

Technical Note

Machine Learning Methods for Woody Volume Prediction in Eucalyptus

Dthenifer Cordeiro Santana ¹, Regimar Garcia dos Santos ¹, Pedro Henrique Neves da Silva ² , Hemerson Pistori ^{2,3} , Larissa Pereira Ribeiro Teodoro ⁴ , Nerison Luis Poersch ⁵, Gileno Brito de Azevedo ⁴, Glauce Taís de Oliveira Sousa Azevedo ⁴ , Carlos Antonio da Silva Junior ^{6,*}  and Paulo Eduardo Teodoro ⁴ 

- ¹ Department of Agronomy, State University of São Paulo (UNESP), Ilha Solteira 15385-000, SP, Brazil; dthenifer.santana@unesp.br (D.C.S.); regimar.garcia@unesp.br (R.G.d.S.)
- ² Faculty of Computing, Federal University of Mato Grosso do Sul (UFMS), Campo Grande 79070-900, MS, Brazil; pedro.henrique.neves@ufms.br (P.H.N.d.S.); pistori@ucdb.br (H.P.)
- ³ Department of Computer Engineering, Universidade Católica Dom Bosco (UCDB), Campo Grande 79117-900, MS, Brazil
- ⁴ Campus de Chapadão do Sul, Federal University of Mato Grosso do Sul (UFMS), Chapadão do Sul 79560-000, MS, Brazil; larissa_ribeiro@ufms.br (L.P.R.T.); gileno.azevedo@ufms.br (G.B.d.A.); glauce.azevedo@ufms.br (G.T.d.O.S.A.); paulo.teodoro@ufms.br (P.E.T.)
- ⁵ Department of Agronomy, Federal University of Fronteira do Sul (UFFS), Cerro Largo 97900-000, RS, Brazil; nerison.poersch@uffs.edu.br
- ⁶ Department of Geography, State University of Mato Grosso (UNEMAT), Sinop 78555-000, MT, Brazil
- * Correspondence: carlosjr@unemat.br

Abstract: Machine learning (ML) algorithms can be used to predict wood volume in a faster and more accurate way, providing reliable answers in forest inventories. The objective of this work was to evaluate the performance of different ML techniques to predict the volume of eucalyptus wood, using diameter at breast height (DBH) and total height (Ht) as input variables, obtained by measuring DBH and Ht of 72 trees of six eucalyptus species (*Eucalyptus camaldulensis*, *E. urophylla*, *E. saligna*, *E. grandis*, *E. urograndis*, and *Corymbria citriodora*). The trees were cut down in two different epochs, rendering 48 samples at 24 months and 24 samples at 48 months, and the volume of each tree was measured using the Smailian method. This research explores five machine learning models, namely artificial neural networks (ANN), K-nearest neighbor (KNN), multiple linear regression (LR), random forest (RF) and support vector machine (SVM), to estimate the volume of eucalyptus wood using DBH and Ht. Artificial neural networks achieved higher correlations between observed and estimated wood volume values. However, the RF outperformed all models by providing lower MAE and higher correlations between observed and estimated wood volume values. Therefore, RF is the most accurate for predicting wood volume in eucalyptus species.

Keywords: tree volume; forestry inventory; shallow learner



Citation: Santana, D.C.; Santos, R.G.d.; da Silva, P.H.N.; Pistori, H.; Teodoro, L.P.R.; Poersch, N.L.; de Azevedo, G.B.; de Oliveira Sousa Azevedo, G.T.; da Silva Junior, C.A.; Teodoro, P.E. Machine Learning Methods for Woody Volume Prediction in Eucalyptus. *Sustainability* **2023**, *15*, 10968. <https://doi.org/10.3390/su151410968>

Academic Editor: Qi Feng

Received: 30 April 2023

Revised: 7 July 2023

Accepted: 11 July 2023

Published: 13 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In Brazil, planted forests cover an area of 9.6 million hectares. Of this total, 7.4 million hectares are destined for the production of eucalyptus to manufacture paper, and about 44% of the eucalyptus plantation is concentrated in the southeastern region of the country [1]. Knowing the volume per hectare of wood produced by the plants is important for regulating the continuous supply of raw material, and enabling adequate competitiveness of the sector. Precisely estimating the diameter, height, volume and biomass of trees has economic and ecological importance, as it makes it possible to know the structure of the population, determine commercially important characteristics of the trees, and calculate the carbon stock of the forests [2]. Wood production estimates can be performed by adjusting statistical models, mainly regressions using growth and production information [3].

Obtaining information on variables related to wood production is essential for the rational use of forest resources, and its precise quantification is essential [4]. During the forest inventory process, a portion of the trees are sampled to measure diameter at breast height (DBH) and total height (Ht). These variables are used to predict the volume of wood using several available approaches [3].

These measurements are carried out through manual measurements in the field. However, the evaluation of the volume of wood demands a lot of time and manpower, in addition to being a destructive process, in which the trees are cut for its quantification. The information from DBH and Ht makes it possible to infer the productive capacity of the stand, and provide information on the application of raw material during forest inventories [5]. In addition, this information helps in improvement programs by allowing the selection of trees with better characteristics in advance. Manual quantifications are difficult in fast-growing plantations, where field-based inventory programs may not be sufficient to observe yields at the total area level. The traditional way of estimating both the growth and productivity is described in [6].

