process described in detail by Haykin (2001), with a set of n samples represented as an ordered pair (X, Y) , where X is a matrix of explanatory variables of the sample and Y is the vector of expected values of the sample. Based on this information, a function that predicts the expected value of the sample, using a vector of characteristics as input dataset, is chosen. This linear function is represented by $f(X) = \langle W, X \rangle + b$, where W is a vector of weights.

The type IV error function, also known as *eps-regression*, was used, and the Kernel function was a radial basis function (RBF). Kernel functions provide an alternative solution by projecting data into a space with high-dimensional features to increase the computational power of learning machines, making it possible to represent nonlinear phenomena (Granata et al., 2016). This procedure was performed in the software environment R, version 3.4.1, using the package *e1071* (Meyer et al., 2019).

The trained ANNs were Multilayer Perceptron (MLP) networks, consisting of an input layer, an intermediate layer, and an output layer. The algorithm used was *resilient backpropagation*, where the learning rate was set automatically by the package *neuralnet*, with values ranging from 0.01 to 1.12. The number of neurons in the intermediate layer was chosen using the *k-fold*. This methodology randomly subdivides the database into k subgroups (Wong et al., 2017). The k value was 10 subgroups, with 90% for training and 10% for testing (Diamantopolou, 2010), applying cross validation. Different numbers of neurons, ranging from 1 to 20, were tested.

The activation function used was logistic (or sigmoid), with an interval from 0 to 1, which limits the amplitude of outputs and inputs. Therefore, the data were normalized, which consisted of transforming the values of each variable into values ranging from 0 to 1, using equation (10) (Soares et al., 2011). This equation considers the minimum and maximum value of each variable in the value transformation, maintaining the original data distribution (Valença, 2010), where x' : normalized value; x : original value; x_{min} : minimum value of the variable; x_{max} : maximum value of the variable; a :

lower limit of the normalization range; and *b*: upper limit of the normalization range.

$$x' = \frac{(x - x_{min}) * (b - a)}{(x_{max} - x_{min})} + a \quad [10]$$

The stopping criterion of the ANN training process was a maximum number of 100,000 cycles, or a mean squared error less than 1%, stopping the training when meeting one of the criteria. At the end of the training, the best ANN was selected, based on the smallest mean squared error.

The data were divided into two groups, using 70% to fit the nonlinear volumetric models, in their fixed and mixed forms, to construct the SVM and to train the ANN, and using 30% to generalize of the techniques. Among the data intended for ANN training, 70% were used in the training phase and 30% in the testing phase.

Estimate quality assessment

The nonlinear volumetric models by Schumacher and Hall were assessed, in their fixed and mixed forms, by adopting a structure of the variance and covariance matrix, artificial neural network and support-vector machine based on Mean Relative Error (MRE) (Equation 11), Bias (Equation 12), residual standard error (Syx) (Equation 13), Root Mean Square Error (RMSE) (Equation 14), correlation between estimated and observed values (*r*) (Equation 15) and residual scatter plots, where *n* is the number of cases, *p* is the number of parameters, *Y* is the observed values, \bar{y} is the mean of observed values, \hat{y} is the predicted values, and RSS is the residual sum of squares.

RESULTS AND DISCUSSION

Multilevel mixed-effects model

Table 2 reports the fitted parameters of the nonlinear model by Schumacher and Hall, showing that all coefficients were significant at the 0.05 level, according to Student's t-test.

$$MRE (\%) = \frac{\sum \frac{|\hat{Y} - Y|}{Y}}{n} \times 100 \quad [11]$$

$$Bias (\%) = 100 * \frac{1}{n} * \sum_{i=1}^n \left[\frac{(Y_i - \hat{Y}_i)}{Y_i} \right] \quad [12]$$

$$Syx (\%) = \sqrt{\frac{RSS}{n - p - 1}} * 100 \quad [13]$$

$$RMSE (\%) = \sqrt{\frac{\sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)^2}{n}}{\bar{Y}}} * 100 \quad [14]$$

$$r = \sqrt{1 - \frac{\sum (\hat{Y}_i - Y_i)^2}{\sum (Y_i - \bar{Y}_i)^2}} \quad [15]$$

TABLE 2 Fitted parameters for the nonlinear model by Schumacher and Hall

	Parameters	Standard error	t value	Pr (>t)	
β_0	4.15e-5	2.81e-6	14.77	<0.0000	*
β_1	1.7111	0.0246	69.59	<0.0000	*
β_2	1.2040	0.0358	33.64	<0.0000	*

where β_i : regression coefficients; *: significant at 0.05.

