Journal of Environmental Management 249 (2019) 109368

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

Computational techniques applied to volume and biomass estimation of trees in Brazilian savanna

Jeferson Pereira Martins Silva[a,*], Mayra Luiza Marques da Silva[b], Evandro Ferreira da Silva[a], Gilson Fernandes da Silva[a], Adriano Ribeiro de Mendonça[a], Christian Dias Cabacinha[c], Emanuel França Araújo[a], Jeangelis Silva Santos[a], Giovanni Correia Vieira[d], Maria Naruna Felix de Almeida[a], Márcia Rodrigues de Moura Fernandes[a]

[a] Federal University of Espírito Santo/UFES, Department of Forestry and Wood Science, Avenue Governor Lindemberg; 316, 29550-000, Jerônimo Monteiro, ES, Brazil
[b] Federal University of São João Del-Rei – UFSJ, Sétimo Moreira Martins, 188, 35702-031, Sete Lagoas, MG, Brazil
[c] Federal University of Minas Gerais/UFMG, Institute of Agrarian Sciences, Avenue University, 1000, 39404-547, Montes Claros, MG, Brazil
[d] Federal Institute of Rondônia/IFRO, Campus Ji-Paraná, Rio Amazonas, 151, 78900-730, Ji-Paraná, RO, Brazil

ARTICLE INFO

Keywords:
Machine learning
Model mixed effect
Forest management
Artificial intelligence techniques
Cerrado

ABSTRACT

The Brazilian Savannah, known as Cerrado, has the richest flora in the world among the savannas, with a high degree of endemic species. Despite the global ecological importance of the Cerrado, there are few studies focused on the modeling of the volume and biomass of this forest formation. Volume and biomass estimation can be performed using allometric models, artificial intelligence (AI) techniques and mixed regression models. Thus, the aim of this work was to evaluate the use of AI techniques and mixed models to estimate the volume and biomass of individual trees in vegetation of Brazilian central savanna. Numerical variables (diameter at height of 1.30 m of ground, total height, volume and biomass) and categorical variables (species) were used for the training and fitting of AI techniques and mixed models, respectively. The statistical indicators used to evaluate the training and the adjustment were the correlation coefficient, bias and Root mean square error relative. In addition, graphs were elaborated as complementary analysis. The results obtained by the statistical indicators and the graphical analysis show the great potential of AI techniques and mixed models in the estimation of volume and biomass of individual trees in Brazilian savanna vegetation. In addition, the proposed methodologies can be adapted to other biomes, forest typologies and variables of interest.

1. Introduction

The savanna of Central Brazil, known locally as Cerrado, has the richest flora among Earth's savannas, with a high number of endemic species (Marris, 2005; Myers et al., 2000). However, severe deforestation occurs in the Brazilian Cerrado, being potentiated by the use of fire to remove remaining vegetation, which is common due to its ease and low cost; this expels carbon fixed in gases and contributes to the greenhouse effect (Barni et al., 2016; Cunha et al., 2016; Günther et al., 2018; Kuch, 2017; Ledo et al., 2018; Moon et al., 2013; Torres et al., 2015; Tozer and Klenk, 2018; Yang et al., 2016).

Even though the global ecological importance of the Cerrado has been acknowledged, there have been few studies focused on volume and biomass quantification of this forest formation (Ribeiro et al., 2011; Sales et al., 2007). One of the reasons for this lack of studies is the specific growth characteristics of Brazilian Cerrado trees, such as the tortuosity and irregularity of the stems from the base to the canopy, which make it difficult to collect data and estimate volume and biomass (Nunes and Görgens, 2016; Özçelik et al., 2010).

Volume and biomass estimation can be performed with allometric models. In these models, the volume and biomass values are estimated as a function of dendrometric variables obtained in the field, such as diameter at a height of 1.30 m from the ground (DBH) and total height (H), or the basic density of wood (BD) (Ferraz et al., 2014; Koala et al., 2017; Ratuchne et al., 2016; Somogyi et al., 2007).

Classical models that use only variables such as DBH and H, based

* Corresponding author.
E-mail addresses: jefersonsb09@hotmail.com (J.P. Martins Silva), mayralmsilva@gmail.com (M.L. Marques da Silva), evandroforestal@gmail.com (E. Ferreira da Silva), fernandes5012@gmail.com (G. Fernandes da Silva), ribeiroflorestal@yahoo.com.br (A. Ribeiro de Mendonça), christian.cabacinha@gmail.com (C.D. Cabacinha), emanuelfa.bj@hotmail.com (E.F. Araújo), jeangelis@gmail.com (J.S. Santos), cvgiovanni@gmail.com (G.C. Vieira), narunafelix@gmail.com (M.N. Felix de Almeida), marcia.rmfe@gmail.com (M.R.d.M. Fernandes).

https://doi.org/10.1016/j.jenvman.2019.109368
Received 17 February 2019; Received in revised form 4 August 2019; Accepted 4 August 2019
Available online 14 August 2019

on dendrometric prototypes, may not perform well in parameter adjustment because there is a low correlation between DBH and H with the volume or biomass of Cerrado trees, due to the tortuosity of the bole. Thus, a factor that can positively influence the accuracy of volume and biomass estimates in the Cerrado is the use of techniques that allow the insertion of other independent variables correlated with the variable of interest.

Thereby, in order to minimize errors in volume and biomass estimation within complex vegetation, such as the Brazilian savanna, the application of artificial intelligence techniques, which are capable of mapping input patterns and discovering hidden patterns in heterogeneous non-linear high-dimensional data (Aertsen et al., 2010; Breiman, 2001; Cano et al., 2017; Görgens et al., 2015; Özçelik et al., 2013, 2010; Reis et al., 2018; Siminski, 2017; Simões and Shaw, 2007), and mixed regression models are alternatives with strong potential, representing robust methods of data analysis that can increase the accuracy of estimates (Xu et al., 2014).

Based on the above, it is believed that the classical models of regression in the literature are not the most suited to estimating the volume and biomass of Cerrado trees. Thus, the objective of this work was to evaluate the potential of artificial intelligence techniques and mixed models to increase the accuracy of the estimation of the volume and biomass of individual trees in Cerrado vegetation.

2. Materials and methods

2.1. Physical aspects of the study area

This research was conducted at 16° 41 ′south latitude and 43° 50′ west longitude in a legal reserve area (29.6 ha) of the Agricultural Sciences Institute of the Federal University of Minas Gerais in the municipality of Montes Claros (Fig. 1).

The climate of the region is Aw according to Köppen and Geiger (1928), tropical semiarid, warm, and dry, with periods of concentrated rainfall between October and March. The annual precipitation is 1060 mm and the mean annual temperature is 24.2 °C (Instituto Nacional de Meteorologia, 2013). The predominant vegetation is the Cerrado *sensu stricto* (Scolforo et al., 2008).

The data of the present work are part of a project whose main objective was to carry out the quantitative and qualitative inventory of the forest, and later to carry out the total removal of the vegetation, in order to accompany the dynamics of forest growth. The study area comprises 1 ha (100 × 100 m) within the area of 29, 6 ha (Fig. 1). Within the study area, all woody trees with DBH equal to or greater than 3.0 cm were selected. All of the trees were botanically identified and labeled with their scientific names. The classification system adopted for the family level was the APG III (2009). The total height of all individuals was also measured using a telescopic measuring rod of 10 m in length.

A total of 919 individuals were measured, covering 19 families, 45 genera, and 48 species. The Fabaceae family had the largest number of species (23.72% of the total). The families that most stood out in terms of the number of individuals were Fabaceae (218), Malpighiaceae (184), and Anarcadiaceae (183), representing 63.65% of the total sampled individuals.

The methodological procedure adopted is presented in Fig. 2 and is divided into four steps: database; calculation of BD, volume, and biomass; methodology processing; and analyzing solution strategies.

2.2. Processing inventory data

The study area, defined as 1 ha, contained 919 individuals. Of these, 504 trees belonging to 17 species were sampled (Table 1). The selected trees were cut and measured respecting two criteria: to cut a minimum of 8 trees per species with DBH ≥3 cm and not to cut trees of species protected by current legislation. The value of importance [VI] was calculated that characterizes the importance of each species in the community. Theoretically, the species with the highest VI values are the most successful in exploiting the habitat resource. The sum of the VI (%) of the 17 species represented 61.70% of the total. Measured trees had DBH (with bark) greater than 3 cm. The commercial height (HC), the total height (H) and the volume of the tree branches were measured up to the minimum diameter of 3 cm.

The first stage of the processing was to structure the database to estimate volume and biomass. This was divided in 70% for the training/adjustment and 30% for validation of the analyzed techniques. In this proportion of 70/30 trees were selected randomly within five classes of DAP, and the number of trees of each species.

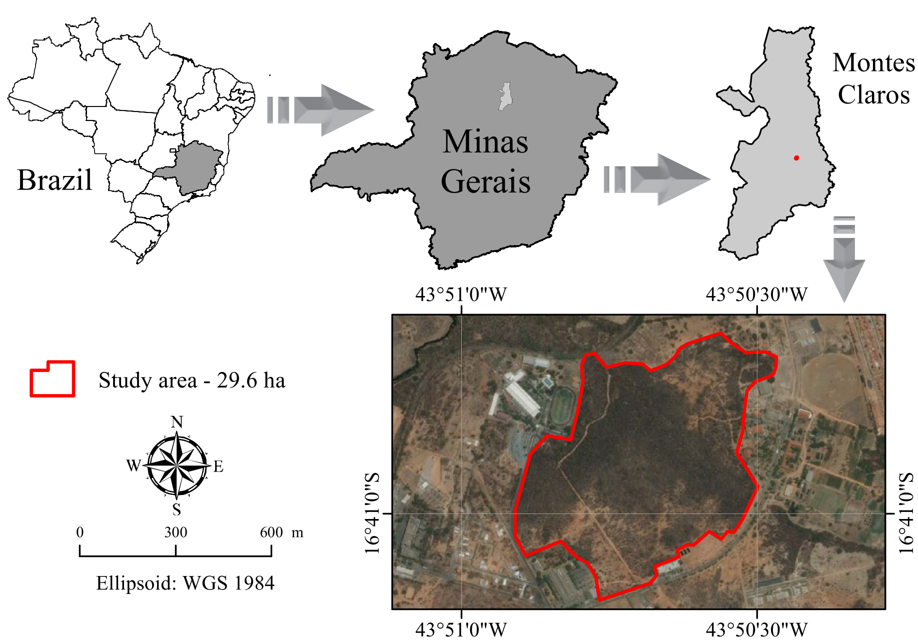

Fig. 1. Location of the study area.