The traditional way of estimating both tree growth and productivity is through regression methods [7]. In this way, the use of machine learning (ML) algorithms are tools with the potential to assist in forest inventory [3], characterization, and future property planning [8]. Oliveira et al. [5] found high accuracy in identifying different eucalyptus species based on their growth, by applying the random forest (RF) method. Another ML technique that provides satisfactory results are artificial neural networks (ANN), especially in predicting the diameter of trees [9,10]. ML models circumvent specific problems of forest data when these are nonlinear [11], and have satisfactory performance in obtaining cubage values in forest species [3].

There are some published studies using ML models to predict growth variables in eucalyptus. However, it is still necessary to evaluate the performance of these models for the wood volume. The objective of this work was to evaluate the performance of different ML techniques to predict the volume of eucalyptus wood, using DBH and Ht as input variables.

2. Materials and Methods

2.1. Data Acquisition

The implementation of the experiment took place in January 2014 at the experimental area of the Federal University of Mato Grosso do Sul (UFMS), in Chapadão do Sul, Mato Grosso do Sul State, Brazil (Figure 1). The experimental area has an average altitude of 820 m, and the soil is classified as medium-textured red oxisol. The climate of the region is tropical humid (Aw), with a rainy season from October to April and a dry season between May and September. Annual precipitation throughout the experiment ranged from 1600 to 2100 mm, with average annual temperatures between 26.2 and 27.1 °C.

To set up the experiment, a chemical analysis of the soil was carried out, the results of which were: pH (CaCl₂) = 4.9; organic matter = 31.5 g dm⁻³; P = 13.6 mg dm⁻³; H + Al: 5.4; K = 0.29 cmol dm⁻³; Ca = 2.8 cmolc dm⁻³; Mg = 0.5 cmolc dm⁻³; cation exchange capacity (CEC) = 9.0 cmolc dm⁻³; and base saturation = 39.9%. The proportions of clay, sand and silt were 46%, 46% and 8%, respectively. After this analysis, limestone was applied to raise base saturation to 60%. The initial fertilization was 300 Kg/ha of formulated 4-14-8 (NPK). Crowning, weeding, ant control and application of herbicides (glyphosate) were performed when necessary.

The experimental design was randomized blocks with four replicates, with 20 plants inside each experimental plot. The treatments were composed of four eucalyptus species (*E. camaldulensis*, *C. citriodora*, *E. saligna*, and *E. grandis*) and the GG100 clone (clone of hybrid from *Eucalyptus urophylla* × *Eucalyptus grandis*).



Figure 1. Location of the experimental area with the different species of eucalyptus ($18^{\circ}41'33''$ S, $52^{\circ}40'45''$ W, with altitude is 820 m), at the Federal University of Mato Grosso do Sul (UFMS), Campus of Chapadão do Sul.

The data used in this research were obtained by measuring the diameter at breast height (DBH) and the height (Ht) of 72 trees from eucalyptus species in each plot. To obtain DBH (cm), the tape was used to measure the circumference at breast height, which was converted to DBH, and the Ht (m) was obtained using a Haglof hypsometer. The trees were cut down in 2 different epochs, rendering 48 samples at 24 months and 24 samples at 48 months, and the wood volume (WV) of each tree was measured using the Smalian method (Equation (1)):

$$WV = \sum_{i=1}^{n-1} \frac{g_i + g_{(i+1)}}{2} * L \quad (1)$$

where V is the individual volume of the tree (m^3); g_i is the transverse area at the position (height) i of the stem; L is the section length; and i is the position at the stem, where $i = 1$ (first) is the base of the stem. In each cubed tree, the diameter at breast height (DBH, cm) at 1.30 m from the ground and the total height (Ht, m) was measured with the aid of a measuring tape.

All attributes were applied to machine learning models as follows: isolated DBH (named DBH set) and isolated Ht (named Ht set), combined DBH and Ht (named DBH + Ht set), combined DBH, Ht and the categorical attribute species (named Species set) and combined DBH, Ht, species and the age of each tree (named All set).

2.2. Machine Learning

To test the data, we used five machine learning models: an artificial neural network (ANN) using a multilayer perceptron algorithm, K-nearest neighbor (KNN), multiple linear regression (MLR), random forests (RF), and support vector machine (SVM). All machine learning models were executed on an Intel Core i5 CPU with 8 Gb RAM, with all hyperparameters set as default, according to the Weka (Version 3.9.5) default library. For all models, a randomized stratified 5-fold cross-validation with 10 repetitions was performed, giving a total of 50 runs for each model (Figure 2).

For each machine learning model applied, the Pearson's coefficient (r) and the mean absolute error (MAE) metrics between the DBH, Ht, tree species, and the measured volume were estimated according Equations (2) and (3), respectively. Subsequently, boxplots for r and MAE values were generated to compare the performance of the models by the Scott-Knott test at 5% probability level. This analysis was performed with R software [12] using the *ExpDes.pt* and *ggplot2* packages.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right) \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

where \bar{x} and \bar{y} are the sample means of the observed and predicted mean values; n is the number of samples, y_i is the observed absolute value for each sample, and \hat{y}_i is the predicted absolute value by the model for each sample.