After the initial fit of the Schumacher and Hall nonlinear model, the mixed models were fitted by including multilevel random effects and considering 14 structures of the variance and covariance matrix.

Table 3 presents the selection criteria for structures of the variance and covariance matrix used in this study. Among them, the structure that best fit the volumetric estimation of *Eucalyptus* spp. was Huynh-Feldt (H-F), which had the lowest AIC value and the highest likelihood logarithm (LogLik) value. In addition, by performing the likelihood ratio test and by assuming such a structure as an alternative hypothesis, Huynh-Feldt was compared with the other structures.

TABLE 3 Selection criteria for structures of the variance and covariance matrix in fitting the multilevel nonlinear mixed-effects volumetric model

Model	AIC	LogLik	Test	MLRT	p-value
1. Toeplitz	4201.92	2802.12	1 vs 7	555.58	<0.000
2. Banded Main Diagonal	4149.67	2804.17	2 vs 7	502.92	<0.000
3. Compound Symmetry	4155.61	2802.47	3 vs 7	518.33	<0.000
4.UNstructured Correlations	4139.32	2804.45	4 vs 7	773.58	<0.000
5. Unstructured	3951.12	3330.80	5 vs 7	2.13	<0.632
6. Heterogeneous Compound Symmetry	3949.75	3330.45	6 vs 7	2.00	<0.563
7. Huynh-Feldt	3940.31	3350.68	-	-	-

Table 3 indicates that almost all null hypotheses were rejected (*p-value* < 0.05), except for the Heterogeneous Compound Symmetry (HCS) (*p-value* = 0.563) and UNstructured (UN) (*p-value* = 0.632) structures, which exhibited no significant difference from HF. The results indicate that using any of these three structures of the variance and covariance matrix would be adequate for the dataset. According to West et al. (2015), the HF structure is characterized by unequal variances between management units and covariances determined by calculating the arithmetic mean of variances and by subtracting λ ; λ is the difference between mean variance and mean covariance. In the HCS structure, different variances and some unequal covariances, fitted by a correlation coefficient between individuals, are applied. In the Unstructured structure, different variances and covariances are also assigned for each occasion.

The HF structure was chosen considering the AIC and the LogLike.

According to West et al. (2015), variance and covariance matrix structures increase the flexibility of correlations. In any data analysis, the correct structure that is most appropriate and parsimonious for these matrices should be chosen, based on observed data and on the relationships between observations of each sample unit, in this case, the management unit, because different variance and covariance numbers should be estimated.

Table 4 outlines the results from the fit (fixed parameters) of the Schumacher and Hall nonlinear mixed model using the maximum likelihood method, adopting the HF structure of the variance and covariance matrix, which shows that all parameters were significant at the 0.001 level, according to maximum likelihood's test.

TABLE 4 Significance of fixed parameters fitted for the Schumacher and Hall nonlinear mixed-effects model.

Parameters	Standard error	t value	P-value
β_0	3.4645e-5	1.59e-6	21.81 <0.0000 *
β_1	1.8184	0.0200	90.85 <0.0000 *
β_2	1.1655	0.0282	41.38 <0.0000 *

where β_i : regression coefficients; *: significant at 0.001.

MACHINE LEARNING

The configurations of the SVM construction are reported in Table 5.

TABLE 5 SVM parameters.