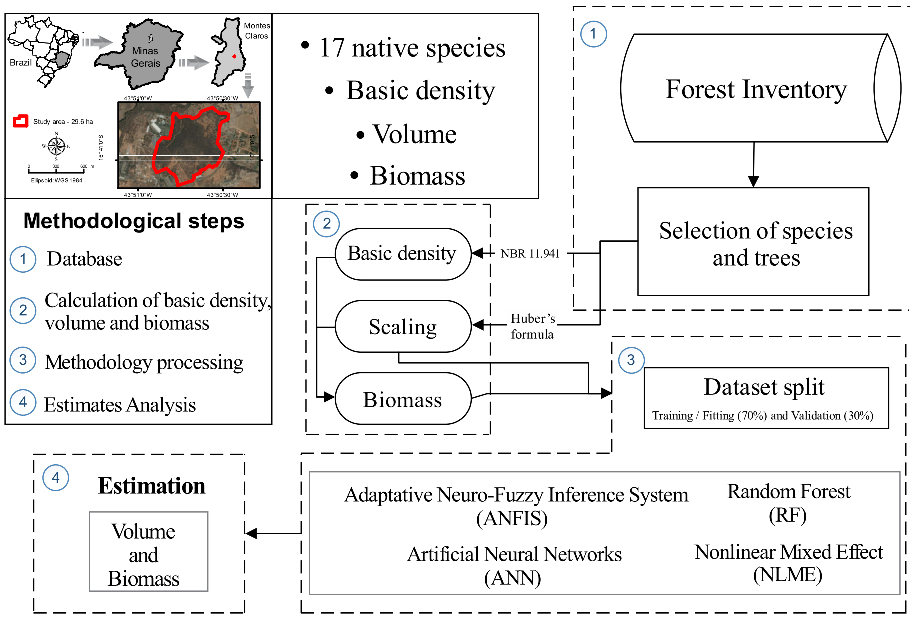

Fig. 2. Flowchart of the methodology used to evaluate volume and biomass estimation of Brazilian savannah trees.

2.3. Calculation of volume and biomass

The total volumes of the trees (Table 1) were obtained by summing the bole and branch volumes using scaled diameter measurements, applying the Huber method. (Kershaw et al., 2016). The Huber method was chosen because of its greater practical ease compared to Newton's method, and for being more accurate than the Smalian method, especially for trees that present tortuosity (Machado and Figueiredo Filho, 2009). Due to the species tortuosity of the Brazilian savanna, the measurement procedure was adapted. Measurement heights were not constant for all trees; the measurers sought to respect the straight segments of the trunks. Thus, at the moment of the cubing, the measurer identified, tree to tree, each segment rectilinearly, measuring its length and diameter in half its length, then applied it to each segment of Huber's formula. This same reasoning was used for both the main bole and the branches, and the total volume of the tree was obtained by adding the partial volumes.

To obtain the wood BD, the trees were cut and 5 wood discs of approximately 3 cm thick were removed at the 0%, 25%, 50%, 75%, and 100% HC positions. In the laboratory, the volume of the sample to calculate the density was obtained based on the principle of Archimedes, by means of water displacement, and the dry weight of each sample was obtained after oven drying at 103 °C (American Society for Testing and Materials, 2002). The wood BD at each sampled position was obtained using the ratio of the dry weight to the volume of the wood. Finally, the BD of the wood of each tree was obtained using the arithmetic mean of the densities at each longitudinal position.

To obtain the dry biomass, the average BD per species was multiplied by the total volume of the tree, considering the bole and branches, according to Brown and Lugo (1984) and Nogueira et al. (2008).

2.4. Application of volume and biomass estimation methods

Using the obtained data, four possible ways of estimating the

Table 1
Species; total trees per species; and minimum, medium, and maximum values of the variables DBH (with bark), total height, commercial height, and importance value.

Species	Number of trees	*DBH* (cm)			*H* (m)			*HC* (m)			*IV (%)*
		Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	
Luehea paniculata	8	3.50	5.02	10.50	3.35	4.75	6.95	1.55	2.95	5.72	1.17
Sclerolobium sp.	10	3.60	5.76	7.96	2.95	4.05	5.21	1.90	2.99	3.80	1.79
Terminalia fagifolia	30	3.34	5.13	9.33	3.30	4.95	13.49	1.79	3.19	5.40	4.31
Copaifera langsdorffii	31	3.02	8.42	17.83	3.23	6.29	9.90	1.80	4.59	8.02	6.84
Maytenus ilicifolia	8	3.02	5.69	8.21	3.75	4.95	5.90	1.77	3.20	4.44	0.74
Heteropteris byrsonimifolia	149	3.02	4.77	12.89	2.64	4.38	13.50	1.43	2.71	6.50	11.59
Tocoyena formosa	19	3.34	4.34	6.40	2.70	4.23	6.50	1.45	2.74	5.00	2.58
Machaerium opacum	86	3.34	7.14	16.23	1.94	4.26	9.70	1.53	3.36	8.60	10.52
Curatella americana	31	3.85	7.11	15.60	1.96	3.91	6.25	1.43	3.16	5.70	5.03
Alibertia edulis	12	3.18	4.20	7.48	2.95	4.27	6.65	1.32	2.42	5.09	1.47
Byrsonima heterophyla	16	3.02	5.96	10.38	3.22	4.56	6.20	1.39	3.04	5.20	1.85
Sebastiana brasiliensis	20	3.01	4.09	5.79	3.88	5.43	7.40	1.45	3.13	5.50	2.07
Combretum leprosum	26	3.06	5.20	14.48	3.00	5.12	8.81	1.40	3.19	6.93	2.15
Jacaranda braziliensis	21	3.02	6.07	10.70	3.11	5.26	7.34	1.43	3.36	5.72	2.54
Magonia pubescens	13	3.02	5.43	9.87	3.30	4.50	6.06	1.30	2.98	4.35	2.63
Acosmium dasycarpum	15	3.02	5.78	10.35	2.63	4.04	5.58	1.42	2.99	4.51	1.86
Plathymenia reticulata	9	4.90	12.10	18.27	4.82	7.10	9.60	3.40	5.87	8.60	2.56

Wherein *Min.* = minimum values, *Mean* = mean values, Max. = maximum values, *DBH* = diameter at a height of 1.30 m from the ground, *H* = total height, *CH* = commercial height, and *IV* = importance value.

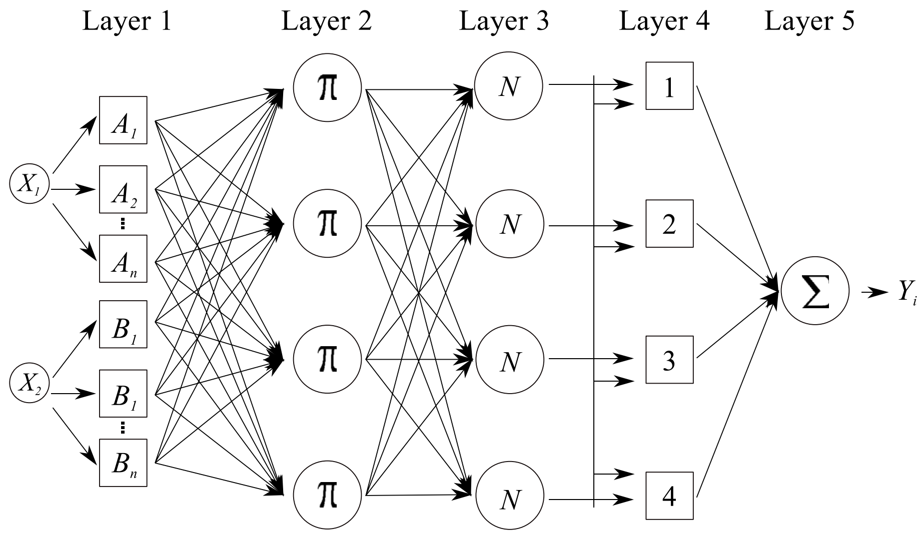

Fig. 3. Structure of a neuro-fuzzy system.
Source: Adapted from Haznedar and Kalinli (2018).

volume and biomass of individual trees were verified and described in the sequence.

2.4.1. Data processing with Adaptative network-based fuzzy inference systems (ANFIS)

The ANFIS (Fig. 3) is the combination of RNA and fuzzy logic, in which the neural network generates rules or functions of pertinence for the fuzzy systems (Simões and Shaw, 2007).

The output of the ANFIS system is modified according to the antecedent and consequent parameters. In the present work, a hybrid approach was considered, in which the parameters of the antecedents were adjusted using the descending gradient method and the consequent ones using the least squares method (Akbarzadeh et al., 2014; Mathur et al., 2016).