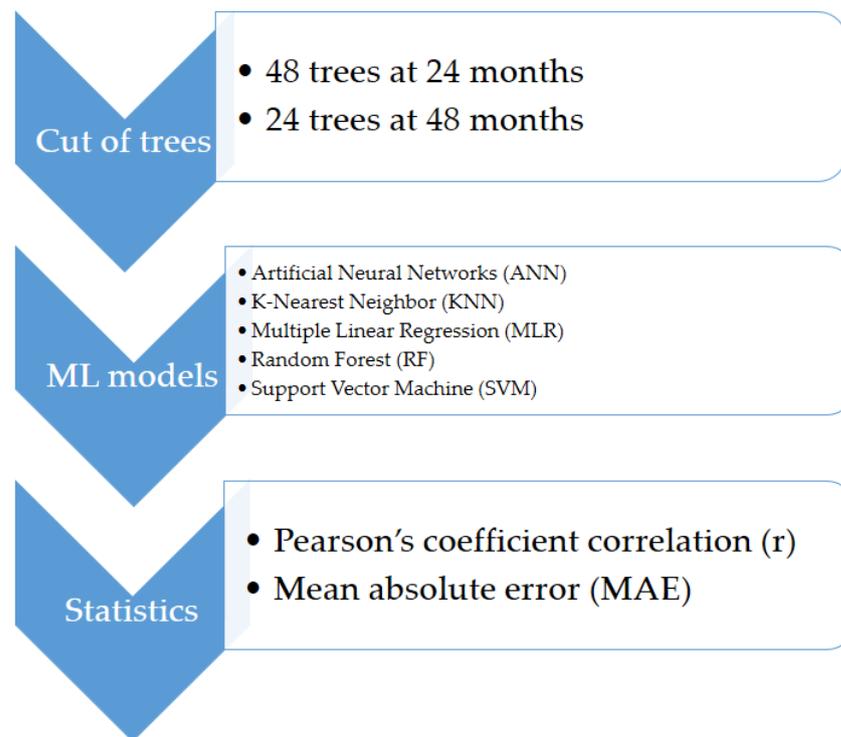


Figure 2. Scheme of procedures performed for data collection, processing and analysis.

2.2.1. Artificial Neural Networks (ANN)

An artificial neural network (ANN) is a machine learning model that simulates human brain behavior through information processing occurring in neurons. Each neuron applies an activation function to process the information, sending it to the next neuron through connection links [10]. ANN models are used in many practical environmental and forest modeling applications, from classification tasks to estimation and prediction. The great advantage of ANNs compared to traditional models is the ability to discover and model relationships between input and output variables without knowing any assumptions about the form of a fitting function [13].

2.2.2. K-Nearest Neighbor (KNN)

The K-nearest neighbor (KNN) is a machine learning classifier that uses Euclidean distance to calculate the similarity between the input data and the training data. KNN models have been used in the Finnish multi-source National Forest Inventory (NFI) since 1990 [14], in New Zealand's coniferous forest areas [15], for forest estimations in the Oregon University/US Forest service's 'coastal landscape analysis and modeling study', and the NASA-funded Upper Midwest Regional Earth Science Application Center (RESAC) [16].

2.2.3. Multiple Linear Regression (MLR)

Multiple linear regression is a statistical method used to predict response values and find the relationship between several independent (or exploratory) variables and the dependent (response) variables [17]. MLR is extensively applied in environmental and forest modeling to estimate dependent variables that are difficult to obtain with direct measurements, such as volume, biomass, and crown diameter, from independent variables such as total height, DBH, and their respective transformations [18,19].

2.2.4. Random Forests (RF)

A random forest is a machine learning classifier that uses a collection of classifiers structured as a tree, with each tree depending on a collection of random variables [20,21]. This model uses a voting scheme among all the trees to classify new instances of the data. RF is applied in many practical applications, from prediction of the predominant kind of tree cover of a forest [22] to ecohydrological modeling [23].

2.2.5. Support Vector Machines (SVM)

A support vector machine (SVM) is a machine learning regression and classifier applied in a wide range of computer vision and pattern recognition problems [24]. Nieto et al. [11] applied SVMs to predict inside-bark volume estimates of standing *E. globulus* trees, using the parameters DBH, height, and age of the tree. SVMs consist of determining a hyperplane to separate the data belonging to two, or more, classes [24].

From the average of the 10 folds of the cross validation, dispersion graphs between the observed versus predicted wood volume and the residuals versus predicted wood volume by the best identified input were generated.

3. Results

The results indicated that every set has a positive correlation with the tree volume (Table 1), with the lowest correlation by KNN using only Ht ($r = 0.8132129$) and the highest correlation obtained by ANN using the combined attributes DBH and Ht ($r = 0.9546309$). However, RF obtained the lowest MAE for each set (Table 2), with the lowest value ($MAE = 0.01938396$) in the All set and correlation ($r = 0.9447906$) differing by only 0.01 from the highest value obtained among all models, ANN using the DBH+Ht set ($r = 0.9546309$ and $MAE = 0.02385956$) and missing 0.4% less.

Table 1. Grouping of Pearson's correlation coefficient (r) between estimated and observed values for the tree volume obtained with different machine learning models and input configurations.

Model	r				
	DBH	Ht	Species	DBH + Ht	All
ANN	0.9137677 Ab	0.8488631 Ac	0.9285823 Ab	0.9546309 Aa	0.9488428 Aa
KNN	0.8378938 Cb	0.8132129 Bc	0.8517617 Bb	0.9374872 Aa	0.8568371 Bb
LR	0.9079053 Aa	0.8138639 Bb	0.9258909 Aa	0.9252162 Aa	0.9455784 Aa
RF	0.8722667 Bb	0.8467742 Ab	0.9424229 Aa	0.9414576 Aa	0.9447906 Aa
SVM	0.9079053 Aa	0.8138639 Bb	0.9246548 Aa	0.9247454 Aa	0.9350522 Aa

Means followed by the same uppercase letters in the same row and the same lowercase letters in the same column do not differ by the Scott-Knott test at 5% probability.