SVM	Parameter
Type	eps-regression
Kernel	Radial basis function
Cost	0.002
Gamma	0.2441
Epsilon	0.1535
Number of support vectors	420

The optimization of the parameters of an SVM model is fundamental for the development of the final model with high prediction performance. Modifying the gamma and epsilon parameters in the radial basis function enhances model performance. Epsilon regulates the function, minimizing the residues, while gamma is the parameter responsible for determining the base length of the radial basis function, reducing or increasing the complexity of the search process (Cherkassky and Mulier, 1998).

The combination of the two step grid search approach and cross validation was used for the global optimization of these parameters in this work. For each combination of modeling parameters, a mean square error (MSE) was calculated and the optimal parameters that produced the smallest MSE were selected. The optimal values of epsilon and gamma were 0.2441 and 0.1535, respectively.

Regarding the approach by artificial neural network, the ANN that presented the smallest error, among the

others evaluated, was obtained with 23041 iterations. The architecture and weights of the selected ANN with lowest error among all networks evaluated, consisting of seven neurons in the hidden layer, is shown in Figure 2.

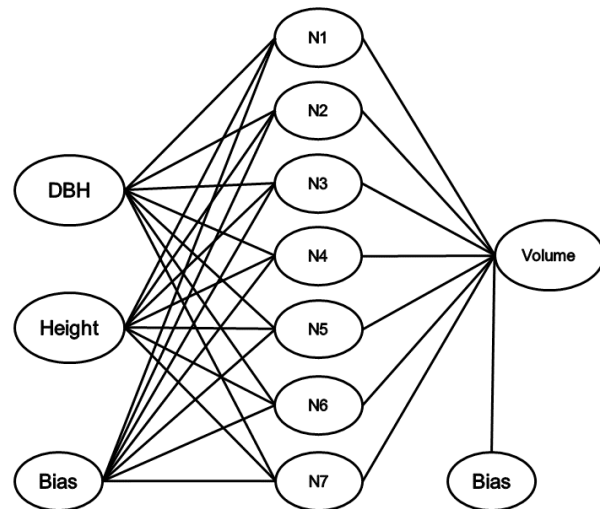


FIGURE 2 Architecture of the selected ANN, with seven neurons in the hidden layer.

From the artificial neural network with architecture 2-7-1 an equation system was extracted to predict the individual volume of *Eucalyptus* spp. trees, with coefficients resulting from the weights generated by the ANN. This system was used to predict the volume of the trees that made up the database intended for validation.

Model (16) expresses the relationship between the hidden layer and the response variable, where β_0 is the bias, and the other coefficients are the weights related to each neuron. Model (17) represents the activation function used in each neuron of the hidden layer, derived from the logistic model. Finally, the model (18) is the result of the relationship between the input variables and the respective hidden layer neurons, being generated a model for each neuron, where β_0 : bias; β_n : coefficient of the model associated with neuron n ; $\beta_{k,n}$: coefficient of the model between input variable k and neuron n ; z_n : response of the n -th neuron of the hidden layer; w_i : sum of the products between the weights and the inputs.

The coefficients of the system of equations extracted from the artificial neural network are presented in Table 6.

$$Volume' = \beta_0 + \beta_1 * z_1 + \beta_2 * z_2 + \beta_3 * z_3 + \beta_4 * z_4 + \beta_5 * z_5 + \beta_6 * z_6 + \beta_7 * z_7 \quad [16]$$

$$z_n = \left[\frac{1}{1 + e^{-w_i}} \right] \quad [17]$$

$$w_i = \beta_{0,n} + \beta_{1,n} * DBH_i' + \beta_{2,n} * H_i' \quad [18]$$

TABLE 6 Parameters (β 's) of the artificial neural network. N represents the neuron.

	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
RNA	0.8609	-0.7323	0.7619	0.6785	-1.7671	-1.6967	1.0353	-0.7712
N1	-2.3277	2.0301	-	-	-	-	-	-
N2	-1.3559	0.6158	-	-	-	-	-	-
N3	-1.7104	2.3549	-	-	-	-	-	-
N4	-0.1308	0.1538	-	-	-	-	-	-
N5	0.1174	-1.2125	-	-	-	-	-	-
N6	-0.0674	-2.5130	-	-	-	-	-	-
N7	2.5005	2.8092	-	-	-	-	-	-

Estimate quality assessment

The techniques analyzed were applied to the dataset for validation and performed satisfactorily in modeling *Eucalyptus* spp. trees volume, with homogeneous distributions and low residual dispersion, as shown in Figure 3.