In layer 1 (Fig. 3), the fuzzification process occurs, in which the universe of real numbers is mapped to the fuzzy domain (Mathur et al., 2016). Each node of this layer creates an association degree (μ_i) for the input variables (x_1 e x_2), which varies from 0 to 1, using pertinence functions (A_j, B_k). The present study used the Gaussian-type function; see Eq. (1) and Eq. (2) (Bilgehan, 2011).

$$O_{1,j} = \mu_{A\,j}(x_1) = \exp\left[-\left(\frac{x_1 - c_j}{a_j}\right)\right] \tag{1}$$

$$O_{1,k} = \mu_{Bk}(x_2) = \exp\left[-\left(\frac{x_2 - c_k}{a_k}\right)\right] \tag{2}$$

wherein a_k, a_j, c_k, and c_j are parameters of the antecedents to be estimated by the descending gradient method (backpropagation);; $\mu_{An}(x_1)$ and $\mu_{Bn}(x_2)$ is the degree of association of the variables x_1 e x_2.

In layer 2, the nodes are fixed and labeled as $O_{2,\ i}$. The output of each node (ω_i) is the product of all input signals; see Eq. (3) (Akbarzadeh et al., 2014; Mathur et al., 2016).

$$O_{2,i} = \varpi_i = \mu_{Ak}(x_1)\mu_{Bk}(x_2) \tag{3}$$

In layer 3 ($O_{3,i}$), the values of ω_i are normalized by means of Eq. (4):

$$O_{3,i} = \varpi_i = \frac{\omega_i}{\omega_1 + \omega_2} \tag{4}$$

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2019.109368.

In layer 4 ($O_{4,i}$), the defuzzification occurs, which is the transformation of the fuzzy values to a real number; see Eq. (5) (Simões and Shaw, 2007):

$$O_{4,i} = \varpi_i f_i = \varpi_i(a_i x_i + b_i x_2 + c_i) \tag{5}$$

wherein the constants a_i, b_i e c_i are parameters of consequents and can be adjusted via linear regression analysis using the least squares method.

In layer 5 ($O_{5,i}$), the sum $\varpi_i f_i$ occurs, which provides the final output (Eq. (6)).

$$O_{5,i} = \sum_j \varpi_i f_i = \frac{\sum_i \omega_i f_i}{\sum_i \omega_i} \tag{6}$$

For ANFIS implementation, it is necessary to provide a data matrix containing the input (independent variables) and output (dependent variable). After loading this data, the fuzzy inference system (FIS) is generated. Two clustering algorithms were tested fuzzy C-means (FCM) (Bezdek, 1981) and subtractive cluster [(SC)] (Chiu et al., 1994) were used to optimize the amount of rules generated in the training. The clustering consists of finding data points with greater similarity in the same cluster and less similarity between different cluster data (Su and Zhao, 2017).

The FCM finds the centers of the clusters (j), minimizing the distance of x_k and v_i, using the objective function represented by Equations (7)–(9):

$$j = \sum_{k=1}^{n}\sum_{i=1}^{c} \mu_{ik}^{m}\|x_k - v_i\|^2 \tag{7}$$

$$v_i = \frac{\sum_{k=1}^{n}[\mu_i(x_k)]^2 x_k}{\sum_{k=1}^{n}[\mu_i(x_k)]^m} \tag{8}$$

$$\mu_{ik} = \frac{1}{\sum_{i=1}^{c}\left(\frac{\|x_k - v_i\|}{\|x_k - v_J\|}\right)^{1/m-1}} \tag{9}$$

Wherein n = number of observations, c = number of clusters, x_k = k-th observation, v_i = i-th center of cluster, v_j = j-th center of cluster, m = exponential constant with a value greater than 1, and μ_{ik} = degree of the association function that initially occurs randomly.

The FCM algorithm requires three parameters: the number of clusters, the exponent value, and the number of iterations. In this work, the cluster number ranging from 2 to 17 was analyzed in intervals of 1, the value of the exponent ranged from 1.10 to 3.10 in intervals of 0.10, and the number of iterations was equal to 100.

The SC algorithm considers each data point as having the potential to be the center of the cluster; this was calculated according to Equation

(10):

$$P_i \Leftarrow P_i - P_k^* e^{-\left(\frac{4}{\eta r_a}\right)} \|x_i - x_k^*\|^2 \quad (10)$$

whereinP_K^*= potential of the points that are already centers of the clusters, x_i = candidate point to be the cluster center x_k^*= centers of clusters formed, r_a = radius of influence of the cluster, η= *squash* factor.

To group the data to estimate the volume and biomass using SC, four parameters had to be specified: a) squash factor, which reduces the probability of peripheral points being considered part of a cluster. The analyzed values varied between 0.60 and 2.70, in intervals of 0.30; b) radius of influence, which is the range that the center of the cluster has under the neighboring points. Points with distances smaller than the radius range are grouped in the same cluster. The analyzed values varied between 0.20 and 0.90, in intervals of 0.10; c) acceptance rate, which considers only points with potential to be centers of the cluster. Their value varied from 0.45 to 0.60, in intervals of 0.05; and d) rejection rate. Points with values below the potential of the first cluster are rejected. Their value varied between 0.10 and 0.25, in intervals of 0.05.

Data processing with ANFIS was performed using to toolbox logic fuzzy of Software Matlab R2016a (MATHWORKS INC, 2018a). This toolbox has the ANFIS module, as proposed by Jang (1993).

The functions of association of the clusters were of the Gaussian type and the linear function was the output. The training algorithm was a hybrid that combined the backpropagation method with the least squares method. The number of training times ranged from 1 to 20. The stopping criterion for the training was an error equal to zero or early stopping. Early stopping is used to improve generalization and avoid data overfitting. The moment the training database error began to decrease and the validation database error began to increase, training stopped.

2.4.2. Data processing with an artificial neural network (ANN)

An ANN of the multilayer perceptron (MLP) type has great potential for function approximation studies (Görgens et al., 2015; Nunes and Görgens, 2016; Reis et al., 2018; Vieira et al., 2018), and therefore one was used in this study. They are composed of an input layer, where the variables are presented to the network; intermediate or hidden layers, where the processing is done; and an output layer, where the result is presented. The intermediate layer is responsible for identifying the nonlinear patterns of the data through the use of activation functions (Braga, 2007).

In this study, the hyperbolic and logistic tangent functions were tested in the intermediate layer and the linear function in the output layer. When using MLP (Fig. 4) in the estimation of the volume and biomass of Brazilian savanna species, it was necessary to adjust the synaptic weights of the connections between the processing units. These were adjusted in an iterative process commonly called learning or training (Braga, 2007). The training algorithms analyzed were the Levenberg-Marquardat and the resilient backpropagation (Haykin, 2003; Braga, 2007).

The inputs were normalized, by means of internal MATLAB procedure, in intervals of −1 to 1 and 0 to 1, respectively, for the activation function hyperbolic and logistic tangent. After normalization, the inputs were entered into the network and passed to the intermediate layer (a_j), where the weighted sum was calculated according to Eq. (11), which is the value of the variable (x_i) times the weight of the variable (w_{ij}) plus the bias of the variable (b_i). The product of this calculation was transmitted to the activation functions of the hyperbolic (C_{jh}) (Eq. (12)) and logistic (C_{jl}) (Eq. (13)) tangents. Finally, the calculations of the intermediate neurons served as input (x_1, x_2, ...x_n) the DBH, HT and species, to the output neurons (Y_K) volume and biomass (Eqs. (14) and (15)).

$$a_j = \sum_{i=1} (x_i w_{ij}) + b_i \quad (11)$$

$$C_{jh} = \frac{e^{a_j} + e^{-a_j}}{e^{a_j} - e^{-a_j}} \quad (12)$$

$$C_{jl} = \frac{1}{1 + e^{a_j}} \quad (13)$$

$$d_k = \sum_{j=1} (c_j w_{jk}) \Bigg) + b_k \quad (14)$$

$$Y_K = d_k \quad (15)$$

For the adjustment of the parameters of the ANNs, the toolbox artificial neural networks of the software Matlab R2016a (MATHWORKS INC, 2018b) were used. Ten networks were trained by altering the number of neurons in the intermediate layer, which varied from 1 to 19, seeking to investigate the application of a simpler structure to a more complex one. According to Braga (2007), it is best to use the simplest structure, as long as it provides accurate estimates. Four criteria for stopping ANN training were established. During training, the criterion that was reached first established the end of the processing. The criteria analyzed were maximum number of times equal to 1,000, maximum training time equal to 300 s, maximum error of 0.01, and early stopping.

2.4.3. Data processing with random forest

Random forest (RF) is a machine learning method that uses a bootstrap approach to construct multiple decision trees. In addition, the trees have a resource selection system in their structure, using only the variables that actually influence the response (Cano et al., 2017).

The RF training mechanism creates several regression trees using the classification and regression tree (CART) algorithm. These trees are generated from bootstrap samples. The bootstrap is a statistical approach used to quantify the uncertainty associated with a given estimator (James et al., 2013).

Data processing with RF was performed using the statistics and machine learning toolbox of Matlab R2016a. In order to estimate the volume and biomass of native Brazilian savanna species using RF (Fig. 5), it was necessary to define the number of trees, the number of variables chosen to start the division, and the number of leaves for each tree formed. The number of trees used as standard was 500 (Wang

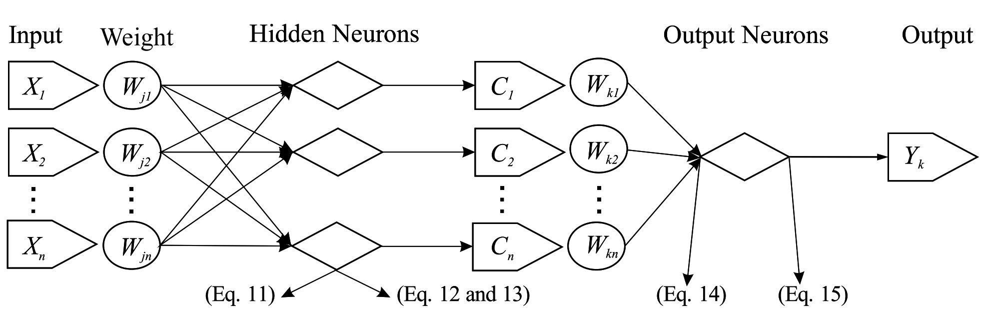

Fig. 4. Structure of an artificial neural network.