Analyzing the inputs, ANN presented the highest correlation for the sets DBH ($r = 0.9137677$), Ht ($r = 0.8488631$), DBH + Ht ($r = 0.9546309$), and All ($r = 0.9488428$); however, for the Species set, the highest correlations were presented by RF ($r = 0.9424229$). There were no statistical differences between all models (ANN, KNN, LR, RF and SVM) for the DBH + Ht set by the Scott-Knott test at a 5% probability. The categorical attribute species, included with the attributes DBH, Ht and Age in the Species and All sets, presented better correlation results than the isolated attributes DBH and Ht. However, for the ANN and KNN models, the DBH + Ht set obtained better results. The use of all attributes resulted

in the best correlation with volume in LR, RF, and SVM models, and LR and RF, the use of species gave a slight increase in the correlation value. The same result can be observed for MAE, with DBH + Ht set obtaining the lowest values for the ANN and KNN models, and the use of all attributes resulted in the lowest values for MAE in the LR, RF, and SVM models, where the introduction of the species categorical attribute showed no significant reduction in the mean error.

Table 2. Grouping of Mean Absolute Errors (MAE) between estimated and observed values for the tree volume obtained with different machine learning models and input configurations.

Model	MAE				
	DBH	Ht	Species	DBH + Ht	All
ANN	0.03127626 Ab	0.03945802 Aa	0.02844863 Ab	0.02385956 Bc	0.02425266 Cc
KNN	0.02784657 Ac	0.03959366 Aa	0.03139129 Ab	0.02189490 Bd	0.03236239 Ab
LR	0.02974847 Ab	0.04262036 Aa	0.02979906 Ab	0.02948298 Ab	0.02634445 Bb
RF	0.02607193 Ab	0.03590287 Ba	0.02000122 Bc	0.01961697 Bc	0.01938396 Dc
SVM	0.02917617 Ab	0.04083021 Aa	0.02921831 Ab	0.02971024 Ab	0.02752851 Bb

Means followed by the same uppercase letters in the same row and the same lowercase letters in the same column do not differ by the Scott-Knott test at 5% probability.

The boxplots for r and MAE between estimated and observed values for the tree volume obtained with different machine learning models (ANN, KNN, LR, RF and SVM) and input sets (DBH, Ht, Species, DBH + Ht and All) are grouped in Figure S1. The results show that the estimation accuracy is higher for the DBH + Ht set (smaller boxes and close to 0.9367074), and the MAE estimate is lower for the DBH + Ht set, with values close to 0.02491293. However, every machine learning model showed outliers for r , with the exception of ANN in Ht. In the boxplots for MAE, LR only showed an outlier in Ht, without outliers in other sets (DBH, Species, DBH + Ht, and All). The Species set only showed outlier in RF, and the All set showed no outliers in ANN and LR models. Every other combination of ML models and input set showed outliers.

The values observed versus predicted for the tree volume obtained with different machine learning models (Ann, KNN, LR, RF, and SVM), and best input configuration (DBH + HT), are contained in Figure 3. Figure 4 contains the residual versus predicted values for the tree volume obtained with different machine learning models. These results corroborate the results presented in Tables 1 and 2. Overall, the RF was the most accurate model. This model's points were closer to the trend line, especially in the best wooden volumes. That is, while the other models have greater waste for the smallest volumes, the RF presented less dispersion, up to 0.2 m³/tree. From this value, all models have greater dispersion.