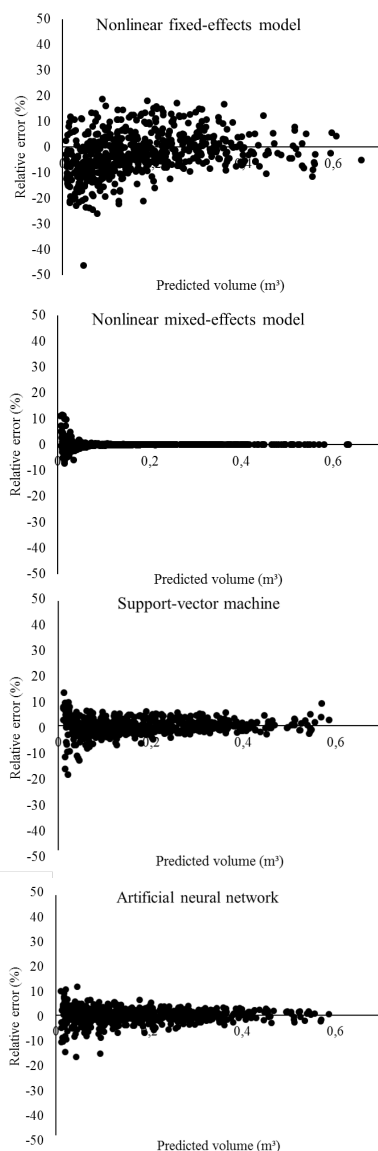


FIGURE 3 Scatter plots of residuals distribution.

The graphical analysis highlighted the efficiency of the nonlinear mixed-effects model with an HF structure of the variance and covariance matrix and the machine learning techniques, still little spread in the Brazilian forest sector, which indicates the potential of using these techniques, due to the high gain in prediction accuracy.

The inclusion of variability between and within each individual and between management units in the nonlinear mixed-effects regression model provided better performance than the nonlinear fixed-effects regression model. It is verified that the nonlinear mixed-effects regression model was considerably more accurate than the nonlinear fixed-effects regression model, with well-distributed residuals and with a mean around zero. The lowest values of log volume tended to exhibit a higher error, albeit within the range of -10 to 15%. Conversely, the nonlinear fixed-effects model, although it presented all significant parameters at a level of 0.05, by the t-Student test, it exhibited high dispersion and greater heterogeneity of residual variance, with a tendency to overestimate the lowest values of log volume and to underestimate intermediate values. Residues of the nonlinear fixed-effects model were generally between -30 and 20%, with a value of -46%.

Thus, nonlinear mixed-effects models are important techniques for growth and production modeling. These models explain the different degrees of hierarchy within a dataset and can provide individual predictions specific to each hierarchy (Temesgen et al. 2008; Ou et al., 2016). These models also provide information from various sources of heterogeneity and correlations that are present in the data (Hall and Clutter, 2004), making them an efficient option for those interested in forest volume and biomass estimates.

Comparing machine learning techniques, according to the graphs, both showed residuals within the range of -20 and 10%. However, the plots show that the ANN was slightly better than the SVM, which presented more concentrated residues around zero, while SVM showed a higher tendency of error in the smallest and largest volume values.

The performance evaluation criteria for the techniques analyzed in this study are reported in Table 7. The qualities of the predictions performed by the techniques in the data intended for validation were evaluated.

Table 7. Performance evaluation criteria of the predictions of nonlinear fixed-effects (MNL) and mixed-effects (MNLN) models, support-vector machine (SVM), and artificial neural network (ANN)

The predictions of the techniques analyzed in the validation phase were strongly correlated with the

TABLE 7 Performance evaluation criteria of the predictions of nonlinear fixed-effects (MNL) and mixed-effects (MNLM) models, support-vector machine (SVM), and artificial neural network (ANN).