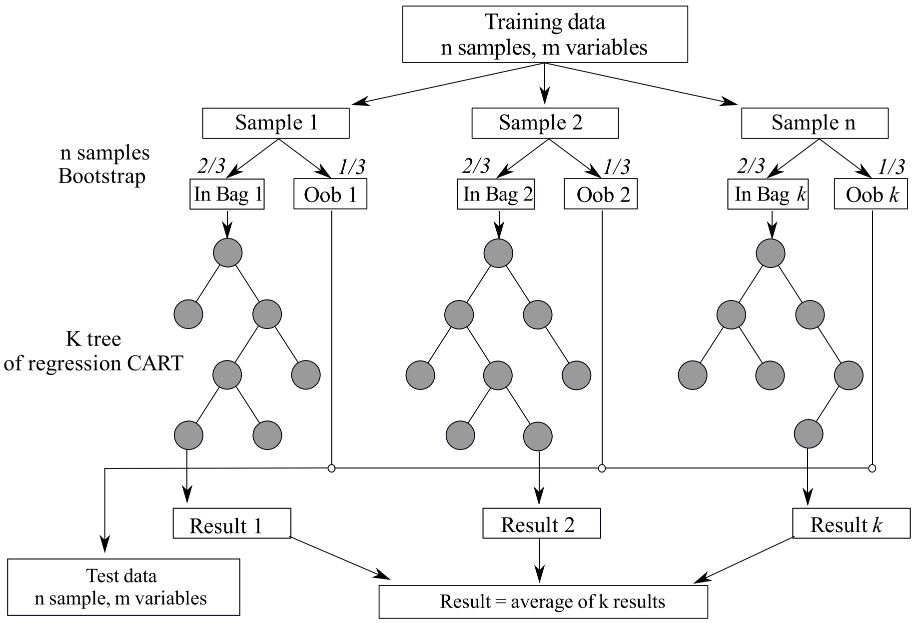

Fig. 5. Structure of a random forest.
Source: Adapted from Ibrahim and Khatib (2017).

Table 2
Statistics used to evaluate the performance of the techniques used to estimate the volume and biomass of Brazilian savannah trees.

Statistics	Formulas	Nº
Correlation coefficient	$r_{y\hat{y}} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(\hat{Y}_i - \hat{Y}_m)}{\sqrt{[\sum_{i=1}^{n}(Y_i - \bar{Y})^2]\left[\sum_{i=1}^{n}(\hat{Y}_i - \hat{Y}_m)^2\right]}}$	(17)
Bias	$B(\%) = \frac{100}{\bar{Y}} \frac{\sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \hat{Y}_i}{n}$	(18)
Relative root mean square error	$RMSE(\%) = \frac{100}{\bar{Y}} \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}}$	(19)
Mean absolute error	$MAE = \frac{\sum_{i=1}^{n}\lvert Y_i - \hat{Y}_i \rvert}{n}$	(20)
Residual prediction deviation	$RPD = \frac{SD}{RMSE}$	(21)

et al., 2016). In this work, 50 to 1000 trees were trained, with a range of 50 trees. The number of variables selected to begin dividing was 1/3 of the number of independent variables and the number of observations per sheets was 5–35, with intervals of 5 sheets. The choice of the best RF configuration was made based on the square root of the relative mean error (rRMSE) of the training and validation data set.

2.4.4. Data processing with regression models

The regression model used in the volume and biomass estimation was that by Schumacher and Hall (1933) (Eq. (16)). This model was adjusted in fixed (MNL) and mixed (NLME) form, in order to check if the inclusion of the random effect (species) would bring gains in terms of the accuracy of the estimation.

$$Y = \beta_0 DAP^{\beta_1} H^{\beta_2} + \varepsilon \quad (16)$$

All possible combinations for the inclusion of the random effect in

Table 3
Basic density in different positions and means of the 17 species analyzed.

Species	Number of trees	Position in the commercial bole					Mean	VC%
		%	5%	0%	5%	100%		
		Basic density g cm$^{3-1}$						
Luehea paniculata	8	0.55	0.58	0.57	0.57	0.58	0.57	2.15
Sclerolobium sp.	10	0.62	0.62	0.59	0.55	0.58	0.59	4.98
Terminalia fagifolia	30	0.67	0.63	0.62	0.61	0.65	0.64	3.79
Copaifera langsdorffii	31	0.65	0.63	0.60	0.59	0.59	0.61	4.38
Maytenus ilicifolia	8	0.66	0.63	0.66	0.62	0.58	0.63	5.26
Heteropteris byrsonimifolia	149	0.66	0.64	0.64	0.63	0.63	0.64	1.91
Tocoyena formosa	19	0.71	0.72	0.72	0.71	0.70	0.71	1.18
Machaerium opacum	86	0.62	0.65	0.63	0.61	0.65	0.63	2.83
Curatella americana	31	0.55	0.54	0.53	0.53	0.53	0.54	1.67
Alibertia edulis	12	0.72	0.70	0.69	0.68	0.68	0.69	2.41
Byrsonima heterophyla	16	0.58	0.53	0.53	0.56	0.45	0.53	9.34
Sebastiana brasiliensis	20	0.60	0.58	0.57	0.55	0.58	0.58	3.15
Combretum leprosum	26	0.69	0.68	0.66	0.66	0.66	0.67	2.11
Jacaranda braziliensis	21	0.60	0.57	0.57	0.58	0.52	0.57	5.19
Magonia pubescens	13	0.72	0.69	0.69	0.70	0.80	0.72	6.44
Acosmium dasycarpum	15	0.45	0.42	0.42	0.48	0.41	0.43	6.61
Plathymenia reticulata	9	0.71	0.66	0.63	0.63	0.72	0.67	6.42

Wherein VC (%) = variation coefficient as a percentage.

Table 4
Configurations of selected techniques used to estimate volume and biomass.

Technical	Strategies	Variable	Parameters		
			Algorithm	Cluster number	Exponent
ANFIS	1	Volume	FCM	3	1.2
		Biomass		3	1.3
	2	Volume		9	1.5
		Biomass		10	1.2
			Algorithm	Activation function	Neuron number
ANN	1	Volume	Levenberg-Marquadt	Hyperbolic tangent	19
		Biomass			6
	2	Volume			5
		Biomass			4
RF	1	Volume	Algorithm	Number of trees	Number of obs/tree
			CART	50	5
		Biomass		950	5
	2	Volume		50	5
		Biomass		450	5

Wherein FCM = fuzzy C-means, CART = classification and regression tree.

the NLME were tested: i.e., effects associated with the DBH and H constant coefficients and their combinations. To adjust the model, the maximum likelihood algorithm was used in the R software (R Core Team, 2017).

2.5. Solution strategies

Volume and biomass were estimated using the ANFIS, ANN, RF, and regression models with fixed and mixed effects. The training/fitting strategy had as independent variables DBH, H, and species, and as the dependent variables volume and biomass. Thus, two training/fitting strategies were adopted, varying the inclusion of the independent species variable, in which Strategy 1 for the volume and biomass estimation used DBP, H, and species, and Strategy 2 only DBP and H.

2.6. Methods for assessing the accuracy of estimates

The ANFIS, RNA, RF, and mixed effects regression models for training/adjustment and validation were evaluated based on the following statistics: the linear correlation coefficient between the observed and estimated values ($r_{Y\hat{Y}}$), relative bias [B (%)], RMSE(%), mean absolute error (MAE), and residual prediction deviation (RPD) (Table 2).Wherein Y_i= dependent variable observed, $\hat{Y}_i$= dependent variable estimated, $\overline{Y}_i$= mean of the observed dependent variable, $\hat{Y}_m$= mean of the estimated dependent variable, n= number of observations, SD = standard deviation of measured values, and RMSE = root mean square error.

In order to complement the statistics, graphs were drawn up that related the values observed and estimated by the techniques, and graphs of percent residuals were made. The error of each observation was calculated as a percentage (Eq. (22)).

$$E(\%) = \frac{Y_i - \hat{Y}_i}{Y_i}100 \tag{22}$$

wherein $E(\%)$ = error of each observation.

3. Results

3.1. Values of the basic wood density of the species studied

Table 3 shows the mean values of the densities per sampled position, the mean value per species, and the variation coefficient (VC%) of the means at the collection positions. The VC% of the means was low in relation to the 17 species that were collected, and the *Byrsonima heterophyla* species presented the highest VC value (9.34%). The species *Acosmium dasycarpum* had the lowest BD (0.43 g cm-3) and *Magonia pubescens* the highest BD (0.72 g cm-3). The overall mean densities were 0.61 g cm-3 and the mean variation coefficient was 4.11%.

The mean BD of the species of the Brazilian savanna fragment studied was 0.612 g cm^{-3}, similar to the wood BD of Cerrado tree species in the state of Tocantins (Silva and Miguel, 2015), with a value of 0.650 g cm^{-3}. Jati et al. (2014) also found results close to the present study with a mean BD of 0.592 g.cm-3 in eight species of the Savana de Roraima in the northern Brazilian Amazon. However, Goulart et al. (2012) found that the species *Stryphnodendron adstringens* (barbatimão) had an average BD of 0.469 g.cm-3. For all species, the variation in density along the bole did not exceed 10%, varying in average at around 4.10%, which can be considered a low variation along the bole.