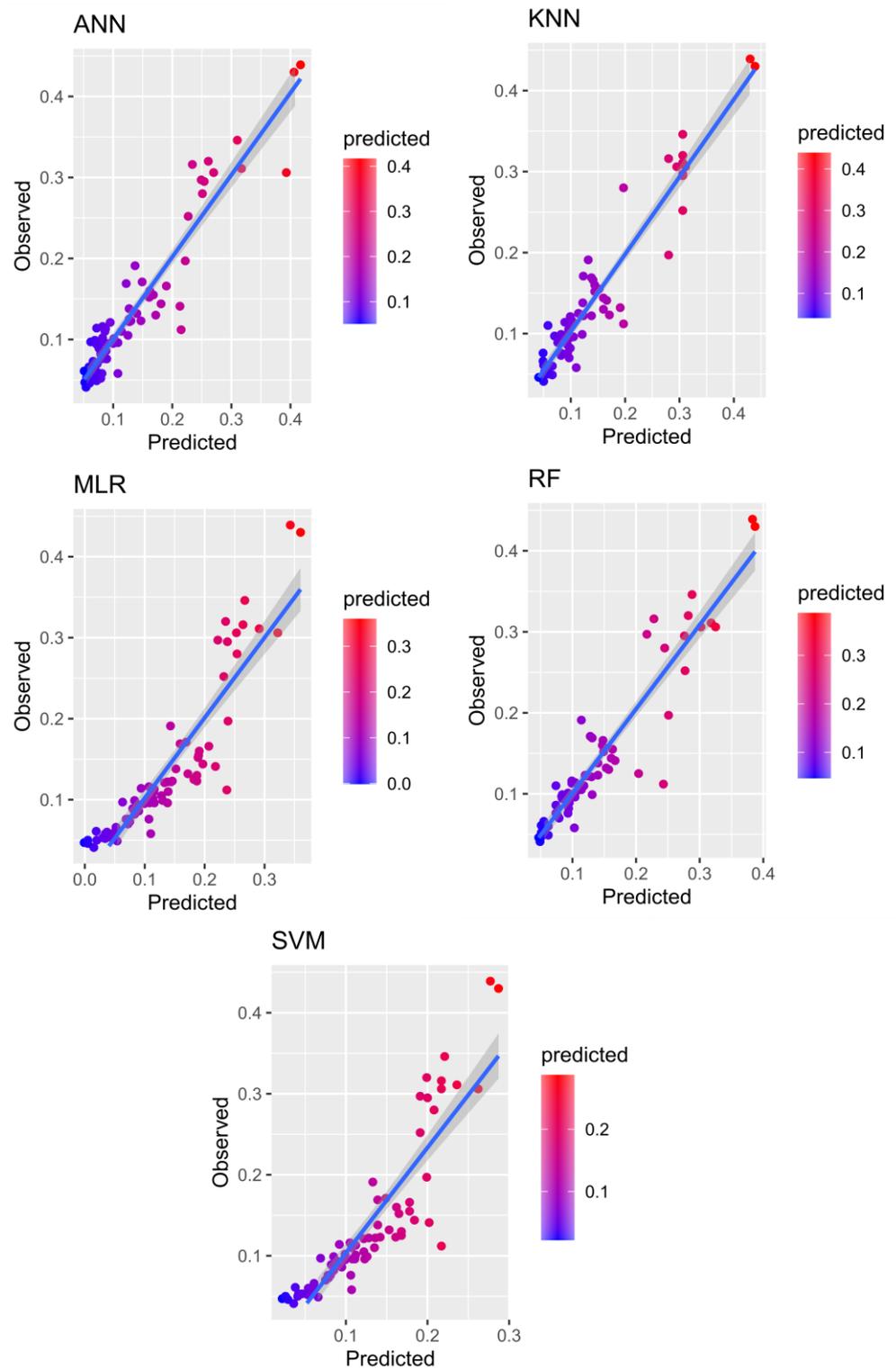


Figure 3. Observed versus predicted values for the tree volume obtained with different machine learning models: artificial neural network (ANN), K-nearest neighbor (KNN), multiple linear regression (MLR), random forest (RF), and support vector machine (SVM), using the best input (DBH + Ht).