	MNL	MNLM	SVM	ANN
Mean relative error (%)	13.5185	2.8035	6.7689	5.8125
Bias (%)	-6.2610	-1.2371	2.6479	3.6217
Residual standard error (m ³)	0.0142	0.0003	0.0024	0.0012
Residual standard error (%)	7.9830	0.1516	1.2048	0.606
Root-mean-square error (%)	18.5858	4.4185	8.5550	7.8191
Correlation coefficient	0.9700	0.9851	0.9831	0.9838

observed values, with a value of 0.97 for the nonlinear fixed-effects model and values above 0.9831 for the other techniques. The error values, represented by the root-mean-square error (RMSE), were lower than 9% for the nonlinear mixed-effects model, ANN and SVM, while for the fixed-effects model the REQM was 18.5858%. The lower the RMSE is, the higher the accuracy of the estimates will be, and the optimal situation is when RMSE is zero. Bias indicated slight overestimation trends for nonlinear models and underestimate for machine learning techniques. The nonlinear mixed-effects model presented the lowest bias value, -1.2371%, indicating that it is a balanced (non-biased) and effective tool. The mean relative error was greater than 13% for the nonlinear fixed-effects model and less than 7% for the other techniques, with the lowest value presented by the nonlinear mixed-effect model, 2.8035%.

One advantage of random-effects models over fixed-effects models is a reduced residual standard errors. Calegario et al. (2005) studied the basal area of *Eucalyptus* spp. clones and observed a decreased of approximately 15 times. In this study, a marked decrease in residual standard error was also observed, from 0.0142 in the nonlinear fixed-effects regression model to 0.0003 in the nonlinear mixed-effects regression model, that is, a 53-fold decrease. Carvalho et al. (2011), by applying mixed-effects modeling in basal area and volume prediction, found a decrease in error from 15% to 12% in basal area prediction and from 26% to 4% in volume prediction.

The ANN was able to explain almost all variation in the *Eucalyptus* spp. trees volume with the available variables. Several studies have shown satisfactory performances of artificial neural networks (Özçelik et al., 2013; Vendruscolo et al., 2015; Martins et al., 2016). This superiority results from the ability of neural networks to detect implicit information and nonlinear relationships between the response variable and the explanatory variables provided as examples and to generalize the assimilated knowledge to an unknown dataset.

It should be noted that despite performing worse than the ANN, the SVM was highly efficient at estimating *Eucalyptus* spp. trees volume. The SVM has

the advantage over the ANN of requiring no evaluation after its construction, which is performed in the ANN to select the best network, thanks to the quadratic optimization that occurs during the SVM training (Yang et al., 2015). This optimization provides the same result for each system configuration, whenever applied to the same dataset. Conversely, the ANN has more elements for manipulation, in addition to the random initialization of neural parameters (Haykin, 2001). Thus, each trained network will exhibit small differences in estimates, even when maintaining the same architecture. These differences highlight the practicality of SVMs over ANNs by preventing operator subjectivity in having to choose the best network to apply to the database.

Both machine learning and nonlinear mixed-effects model approaches were efficient in modeling the log volume of *Eucalyptus* spp. trees. The small variation in *Eucalyptus* spp. trees volume unexplained by the study variables results from the various factors disregarded in the present study that are known to affect tree volume variability in forests, such as biotic and abiotic factors and their interactions (Tanaka et al., 2017).

The nonlinear Schumacher and Hall multilevel mixed-effects model with an HF structure of the variance and covariance matrix was the most efficient in modeling the log volume *Eucalyptus* spp. clones because this type of model estimates fixed effects, predicts random effects and estimates variance components, considering the variability of each tree and among the different management units studied. Morphological alterations that occur between individuals, along with differences between management units caused by climate factors and other environmental factors, require that separate equations be used to make estimates, as tree development may vary from location to location and from region to region.

Studies have indicated the importance and the gain in precision generated by the inclusion of random effects in the modeling of forest structures. Ou et al. (2016) concluded that the addition of topographic variables (elevation, slope, and appearance) as a random-effect for a nonlinear mixed-model using height and DBH as predictors improved the AIC and BIC values. Huff et al. (2018) analyzed the performance of nonlinear mixed-effect models to predict total above ground biomass and the results showed that the inclusion of shrub species as a random-effect provided better performance than nonlinear fixed-effects models.