Table 5
Training and validation statistics for ANFIS, ANN, RF, NLME, and MNL volume estimation in Strategies 1 and 2.

Group	Technique	$r_{Y\hat{Y}}$	rRMSE (%)	B (%)	MAE	RPD
Strategy 1						
		Vol./Bio.	Vol./Bio.	Vol./Bio.	Vol./Bio.	Vol./Bio.
Training/Fitting	ANFIS	0.984/0.985	13.22/13.05	0.00/0.00	0.0004/0.2447	5.68/5.80
	ANN	0.986/0.986	12.33/12.47	−0.06/0.17	0.0004/0.2410	6.09/6.07
	RF	0.979/0.978	16.82/17.86	0.39/0.16	0.0004/0.3018	4.46/4.24
	NLME	0.984/0.985	13.55/13.15	−0.03/-0.10	0.0003/0.1890	5.54/5.75
Validation	ANFIS	0.981/0.982	13.16/12.99	−1.24/-1.10	0.0001/0.2303	5.10/5.22
	ANN	0.981/0.982	13.02/12.88	−1.34/-0.89	0.0004/0.2303	5.16/5.27
	RF	0.977/0.971	14.87/16.92	−3.56/-3.20	0.0005/0.3332	4.52/4.00
	NLME	0.981/0.981	12.95/13.28	1.17/0.77	0.0004/0.2846	5.18/5.11
Strategy 2						
		Vol./Bio.	**Vol./Bio.**	**Vol./Bio.**	**Vol./Bio.**	**Vol./Bio.**
Training/Fitting	ANFIS	0.981/0.979	14.55/15.35	0.00/0.00	0.0004/0.2788	5.14/4.90
	ANN	0.981/0.979	14.86/15.17	0.00/-0.04	0.0004/0.2726	5.03/4.95
	RF	0.983/0.982	13.94/14.83	0.11/0.31	0.0003/0.2384	5.37/5.07
	MNL	0.979/0.974	15.07/17.00	0.17/-0.30	0.0004/0.2564	4.96/4.42
Validation	ANFIS	0.974/0.969	16.12/17.95	−0.10/-0.21	0.0004/0.3187	4.33/3.96
	ANN	0.975/0.970	15.64/17.34	0.00/0.01	0.0004/0.3123	4.46/4.09
	RF	0.968/0.957	17.63/20.49	−0.58/0.01	0.0005/0.3625	3.96/3.47
	MNL	0.972/0.969	16.94/16.63	0.57/-0.10	0.0005/0.2752	4.12/4.27

Wherein r = correlation coefficient, rRMSE (%) = square root correlation coefficient of the mean error in percentage, B (%) = Bias, MAE = mean absolute error, RPD = residual prediction deviation, Vol./Bio. = values obtained from volume and biomass estimation.

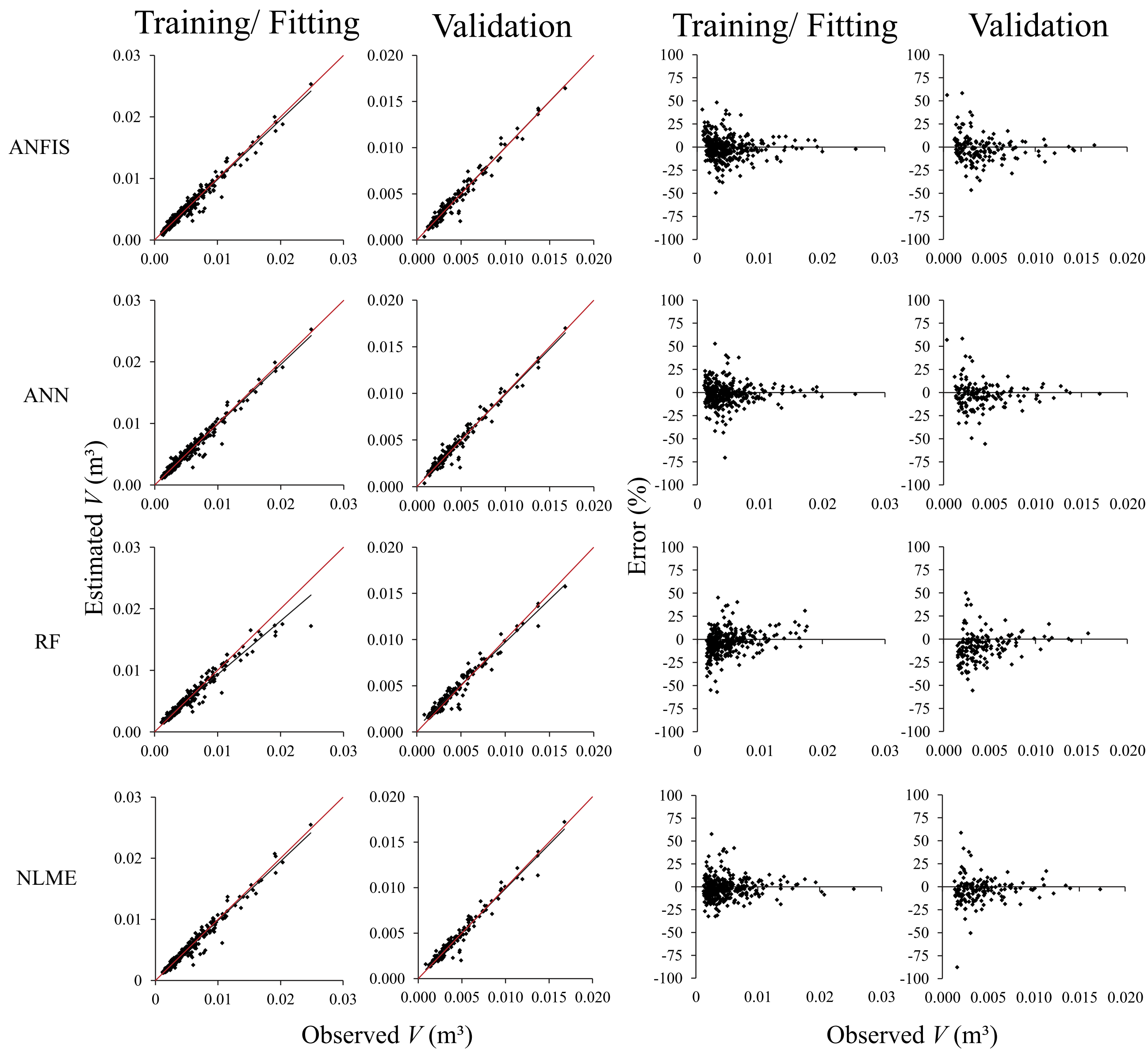

Fig. 6. Graphical analysis of the volume (V) estimated versus the volume (V) observed and graph of the residual, in percentage, versus values observed for Strategy 1.

In absolute values, the mean difference between the highest and lowest densities was 0.059 g.cm-3, with an emphasis on the *Tocoyena formosa* and *Byrsonima heterophyla* species, which had the lowest and highest variations, respectively (Table 3).

3.2. Strategies 1 and 2 results for the estimation of the volume of trees in the Brazilian savanna

Parameter variations resulted in a total of 1,402, 760, and 700 combinations for the ANFIS, ANN, and RF techniques, respectively. The best configuration of each technique was selected based on training and validation errors. The parameters obtained are presented in Table 4.

In NLME, all possible combinations for the inclusion of the random effect in the model—i.e., effects associated with the DBH and H constant coefficients and their combinations—were tested. The model with a random effect in the H variable presented the best performance in terms of accuracy in estimating both volume and biomass. For the Schumacher and Hall nonlinear (MNL) models, all coefficients were significant ($p < 0.05$). The values of the fixed and random effects coefficients and of model nonlinear Schumacher and Hall model are given in Appendix A of Supplementary data.

In Strategies 1 and 2, the statistics used to evaluate the accuracy of the methods studied, rRMSE (%) and V (%), did not show great variation between the methodologies used to estimate volume and biomass (Table 5).

Among the techniques studied, NLME, ANN, and ANFIS were more accurate in estimating volume and biomass, in the training and validation sets, in the two strategies when compared to RF. The inclusion of the species variable (Strategy 1) prompted an improvement in the performance of the techniques studied when analyzing the statistics $r_{Y\hat{Y}}$, rRMSE (%), and MAE (Table 5).

Regarding B statistics (%), the ANN showed a better performance in Strategies 1 and 2, but the NLME obtained a better result for the validation data set (1.17%) in Strategy 1. RPD was greater than 2.5 in all techniques, regardless of the strategy used, indicating excellent predictions, according to Viscarra Rossel et al. (2006).