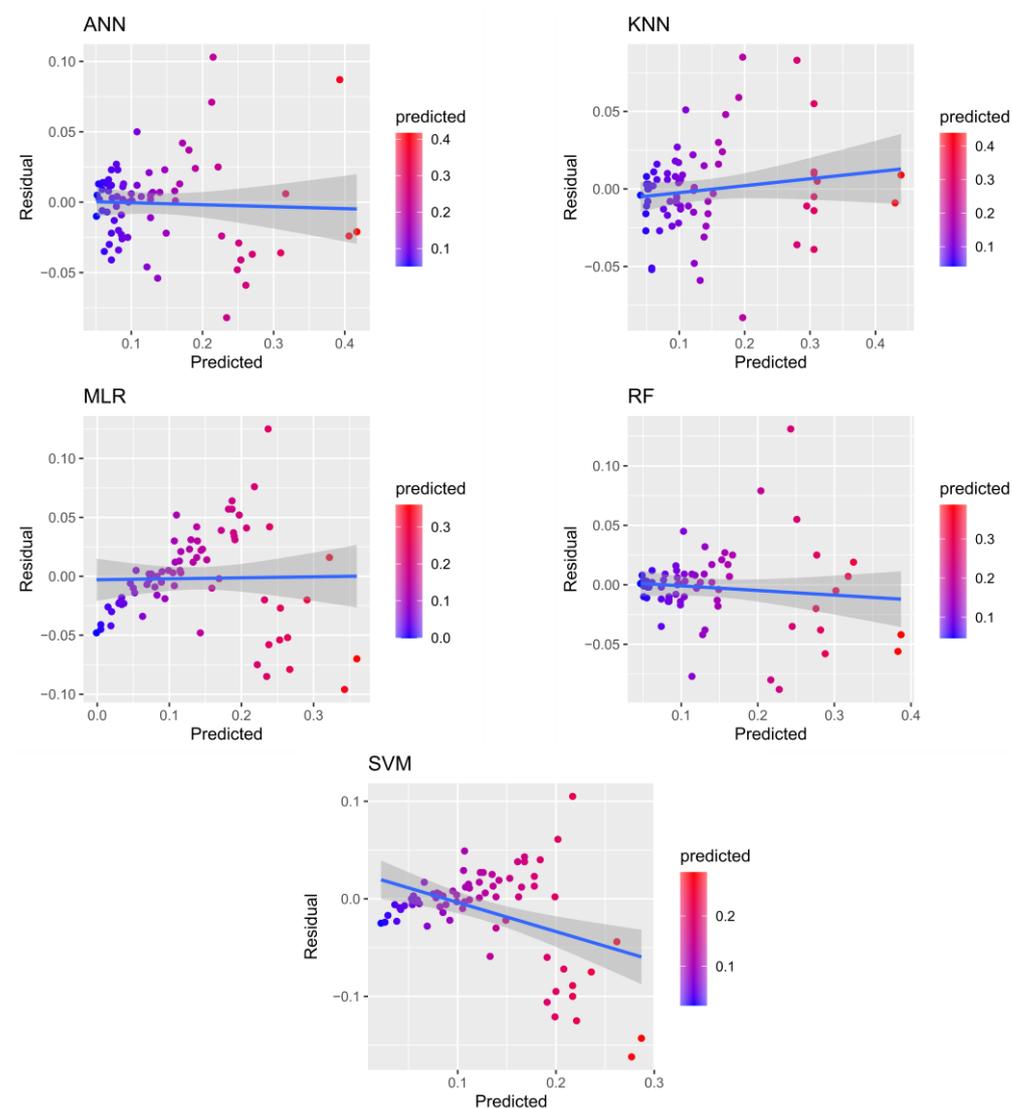


Figure 4. Residual versus predicted values for the tree volume obtained with different machine learning models: artificial neural network (ANN), K-nearest neighbor (KNN), multiple linear regression (MLR), random forest (RF), and support vector machine (SVM), using the best input (DBH + Ht).

4. Discussion

The use of different methods to predict the cubage is a recurrent activity in planted eucalyptus forests. These methods have gained importance in reducing costs in forest management planning, and improving the accuracy of production estimates for the observed areas. The DBH and Ht variables are the main data used to estimate the cubage [25].

The relationship between diameter and height is widely used for both linear and nonlinear regression models, where they present good results [25]. However, ML methods have shown better results because they can learn and generalize data with tolerance to noise and identify nonlinear factors [26]. Environmental factors can directly affect DBH and Ht; this interaction has a nonlinear relationship, because this traditional equations are not able to accurately predict the relationship between these variables [27]. This fact leads to the application of non-parametric methods, such as KNN, LR, RF, ANN, and SVM. In order to compare the predicted and observed data in the field, the Smalian method was chosen, used by several authors to obtain the cubage [28–30].

In all models, a high correlation was observed in the volume estimates, with the ANN providing the highest accuracy. This result corroborates the findings reported by Almeida et al. [31], demonstrating that using ANNs is a good alternative to traditional regression

models for many factors in eucalyptus stands [32]. Artificial neural networks are widely used in forest productivity modeling, as they deal well with the complex relationships between independent and dependent variables [33]. Using the Smalian method to calculate the marketable volume of trees, combined with neural networks to predict diameter and volume which obtain results close to the real ones, are satisfactory techniques to obtain such answers with precision [4].

However, RF obtained lower MAE for all analyzed variables and higher means for the Species set. The RF model has been little used in Brazilian eucalyptus studies, proving to be an accurate model for such estimates and, hence, a good tool for forest management planning [3]. The RF method presents itself as an efficient method for predicting the cubage of eucalyptus, especially when using the attributes DBH, Ht, and species as input data, with satisfactory results found by Araújo et al. [34].

Oliveira et al. [5] found good results from using RF to identify eucalyptus species using growth information. RF also showed good performance in predicting the volume of eucalyptus, increasing the efficiency in monitoring the stem volume of forest plantations [3]. Maire et al. [35] found an r^2 of 0.90 for volume and 0.92 for height using RF, which showed better results. RF has the ability to process data from different sources: satellite, terrestrial, staggered, non-uniform, etc., with high accuracy, and avoids model overfitting [36].

Another important point is that the inclusion of species information in the RF model did not provide a gain in the accuracy of wood volume prediction. This result suggests that the model can be used in other eucalyptus species. The model to run on Free Weekka software is available at the link: https://drive.google.com/drive/folders/11YDwz2Z5hiNdbGtEt7Y3_SyiljqnJXbO?usp=sharing, accessed on 29 April 2023. The step-by-step to use this model is described in the Supplementary Material.

The methods traditionally used in the cubage estimation, although consolidated, have low efficiency when nonlinear factors are considered. New forecasting techniques have become promising, due to their assertiveness in estimating wood volume. Our findings support the applicability of the methods tested here. In this way, the use of a higher number of independent variables, such as climatic, geographic, remote sensing, and soil factors, is encouraged, bringing higher accuracy in estimates and decreasing operating costs.

5. Conclusions

This research explored machine learning algorithms to estimate the volume of eucalyptus wood using DBH and Ht. Artificial neural networks had a good performance, by achieving a higher r between observed and estimated values of wood volume. However, random forest had a lower MAE and slightly lower r than ANN. Therefore, we consider RF the most accurate model for predicting wood volume in eucalyptus species. It is important to highlight that the inclusion of the categorical variable species did not statistically improve the prediction of this model. This indicates that the model can be applied to other eucalyptus species different from those tested in this research.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/su151410968/s1>, Supplementary Material contains the boxplots of the Pearson correlation coefficient (r) and Mean Absolute Error (MAE) values between estimated and observed values for the tree volume obtained with different machine learning models (ML). In addition, it contains information on how the best ML model can be accessed and used by interested readers.