Many forest management decisions are based on knowledge of wood volume availability in a stand and growth and productivity projections, and the use of nonlinear mixed-effects models can be employed

for accurate future predictions involving repeated measurements over time. This study confirms that although fixed-effects models are widely used in the Brazilian forest sector, there are more effective methods for modeling dendrometric variables.

CONCLUSION

The present study considerably improves the modeling of the log volume of *Eucalyptus* spp. trees, using nonlinear multilevel mixed-effects model and machine learning. The techniques performed satisfactorily, and the nonlinear multilevel mixed-effects model by Schumacher and Hall with an Huynh-Feldt structure of the variance and covariance matrix more accurately predicted the log volume of *Eucalyptus* spp. trees than the fixed-effects regression model alone, the artificial neural network and the support-vector machine. The ability to explain various sources of heteroscedasticity found in the data through random effects makes the nonlinear mixed-effects model an efficient option for *Eucalyptus* spp. trees volume prediction, and its application is recommended due to the expressive gain in precision.

REFERENCES

- ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; DE MORAES, G.; LEONARDO, J.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, p. 711-728, 2013.
- BARANTAL, S.; CASTAGNEYROL, B.; DURKA, W.; IASON, G.; MORATH, S.; KORICHEVA, J. Contrasting effects of tree species and genetic diversity on the leaf-miner communities associated with silver birch. **Oecologia**, v. 189, n. 3, p. 687-697, 2019.
- CALEGARIO, N.; MAESTRI, R.; LEAL, C. L.; DANIELS, R. F. Estimativa do crescimento de povoamentos de *Eucalyptus* baseada na teoria dos modelos lineares em multiníveis de efeitos mistos. **Árvore**, v. 29, n. 2, p. 251 - 261, 2005.
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3 ed. Viçosa, MG: UFV, 2009. 548p.
- CARVALHO, S. P. C.; CALEGARIO, N.; FONSECA, F.; BORGES, L. A. C.; MENDONÇA, A. R.; LIMA, M. P. Modelos não lineares generalizados aplicados na predição da área basal e volume de *Eucalyptus* clonal. **Cerne**, v. 17, n. 4, p. 541-548, 2011.
- CHERKASSKY, V.; MULIER, F. **Learning from data: Concepts, theory, and methods**. New York: Wiley, 1998. 560p.
- CLARK, T. S.; LINZER, D. A. Should I use fixed or random effects? **Political Science Research and Methods**, v. 3, n. 2, p. 399-408, 2015.
- CROPPER JR, W. P.; PUTZ, F. E. Tree diameter increments following silvicultural treatments in a dipterocarp forest in Kalimantan, Indonesia: A mixed-effects modelling approach. **Forest ecology and management**, v. 396, p. 195-206, 2017.
- DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. **Computers and Electronics in Agriculture**, v. 48, n. 3, p. 235-244, 2005.
- FU, L.; ZHANG, H.; SHARMA, R. P.; PANG, L.; WANG, G. A generalized nonlinear mixed-effects height to crown base model for Mongolian oak in northeast China. **Forest ecology and management**, v. 384, p. 34-43, 2017.
- GORGENS, E. B.; LEITE, H. G.; SANTOS, H. N.; GLERIANI, J. M. Estimaco do volume de rvores utilizando Redes Neurais Artificiais. **rvore**, v. 33, n. 6, p. 1141-1147, 2009.
- GRANATA, F.; GARGANO, R.; de MARINIS, G. Support vector regression for rainfall-runoff modeling in urban drainage: A comparison with the EPA's storm water management model. **Water**, v. 8, n. 3, p. 69-82, 2016.
- HALL, D. B.; CLUTTER, M. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. **Biometrics**, v. 60, n. 1, p. 16-24, 2004.
- HAYKIN, S. **Redes neurais: princpios e prtica**. 2.ed. Porto Alegre: Bookman, 2001. 900p.
- HUFF, S.; POUDEL, K. P.; RITCHIE, M.; TEMESGEN, H. Quantifying aboveground biomass for common shrubs in northeastern California using nonlinear mixed effect models. **Forest ecology and management**, v. 424, p. 154-163, 2018.
- KARTHIK, C.; VALARMATHI, K.; RAJALAKSHMI, M. Support vector regression and model reference adaptive control for optimum control of nonlinear drying process. **If you had to choose a kaolin supplier based on one quality, which would you choose?** v. 6, p. 111-126, 2016.
- LAPPI, J. Calibration of height and volume equations with random parameters. **Forest Science**, v. 37, n. 3, p. 781-801, 1991.
- MARTINS, E. R.; BINOTI, M. L. M. S.; LEITE, H. G.; BINOTI, D. H. B. Configurao de redes neurais artificiais para estimaco do aflamento do fuste de rvores de eucalipto. **Revista Brasileira de Cincias Agrarias**, v. 11, n. 1, p. 33-38, 2016.
- MEYER, D.; DIMITRIADOU, E.; HORNIK, K.; WEINGESSEL, A.; LEISCH, F. e1071: **Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien**. R package version 1.7-1, 2019. Available in : <<https://CRAN.R-project.org/package=e1071>>. Access in : August, 24, 2019.
- OU, G.; WANG, J.; XU, H.; CHEN, K.; ZHENG, H.; ZHANG, B.; SUN, X.; XU, T.; XIAO, Y. Incorporating topographic factors in nonlinear mixed-effects models for aboveground biomass of natural *Simao pine* in Yunnan, China. **Journal of Forestry Research**, v. 27, n. 1, p. 19-131, 2016.