When analyzing the graphs of the relationship between the observed and estimated values of volume and biomass by the techniques in Strategies 1 and 2 (Figs. 6–9), it was observed that the ANFIS, ANN, and NLME techniques presented points closer to the 45° line (red line) when

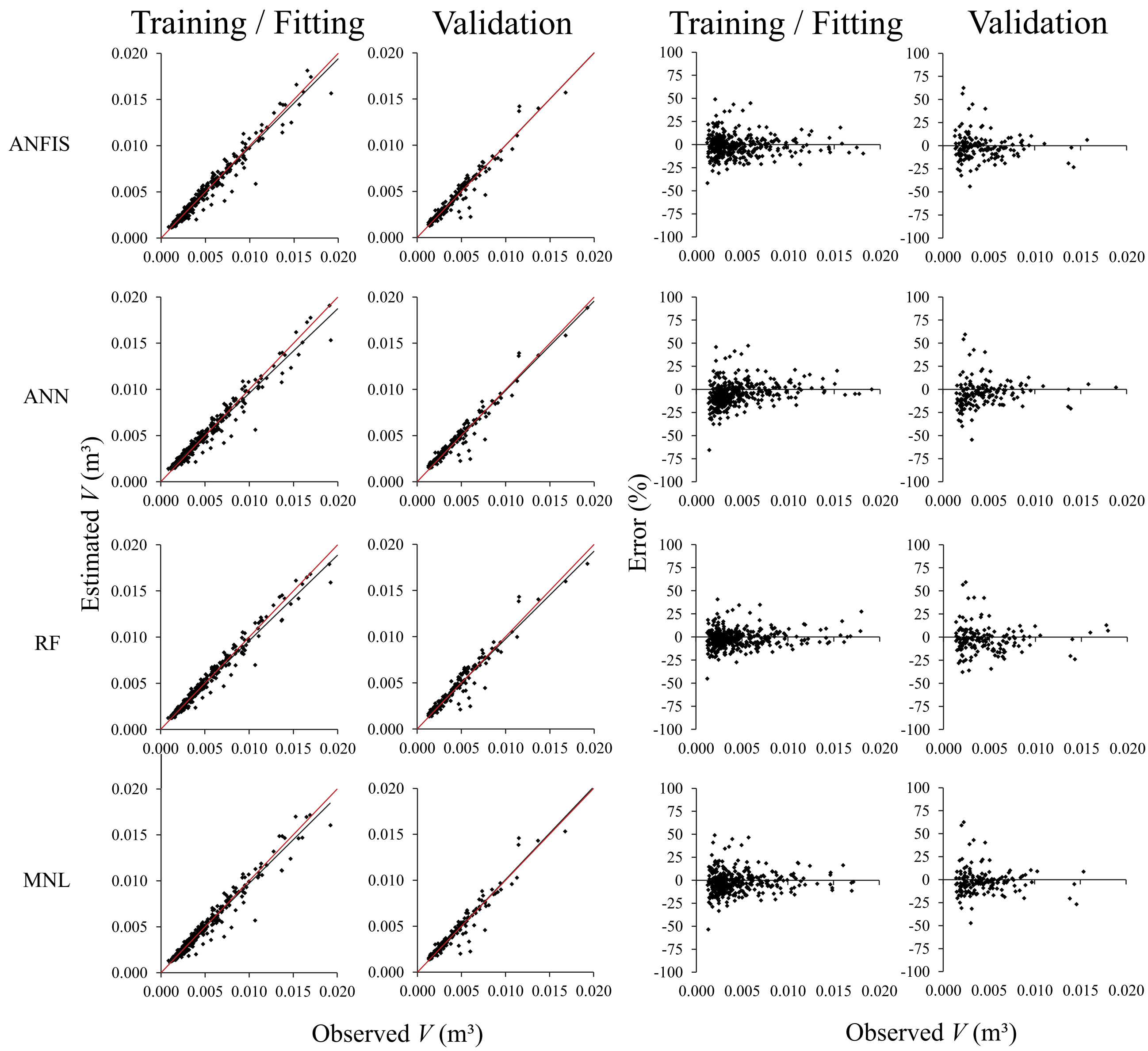

Fig. 7. Graphical analysis of the volume (V) estimated versus the volume (V) observed graph of the residuals in percentage versus values observed for Strategy 2.

compared to RF.

There was no trend of overestimating or underestimating the volume and biomass for the set of training and validation data of the ANFIS, ANN, and NLME techniques. The residual graphs presented, on average, 94.02% of the error contained in the range of ± 20%. RF showed inferior performance to the other techniques, with a tendency to underestimate the larger volumes and biomasses of trees; in addition, approximately 87.29% of the error was concentrated in the range of ± 25%.

4. Discussion

The artificial intelligence techniques employed in this study presented good statistical indicators in the estimation of the volume and biomass of Brazilian savanna trees (Table 5). This result can be attributed to the ability of artificial intelligence techniques to capture the nonlinearity present in the data, as they can approximate complex functions. According to Vieira et al. (2018), this is an important characteristic for the modeling of forest biological problems, since they do not usually present linear behavior.

When analyzing the statistical indicators ($r_{Y\hat{Y}}$, rRMSE (%), B (%) MAE and RPD) (Table 5) was possible to observe that the ANFIS and ANN techniques were superior to the RF technique. This result may have been influenced by the learning method. The ANFIS and ANN techniques are trained by error-back-propagation algorithms, which makes them able to perform input-output mapping for problems of any kind (Haykin, 2003).

In the mechanism of RF operation, the response is given by the mean of the observed values. In this case, it may lead to overestimations in the lower values and underestimations in the larger ones (Nunes and Görgens, 2016). This can be observed in the graphical analyses of the estimates (Figs. 6–9). This trend was also observed in the works of Nunes and Görgens (2016) and Zhang and Lu (2012).

The mixed models provided individual estimates of volume and biomass for each species and allowed for the introduction of variance and covariance structures into the random variable (species). These advantages make mixed models a good alternative in estimating the volume and biomass of woody vegetation with complex growth, such as in the Cerrado.

The use of Schumacher and Hall (1933) model in the estimation of

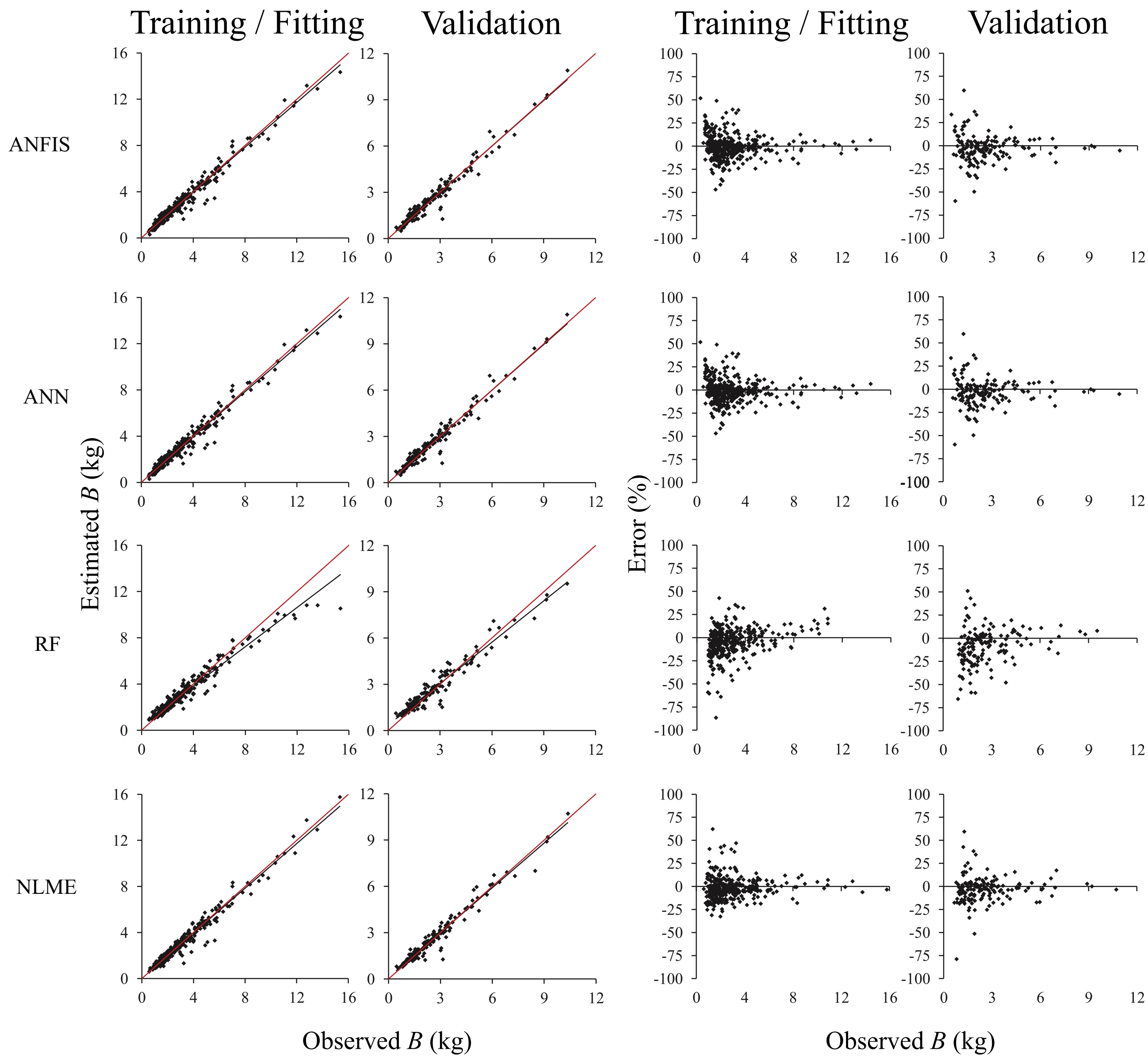

Fig. 8. Graphical analysis of the biomass (B) estimated versus biomass (B) observed and graph of errors (%) of Strategy 1.

volume and biomass for Brazilian savanna species resulted in accurate estimates. The use of species as a random effect provided a reduction in rRMSE of 8.90% for volume and 20.00% for biomass. Research like that of Gouveia et al. (2015) obtained superior performance in the NLME compared to the traditional model of Schumacher and Hall (1933), with a reduction in the residual standard error of 0.0591 to 0.0023.

In the Brazilian savanna, sampling of volume through scaling measurements is an exhausting activity, due to the typology of the vegetation and climate of the region. In addition, identifying species further burdens this activity. However, considering the good results of Strategy 2, which does not consider species as an independent variable, it is possible to accurately estimate the volume of Brazilian savanna trees even without considering the species sampled. However, the gain in terms of accuracy in estimating biomass with the use of the species was considerable.