Author Contributions: Conceptualization, D.C.S., R.G.d.S., L.P.R.T. and P.E.T.; methodology, D.C.S., H.P., C.A.d.S.J., P.E.T., N.L.P., L.P.R.T. and P.E.T.; formal analysis, D.C.S., P.H.N.d.S., G.T.d.O.S.A. and G.B.d.A.; investigation, D.C.S., L.P.R.T., P.E.T. and C.A.d.S.J.; writing—original draft preparation, D.C.S., L.P.R.T. and L.P.R.T.; writing—review and editing, H.P., C.A.d.S.J. and P.E.T.; supervision, P.E.T. and D.C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001, National Council for Research and Development (CNPq). We would also like to thank the anonymous reviewers for providing insights to improve the manuscript. We are also thankful to the research laboratory of the Federal University of Mato Grosso do Sul (UFMS), Dom Bosco Catholic University (UCDB), and State University of Mato Grosso (UNEMAT)—<https://pesquisa.unemat.br/gaaf/>. Thanks to Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT) to numbers 88/2021, and 07/2022, and SIAFEM numbers 30478 and 31333; Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso (FAPEMAT) for the financial support of the research project (0001464/2022 and 000125/2023); and CNPq Research Productivity Scholars (processes 309250/2021-8; 306022/2021-4; 303767/2020-0; 304979/2022-8).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the study design; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. PEVS 2020: Com Crescimento de 17.9%, Valor da Produção de Silvicultura e Extração Vegetal Chega a R\$ 23,6 Bilhões. Agência de Notícias. Available online: <https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/31802-pevs-2020-com-crescimento-de-17-9-valor-da-producao-de-silvicultura-e-extracao-vegetal-chega-a-r-23-6-bilhoes> (accessed on 18 March 2022).
2. Gonzalez-Benecke, C.A.; Fernández, M.P.; Gayoso, J.; Pincheira, M.; Wightman, M.G. Using Tree Height, Crown Area and Stand-Level Parameters to Estimate Tree Diameter, Volume, and Biomass of *Pinus radiata*, *Eucalyptus globulus* and *Eucalyptus nitens*. *Forests* **2022**, *13*, 2043. [CrossRef]
3. da Silva, V.S.; Silva, C.A.; Mohan, M.; Cardil, A.; Rex, F.E.; Loureiro, G.H.; Klauberg, C. Combined Impact of sample size and modeling approaches for predicting stem volume in *Eucalyptus* spp. forest plantations using field and LiDAR data. *Remote Sens.* **2020**, *12*, 1438. [CrossRef]
4. Soares, F.A.A.M.N.; Flôres, E.L.; Cabacinha, C.D.; Carrijo, G.A.; Veiga, A.C.P. Recursive diameter prediction for calculating merchantable volume of *Eucalyptus* clones without previous knowledge of total tree height using artificial neural networks. *Appl. Soft Comput. J.* **2012**, *12*, 2030–2039. [CrossRef]
5. de Oliveira, B.R.; da Silva, A.A.P.; Teodoro, L.P.R.; de Azevedo, G.B.; Azevedo, G.T.D.O.S.; Baio, F.H.R.; Teodoro, P.E. *Eucalyptus* growth recognition using machine learning methods and spectral variables. *For. Ecol. Manag.* **2021**, *497*, 119496. [CrossRef]
6. Kainer, D.; Stone, E.A.; Padovan, A.; Foley, W.J.; Külheim, C. Accuracy of Genomic Prediction for Foliar Terpene Traits in *Eucalyptus polybractea*. *G3 Genes Genomes Genet.* **2018**, *8*, 2573. [CrossRef]
7. da Silva, A.K.V.; Borges, M.V.V.; Batista, T.S.; da Silvia Junior, C.A.; Furuya, D.E.G.; Prado Osco, L.; Pistori, H. Predicting *eucalyptus* diameter at breast height and total height with uav-based spectral indices and machine learning. *Forests* **2021**, *12*, 582. [CrossRef]
8. Vega, M.; Harrison, P.; Hamilton, M.; Musk, R.; Adams, P.; Potts, B. Modelling wood property variation among Tasmanian *Eucalyptus nitens* plantations. *For. Ecol. Manag.* **2021**, *491*, 119203. [CrossRef]
9. Diamantopoulou, M.J.; Özçelik, R.; Crecente-Campo, F.; Eler, Ü. Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. *Biosyst. Eng.* **2015**, *133*, 33–45. [CrossRef]
10. Özçelik, R.; Diamantopoulou, M.J.; Brooks, J.R.; Wiant, H.V. Estimating tree bole volume using artificial neural network models for four species in Turkey. *J. Environ. Manag.* **2010**, *91*, 742–753. [CrossRef]
11. García Nieto, P.J.; Martínez Torres, J.; Araújo Fernández, M.; Ordóñez Galán, C. Support vector machines and neural networks used to evaluate paper manufactured using *Eucalyptus globulus*. *Appl. Math. Model.* **2012**, *36*, 6137–6145. [CrossRef]
12. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2014.
13. Karatepe, Y.; Diamantopoulou, M.J. Investigation of Parametric and Artificial Neural Network Modeling Approaches for Total Tree Height Prediction in Cedar Plantations. 2021. Available online: <https://www.researchsquare.com/article/rs-96662/v2> (accessed on 10 April 2023).
14. Designing a Satellite Image-Aided National Forest Survey in Finland [NFI]. Available online: <https://agris.fao.org/agris-search/search.do?recordID=SE9100028> (accessed on 18 March 2022).
15. Trotter, C.M.; Dymond, J.R.; Goulding, C.J. Estimation of timber volume in a coniferous plantation forest using Landsat TM. *Int. J. Remote Sens.* **1997**, *18*, 2209–2223. [CrossRef]