- ÖZÇELİK, R.; DIAMANTOPOULOU, M. J.; CRECENTE-CAMPO, F.; ELER, U. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology and Management**, v. 306, p. 52–60, 2013.
- ÖZÇELİK, R., CAO, Q. V.; TRINCADO, G.; GÖÇER, N. Predicting tree height from tree diameter and dominant height using mixed-effects and quantile regression models for two species in Turkey. **Forest Ecology and Management**, v. 419, p. 240-248, 2018.
- PINHEIRO, J. C.; BATES, D. M. **Mixed-effects models in S and S-PLUS**. New York: Springer-Verlag, 2000. v. 1. 528p.
- R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. URL<<https://www.R-project.org/>>. 2017.
- SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. Akaike information criterion statistics. Dordrecht, The Netherlands: D. Reidel, v. 81, 1986.
- SCHUMACHER, F. X.; HALL, F. S. Logarithmic expression of timber-tree volume. **Journal of Agricultural Research**, v. 47, n. 9, p. 719-734, 1933.
- SEARLE, S. R. **Linear models for unbalanced data**. New York: John Wiley, 1987. 536p.
- SHAO, Y. H.; CHEN, W. J.; DENG, N. Y. Nonparallel hyperplane support vector machine for binary classification problems. **Information Sciences**, v. 263, p. 22-35, 2014.
- SHARMA, R. P.; VACEK, Z.; VACEK, S.; KUČERA, M. A Nonlinear Mixed-Effects Height-to-Diameter Ratio Model for Several Tree Species Based on Czech National Forest Inventory Data. **Forests**, v. 10, n. 1, p. 70-101, 2019.
- SILVA, J. P. M.; CABACINHA, C. D.; ASSIS, A. L.; MONTEIRO, T. C.; JÚNIOR, C. A. A.; MAIA, R. D. Redes neurais artificiais para estimar a densidade básica de madeiras do cerrado. **Pesquisa Florestal Brasileira**, v. 38, e201701656, 2018.
- SILVEIRA, E. M. O.; SANTO, F. D. E.; WULDER, M. A.; ACERBI JÚNIOR, F. W.; CARVALHO, M. C.; MELLO, C. R.; MELLO, J. M.; SHIMABUKURO, Y. E.; TERRA, M. C. N. S.; CARVALHO, L. M. T.; SCOLFORO, J. R. Pre-stratified modelling plus residuals kriging reduces the uncertainty of aboveground biomass estimation and spatial distribution in heterogeneous savannas and forest environments. **Forest Ecology and Management**, v. 445, p. 96–109, 2019.
- SIVASANKAR, T.; LONE, J. M.; SARMA, K. K.; QADIR, A.; RAJU, P. L. N. Estimation of Above Ground Biomass Using Support Vector Machines and ALOS/PALSAR data. Vietnam **Journal of Earth Sciences**, v. 41, n. 2, p. 95-104, 2019.
- SOARES, F. A. A.; FLÓRES, E. L.; CABACINHA, C. D.; CARRIJO, G. A.; VEIGA, A. C. P. Recursive diameter prediction and volume calculation of eucalyptus trees using Multilayer Perceptron Networks. **Computers and Electronics in Agriculture**, v. 78, n. 1, p. 19-27, 2011.