In many studies, the superiority of AI techniques is reported to provide more accurate estimates than classical regression models (Nunes and Görgens, 2016). It can be seen from the results of the present study that more accurate regression techniques, such as mixed models, provide comparable or even better estimates than AI techniques, such as the values of rRMSE (%) and bias (%) showed in the validation group data (Table 5).

The artificial intelligence techniques and mixed regression models evaluated in this study presented superior results than the results found in other studies. Ribeiro et al. (2011) found a coefficient of determination of 89.80% for the best allometric regression model that included DBH and density as independent variables. Rezende et al. (2006) obtained a coefficient of determination of 98.01% and a standard error of 25.00% for volume estimation and a determination coefficient of 98.64% and standard error of 25.66% for biomass estimation using models of regression.

5. Conclusion

Because the modeling of forest resources commonly presents complex relationships among the variables, artificial intelligence techniques such as ANFIS, ANN, RF, and NLME may be good alternative modeling techniques. These techniques were able to estimate the volume and biomass of different species in the Brazilian savanna. Although the superiority of the ANFIS, ANN, and NLME techniques over RF was

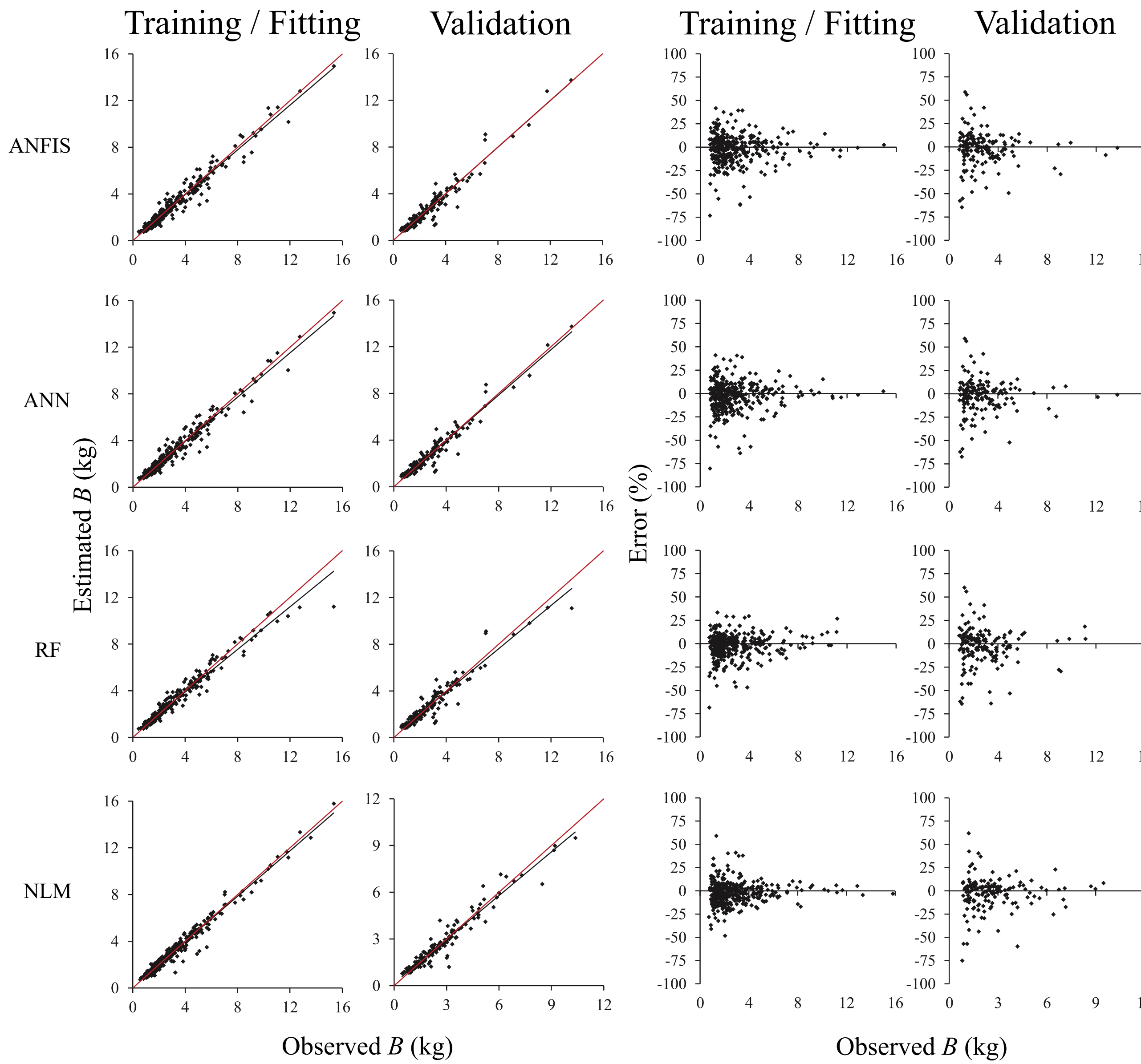

Fig. 9. Graphical analysis of the estimated biomass (B) versus biomass (B) observed and graph of errors (%) of Strategy 2.

observed, the four techniques have great potential for use as auxiliary tools in forest measurement and management; they can estimate variables that contribute to the knowledge and planning of the use of forest resources. In addition, the proposed methodologies can be adapted to other biomes, forest typologies, and variables of interest.

Acknowledgments

The authors express their gratitude towards the Foundation for the Support of Research and Innovation of Espírito Santo (FAPES) and the Coordination for the Improvement of Higher Education Personnel (CAPES) for awarding the grants required for the present study and the National Council for Scientific and Technological Development (CNPq) by the funding of research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2019.109368.

References

Aertsen, W., Kint, V., van Orshoven, J., Özkan, K., Muys, B., 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. Ecol. Model. 221, 1119–1130. https://doi.org/10.1016/j.ecolmodel.2010.01.007.

APG III, 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot. J. Linn. Soc. 161, 105–121. https://doi.org/10.1111/j.1095-8339.2009.00996.x.

Akbarzadeh, S., Arof, A.K., Ramesh, S., Khanmirzaei, M.H., Nor, R.M., 2014. Prediction of conductivity by adaptive neuro-fuzzy model. PLoS One 9, e92241. https://doi.org/10.1371/journal.pone.0092241.

American society for testing and materials, 2002. Standard Test Methods for Specific Gravity of Wood and Wood-Based Materials. West Conshohocken, Pensilvânia.

Barni, P.E., Manzi, A.O., Condé, T.M., Barbosa, R.I., Fearnside, P.M., 2016. Spatial distribution of forest biomass in Brazil's state of Roraima, northern Amazonia. For. Ecol. Manage. 377, 170–181. https://doi.org/10.1016/j.foreco.2016.07.010.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Springer, Boston, MA.

Bilgehan, M., 2011. Comparison of ANFIS and NN models—with a study in critical buckling load estimation. Appl. Soft Comput. 11, 3779–3791. https://doi.org/10.1016/j.asoc.2011.02.011.