16. Reese, H.; Nilsson, M.; Sandstro, P. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Comput. Electron. Agric.* **2002**, *37*, 37–55. [CrossRef]
17. Alexopoulos, E.C. Introduction to multivariate regression analysis. *Hippokratia* **2010**, *14*, 23–28.
18. MacHado, M.V.; Tommaselli, A.M.G.; Tachibana, V.M.; Martins-Neto, R.P.; Campos, M.B. Evaluation of multiple linear regression model to obtain dbh of trees using data from a lightweight laser scanning system on-board a uav. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 449–454. [CrossRef]
19. Zhou, R.; Wu, D.; Zhou, R.; Fang, L.; Zheng, X.; Lou, X. Estimation of DBH at forest stand level based on multi-parameters and generalized regression neural network. *Forests* **2019**, *10*, 778. [CrossRef]
20. Breiman, L. Random forests. *Random For.* **2019**, *1*, 1–33.
21. Cutler, A.; Cutler, D.R.; Stevens, J.R. Ensemble Machine Learning. *Methods Appl.* **2012**, *1*, 1–332.
22. Agrawal, S.; Rana, S.; Ahmad, T. Random forest for the real forests. *Adv. Intell. Syst. Comput.* **2016**, *381*, 301–309.
23. Peters, J.; De Baets, B.; Verhoest, N.E.; Samson, R.; Degroeve, S.; De Becker, P.; Huybrechts, W. Random forests as a tool for ecohydrological distribution modelling. *Ecol. Modell.* **2007**, *207*, 304–318. [CrossRef]
24. Nalepa, J.; Kawulok, M. Selecting training sets for support vector machines: A review. *Artif. Intell. Rev.* **2019**, *52*, 857–900. [CrossRef]
25. Temesgen, H.; Zhang, C.H.; Zhao, X.H. Modelling tree height-diameter relationships in multi-species and multi-layered forests: A large observational study from Northeast China. *For. Ecol. Manag.* **2014**, *316*, 78–89. [CrossRef]
26. Were, K.; Bui, D.T.; Dick, Ø.B.; Singh, B.R. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afrotropical landscape. *Ecol. Indic.* **2015**, *52*, 394–403. [CrossRef]
27. Shen, J.; Hu, Z.; Sharma, R.P.; Wang, G.; Meng, X.; Wang, M.; Fu, L. Modeling height-diameter relationship for poplar plantations using combined-optimization multiple hidden layer back propagation neural network. *Forests* **2020**, *11*, 442. [CrossRef]
28. Campos, O.J.D. Cubagem de árvores. *Master Diss.* **2014**, 87. Available online: <https://repositorio.ufsc.br/bitstream/handle/123456789/123279/327161.pdf?sequence=1&isAllowed=y> (accessed on 9 April 2022).
29. Leal, F.A.; Cabacinha, C.D.; Vinícius, R.; Castro, O.; Aparecido, E.; Matricardi, T. AMOSTRAGEM DE ÁRVORES DE EUCALYPTUS NA CUBAGEM 1 Introdução 2 Material e método. *Rev. Bras. Biom.* **2015**, *33*, 91–103.
30. Müller, M.D.; Salles, T.T.; Paciullo, D.S.C.; Brighenti, A.M.; Castro, C.D. Equações De Altura, Volume E Afilamento Para Eucalipto E Acácia Estabelecidos Em Sistema Silvopastoril. *Floresta* **2014**, *44*, 473. [CrossRef]
31. de Almeida, M.R.D.; Silva, J.N.M.; de Barros, P.L.C.; da Silva Almeida, E.; da Silva, D.A.S.; de Sousa, C.S.C. Adjustment and selection of volumetric models for commercial species in Ipixuna. *Rev. Em Agronegócio E Meio Ambiente* **2020**, *13*, 259–278.
32. da Silva Binoti, M.L.M.; Binoti, D.H.B.; Leite, H.G. Height of Even-Aged Stands of Eucalyptus. *Rev. Árvore* **2013**, *37*, 639–645.
33. de Freitas, E.C.S.; de Paiva, H.N.; Neves, J.C.L.; Marcatti, G.E.; Leite, H.G. Modeling of eucalyptus productivity with artificial neural networks. *Ind. Crops Prod.* **2020**, *146*, 112149. [CrossRef]
34. David, R.A.R.; Santos, A.C.G.D.; Ferreira, M.A.; Freitas, D.; Dias, N.D.S.; Camargos, B.H.L.; Gomide, L.R. Aplicação De Técnicas De Regressão Linear E Aprendizagem De Máquinas Na Predição Da Altura Total De Árvores De Eucalyptus Spp. In *Silvicultura E Manejo Florestal: Técnicas De Utilização E Conservação Da Natureza-Volume 1*; Editora Científica Digital: Guarujá, Brazil, 2021; Volume 1, pp. 29–43.
35. Le Maire, G.; Marsden, C.; Nouvellon, Y.; Grinand, C.; Hakamada, R.; Stape, J.L.; Laclau, J.P. MODIS NDVI time-series allow the monitoring of Eucalyptus plantation biomass. *Remote Sens. Environ.* **2011**, *115*, 2613–2625. [CrossRef]
36. Prasad, N.R.; Patel, N.R.; Danodia, A. Crop yield prediction in cotton for regional level using random forest approach. *Spat. Inf. Res.* **2021**, *29*, 195–206. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.