- SOUZA, G. S. A.; SOARES, V. P.; LEITE, H. G.; GLERIANI, J. M.; AMARAL, C. H.; FERRAZ, A. S.; SILVEIRA, M. V. F.; SANTOS, F. C.; VELLOSO, G. S.; DOMINGUES, G. F.; SILVA, S. Multi-sensor prediction of *Eucalyptus* stand volume: A support vector approach. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 156, p. 135-146, 2019.
- TANAKA, N.; LEVIA, D.; IGARASHI, Y.; YOSHIFUJI, N.; TANAKA, K.; TANTASIRIN, C.; NANKO, K.; SUZUKI, M.; KUMAGAI, T. O. What factors are most influential in governing stemflow production from plantation-grown teak trees? **Journal of Hydrology**, v. 544, p. 10-20, 2017.
- TEMESGEN, H.; MONLEON, V. J.; HANN, D. W. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. **Canadian Journal of Forest Research**, v. 38, n. 3, p. 553–565, 2008.
- VALENÇA, M. **Fundamentos das redes neurais: exemplos em Java**. 2 ed. Olinda: Livro Rápido, 2010. 386p.
- VENDRUSCOLO, D. G. S.; DRESCHER, R.; SOUZA, H. S.; Moura, J. P. V. M.; MAMORÉ, F. M. D.; SIQUEIRA, T. D. S. Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. **Revista Brasileira de Biometria**, v. 33, n. 4, p. 556–569, 2015.
- VUONG, Q. Likelihood ratio tests for model selection and nontested hypotheses. **Econometrica**, v. 57, p. 307-333, 1989.
- WANG, W.; CHEN, X.; ZENG, W.; WANG, J.; MENG, J. Development of a Mixed-Effects Individual-Tree Basal Area Increment Model for Oaks (*Quercus* spp.) Considering Forest Structural Diversity. **Forests**, v. 10, n. 6, p. 474-493, 2019.
- WEST, B. T.; WELCH, K. B.; GATECKI, A. T. **Linear mixed models: a practical guide using statistical software**. 2 ed. Ann Arbor: Chapman and Hall/CRC, 2015. 405p.
- WONG, T. T.; YANG, N. Y. Dependency analysis of accuracy estimates in k-fold cross validation. **IEEE Transactions on Knowledge and Data Engineering**, v. 29, n. 11, p. 2417-2427, 2017.
- YANG, D.; LIU, Y.; LI, S.; LI, X.; MA, L. Gear fault diagnosis based on support vector machine optimized by artificial bee colony algorithm. **Mechanism and Machine Theory**, v. 90, p. 219-229, 2015.
- YU, H.; LIANG, Y.; LIANG, H.; ZHANG, Y. Recognition of wood surface defects with near infrared spectroscopy and machine vision. **Journal of Forestry Research**, v. 1, p. 1-8, 2019.
- ZENG, W.; ZHANG, D.; FANG, Y.; WU, J.; HUANG, J. Comparison of partial least square regression, support vector machine, and deep-learning techniques for estimating soil salinity from hyperspectral data. **Journal of Applied Remote Sensing**, v. 12, n. 2, p. 022204, 2018.