Braga, A. de P., 2007. Redes Neurais Artificiais. Teoria e Aplicações. Editora LTC, Rio de Janeiro.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Brown, S., Lugo, A.E., 1984. Biomass of tropical forests: a new estimate based on forest volumes. Sci. New Ser. 223, 1290–1293.
Cano, G., Garcia-Rodriguez, J., Garcia-Garcia, A., Perez-Sanchez, H., Benediktsson, J.A., Thapa, A., Barr, A., 2017. Automatic selection of molecular descriptors using random forest: application to drug discovery. Expert Syst. Appl. 72, 151–159. https://doi.org/10.1016/j.eswa.2016.12.008.
Chiu, S.L., Center, R.S., Oaks, T., 1994. Fuzzy model identification based on cluster estimation. J. Intell. Fuzzy Syst. Appl. Eng. Technol. 2, 267–278.
Cunha, C.S., Lopes, N.L., Veloso, C.M., Jacovine, L.A.G., Tomich, T.R., Pereira, L.G.R., Marcondes, M.I., 2016. Greenhouse gases inventory and carbon balance of two dairy systems obtained from two methane-estimation methods. Sci. Total Environ. 571, 744–754. https://doi.org/10.1016/j.scitotenv.2016.07.046.
Ferraz, A.S., Soares, V.P., Soares, C.P.B., Ribeiro, C.A.A.S., Binoti, D.H.B., Leite, H.G., 2014. Estimativa do estoque de biomassa em um fragmento florestal usando imagens orbitais. Floresta e Ambient 21, 286–296. https://doi.org/10.1590/2179-8087.052213.
Görgens, E.B., Montaghi, A., Rodriguez, L.C.E., 2015. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. Comput. Electron. Agric. 116, 221–227. https://doi.org/10.1016/j.compag.2015.07.004.
Goulart, S.L., Mori, F.A., Ribeiro, A.de O., Couto, A.M., Arantes, M.D.C., Mendes, L.M., 2012. Análises químicas e densidade básica da madeira de raiz, fuste e galho de barbatimão [(Stryphnodendron adstringens) Coville] de bioma cerrado. CERNE 18, 59–66. Available from: https://doi.org/10.1590/S0104-77602012000100008.
Gouveia, J.F., Da Silva, J.A.A., Ferreira, R.L.C., Gadelha, F.H.L., Lima Filho, L.M. de A., 2015. Modelos volumétricos mistos em clones de eucalyptus no polo gesseiro do araripe, pernambuco. floresta 45, 587. https://doi.org/10.5380/rf.v45i3.36844.
Günther, A., Böther, S., Couwenberg, J., Hüttel, S., Jurasinski, G., 2018. Profitability of direct greenhouse gas measurements in carbon credit schemes of peatland rewetting. Ecol. Econ. 146, 766–771. https://doi.org/10.1016/j.ecolecon.2017.12.025.
Haykin, S., 2003. Redes neurais: princípios e prática, second ed. Bookman Editora, Porto Alegre.
Haznedar, B., Kalinli, A., 2018. Training ANFIS structure using simulated annealing algorithm for dynamic systems identification. Neurocomputing 302, 66–74. https://doi.org/10.1016/j.neucom.2018.04.006.
Ibrahim, I.A., Khatib, T., 2017. A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm. Energy Convers. Manag. 138, 413–425. https://doi.org/10.1016/j.enconman.2017.02.006.
James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning, Design. Springer Texts in Statistics. Springer, New York, New York, NY. https://doi.org/10.1007/978-1-4614-7138-7.
Jang, J.S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man. Cybern 23 (3), 665–685. Available from: https://doi.org/10.1109/21.256541.
Jati, S.R., Fearnside, P.M., Barbosa, R.I., 2014. Densidade da madeira de árvores em savanas do norte da Amazônia brasileira. Acta Amaz 44, 79–86. Available from: https://doi.org/10.1590/S0044-59672014000100008.
Kershaw, J.A., Ducey, M.J., Beers, T.W., Husch, B., 2016. Forest Mensuration. John Wiley & Sons, Ltd, Chichester, UK. https://doi.org/10.1002/9781118902028.
Köppen, W., Geiger, R., 1928. Klimate der Erde. Verlag Justus Perthes, Gotha Wall-map 150cm × 200cm.
Koala, J., Sawadogo, L., Savadogo, P., Aynekulu, E., 2017. Allometric Equations for Below-Ground Biomass of Four Key Woody Species in West African Savanna-Woodlands, vol. 51. pp. 1–15.
Kuch, D., 2017. “Fixing” climate change through carbon capture and storage: situating industrial risk cultures. Futures 92, 90–99. https://doi.org/10.1016/j.futures.2017.02.001.
Ledo, A., Heathcote, R., Hastings, A., Smith, P., Hillier, J., 2018. Perennial-GHG: a new generic allometric model to estimate biomass accumulation and greenhouse gas emissions in perennial food and bioenergy crops. Environ. Model. Softw 102, 292–305. https://doi.org/10.1016/j.envsoft.2017.12.005.
Machado, S. do A., Figueiredo Filho, A., 2009. Dendrometria, second ed. Unicentro, Guarapuava.
Marris, E., 2005. The forgotten ecosystem. Nature 437, 944–945. https://doi.org/10.1038/437944a.
Mathur, N., Glesk, I., Buis, A., 2016. Comparison of adaptive neuro-fuzzy inference system (ANFIS) and Gaussian processes for machine learning (GPML) algorithms for the prediction of skin temperature in lower limb prostheses. Med. Eng. Phys. 38, 1083–1089. https://doi.org/10.1016/j.medengphy.2016.07.003.
MATHWORKS INC, I., 2018a. Logic Fuzzy Toolbox: for Use with MATLAB®: User's Guide.
MATHWORKS INC, I., 2018b. Neural Network Toolbox: for Use with MATLAB®: User's Guide.
Moon, B.E., Choi, E.G., Kim, C.H., Kim, J.K., Ryou, Y.S., Kim, H.T., 2013. An analysis of local quantity of carbon absorption, fixing and discharge by using GIS. IFAC Proc 46, 171–175. https://doi.org/10.3182/20130327-3-JP-3017.00039.
Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858. https://doi.org/10.1038/35002501.
Instituto Nacional de Meteorologia, 2013. Dados Meteorológicos. WWW Document.
Nogueira, E., Fearnside, P., Nelson, B., Barbosa, R., Keizer, E., 2008. Estimates of forest biomass in the Brazilian Amazon : new allometric equations and adjustments to biomass from wood-volume inventories. For. Ecol. Manage. 256, 1853–1867. https://doi.org/10.1016/j.foreco.2008.07.022.
Nunes, M.H., Görgens, E.B., 2016. Artificial intelligence procedures for tree taper estimation within a complex vegetation mosaic in Brazil. PLoS One 11, 1–16. https://doi.org/10.1371/journal.pone.0154738.
Özçelik, R., Diamantopoulou, M.J., Brooks, J.R., Wiant, H.V., 2010. Estimating tree bole volume using artificial neural network models for four species in Turkey. J. Environ. Manag. 91, 742–753. https://doi.org/10.1016/j.jenvman.2009.10.002.
Özçelik, R., Diamantopoulou, M.J., Crecente-Campo, F., Eler, U., 2013. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. For. Ecol. Manage. 306, 52–60. https://doi.org/10.1016/j.foreco.2013.06.009.
R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.
Ratuchne, L.C., Koehler, H.S., Watzlawick, L.F., Sanquetta, C.R., Schamne, P.A., 2016. Estado da Arte na Quantificação de Biomassa em Raízes de Formações Florestais. Floresta e Ambient 23, 450–462. https://doi.org/10.1590/2179-8087.131515.
Reis, L.P., de Souza, A.L., dos Reis, P.C.M., Mazzei, L., Soares, C.P.B., Miquelino Eleto Torres, C.M., da Silva, L.F., Ruschel, A.R., Rêgo, L.J.S., Leite, H.G., 2018. Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. Ecol. Eng. 112, 140–147. https://doi.org/10.1016/j.ecoleng.2017.12.014.
Rezende, A.V., Vale, A.T. do, Sanquetta, C.R., Figueiredo-Filho, A., Felfili, J.M., 2006. Comparison of mathematical models to volume, biomass and carbon stock estimation of the woody vegetation of a cerrado sensu stricto in Brasilia. DF. Sci. For. 71, 65–76.
Ribeiro, S.C., Fehrmann, L., Soares, C.P.B., Jacovine, L.A.G., Kleinn, C., de Oliveira Gaspar, R., 2011. Above- and belowground biomass in a Brazilian Cerrado. For. Ecol. Manage. 262, 491–499. https://doi.org/10.1016/j.foreco.2011.04.017.
Sales, M.H., Souza, C.M., Kyriakidis, P.C., Roberts, D.A., Vidal, E., 2007. Improving spatial distribution estimation of forest biomass with geostatistics: a case study for Rondônia, Brazil. Ecol. Model. 205, 221–230. https://doi.org/10.1016/j.ecolmodel.2007.02.033.
Schumacher, F.X., Hall, F. dos S., 1933. Logarithmic expression of timber-tree volume. J. Agric. Res. 47, 719–734.
Scolforo, J.R.S., Mello, J.M., de, Oliveira, A.D. de, 2008. Inventário florestal de Minas Gerais: Cerrado - Florística, Estrutura, Diversidade, Similaridade, Distribuição diâmetrica e de Altura, Volumetria, Tendências de Crescimento e Áreas Aptas para o Manejo Florestal, first ed. Lavras, Minas Gerais.
Silva, C.J. da, Vale, A.T. do, Miguel, E.P., 2015. Densidade básica da madeira de espécies arbóreas de Cerradão no estado de Tocantins. Pesqui. Florest. Bras. 35, 63. https://doi.org/10.4336/2015.pfb.35.82.822.
Siminski, K., 2017. Interval type-2 neuro-fuzzy system with implication-based inference mechanism. Expert Syst. Appl. 79, 140–152. https://doi.org/10.1016/j.eswa.2017.02.046.
Simões, M.G., Shaw, I.S., 2007. Controle E Modelagem Fuzzy, second ed. Edgard Blucher, São Paulo.
Somogyi, Z., Cienciala, E., Mäkipää, R., Muukkonen, P., Lehtonen, A., Weiss, P., 2007. Indirect methods of large-scale forest biomass estimation. Eur. J. For. Res. 126, 197–207. https://doi.org/10.1007/s10342-006-0125-7.
Su, S., Zhao, S., 2017. An optimal clustering mechanism based on Fuzzy-C means for wireless sensor networks. Sustain. Comput. Informatics Syst. https://doi.org/10.1016/j.suscom.2017.08.001.
Torres, C.M.M.E., Kohmann, M.M., Fraisse, C.W., 2015. Quantification of greenhouse gas emissions for carbon neutral farming in the Southeastern USA. Agric. Syst. 137, 64–75. https://doi.org/10.1016/j.agsy.2015.03.002.
Tozer, L., Klenk, N., 2018. Discourses of carbon neutrality and imaginaries of urban futures. Energy Res. Soc. Sci. 35, 174–181. https://doi.org/10.1016/j.erss.2017.10.017.
Vieira, G.C., de Mendonça, A.R., da Silva, G.F., Zanetti, S.S., da Silva, M.M., dos Santos, A.R., 2018. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. Sci. Total Environ. 619–620, 1473–1481. https://doi.org/10.1016/j.scitotenv.2017.11.138.
Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131, 59–75. https://doi.org/10.1016/j.geoderma.2005.03.007.
Wang, L., Zhou, X., Zhu, X., Dong, Z., Guo, W., 2016. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. Crop J 4, 212–219. https://doi.org/10.1016/j.cj.2016.01.008.
Xu, H., Sun, Y., Wang, X., Fu, Y., Dong, Y., Li, Y., 2014. Nonlinear mixed-effects (NLME) diameter growth models for individual China-fir (Cunninghamia lanceolata) trees in southeast China. PLoS One 9. https://doi.org/10.1371/journal.pone.0104012.
Yang, Q., Han, F., Chen, Y., Yang, H., Chen, H., 2016. Greenhouse gas emissions of a biomass-based pyrolysis plant in China. Renew. Sustain. Energy Rev. 53, 1580–1590. https://doi.org/10.1016/j.rser.2015.09.049.
Zhang, G., Lu, Y., 2012. Bias-corrected random forests in regression. J. Appl. Stat. 39, 151–160. https://doi.org/10.1080/02664763.2011.578621.