UFSM

Ciência Florestal

ISSN 1980-5098
OPEN ACCESS

Ci. Fl., Santa Maria, v. 35, e90612, p. 1-21, 2025 • https://doi.org/10.5902/1980509890612
Submitted: 16th/01/2025 • Approved: 26th/08/2025 • Published: 29th/10/2025

Articles

Accuracy of tree profile estimation in planted forests

Acurácia na estimação do perfil de árvores em florestas plantadas

Thomaz Correa e Castro da Costa[I]
Henrique Coelho Mendes[II]

[I]Embrapa, Sete Lagoas, MG, Brazil
[II]Zup I.T. Serviços Tecnologia e Inovação S.A., Uberlândia, MG, Brazil

ABSTRACT

Theoretically, the tree shape is composed of three geometric segments: a neiloid at the base, a paraboloid in the middle, and a cone at the top. Modeling of this form aims to estimate the allocation of trees to different wood products, such as sawtimber, roundwood, pulp logs, firewood, or charcoal. Taper models can be categorized as non-segmented, segmented, sigmoidal, form exponents, and, more recently, adjusted by artificial neural networks (ANNs). Most studies have focused on selecting the best model by evaluating its performance under homogeneous conditions, such as single species, clone, spacing, and age. This study tested the hypothesis that within variations among trees under the same experimental setup can produce modeling errors that are as significant as those resulting from differences among trees across distinct conditions. To this end, data from various species, clones, ages, and spacing under a wide range of edaphic and climatic conditions were integrated. Different taper models were applied and their errors were compared using the most common performance metrics in the literature: the percentage residual standard error (Syx%) and, when unavailable, the root mean square error (RMSE). The results showed that modeling with integrated datasets produced intermediate error values compared with those reported in other studies. This suggests that, even when spacing, age, and environmental conditions are held constant, shape variation among trees of the same species or clones limits gains in precision, even when using models better suited to specific stem profiles.

Keywords: Taper; Tree shape; Derived from Kozak; ANN

RESUMO

A forma da árvore é teoricamente composta por três segmentos geométricos: um sólido neilóide na base, um parabolóide na porção central e um cone na ponta. A modelagem dessa forma visa estimar a repartição da árvore em diferentes produtos madeireiros, como madeira serrada, roliça, fustes para celulose, lenha ou carvão. Os modelos de afilamento (taper) podem ser classificados como não segmentados, segmentados, sigmoides, de expoente-forma e, mais recentemente, ajustados por redes neurais artificiais (RNA). A maioria dos estudos busca o melhor modelo avaliando seu desempenho em condições homogêneas, como mesma espécie, clone, espaçamento e idade. Este trabalho testa a hipótese de que as variações internas, entre árvores dentro de uma mesma condição experimental, podem gerar erros na modelagem tão relevantes quanto as variações entre árvores em condições distintas. Para isso, foram integrados dados de diferentes espécies, clones, idades e espaçamentos, sob variadas condições edáficas e climáticas. Diferentes modelos de afilamento foram aplicados e seus erros foram comparados com as métricas mais utilizadas na literatura: o erro padrão residual percentual (Syx%) e, quando ausente, o erro quadrático médio (RMSE). Os resultados mostraram que a modelagem integrando as diferentes amostras apresentou erros intermediários no rank dos erros encontrados nas demais referencias, indicando que existem variações de forma em árvores da mesma espécie, e em clones, fixando-se o espaçamento e a idade, nas mesmas condições edáficas e climáticas, que impedem um aumento de precisão, quando comparadas aos dados analisados conjuntamente, mesmo utilizando modelos mais aderentes aos perfis analisados.

Palavras-chave: Afilamento; Forma da árvore; Derivada de Kozak; RNA

1 INTRODUCTION

Scaling of sample trees in forest inventories is necessary when a forest product is intended as round wood and/or sawn timber. This method involves measuring the shape of the tree to obtain taper functions, enabling the segmentation of the tree into various forest products along with the respective volumes and quantities that the forest could potentially supply.

The most comprehensive reviews and comparisons of taper models were conducted by Andrade (2014), Salekin; Catalán; Boczniewicz *et al*. (2021), Andrade; Terra; Carvallho (2022). The first study evaluated the efficiency of 18 models applied to the *Eucalyptus urophylla x Eucalyptus grandis* hybrid by referencing each model. The second reviewed 910 references on the subject, all in English, categorizing taper models from the early 20th century to the present day, utilizing AI tools. The third reviewed and compared 46 tapered models applied to *Corymbia citriodora*. Other relevant references, particularly from Brazil, include studies on the application of artificial neural networks (ANNs), as discussed by Soares *et al*. (2011), Sakici; Ozdemir (2018), Socha; Netzel; Cywicka (2020), Sandoval; Acuña (2022), Seki (2023).

The earliest taper models belong to the “non-segmented” category, which assumes a single function to describe the entire stem. Typically, the shape of a tree resembles a neiloid solid at the base, a paraboloid in the middle, and a cone at the top (Souza; Silva; Xavier *et al*., 2008). This complex geometry can result in more than one inflection point and a reversal in the rate of variation along the trunk surface, which mathematically occurs when the second derivative of a function changes sign.

To better fit the natural shape of the trunk, which changes trends at inflection points, sigmoidal models have been proposed (Guimarães; Leite, 1992), along with segmented models such as those by Max and Burkhart, Demaerschalk, and Kozak, as well as form-exponent models (Kozak *et al*., 1988, cited by Andrade, 2014), Souza *et al*. (2008), and Andrade; Terra; Carvalho (2022). In the field of machine learning and artificial intelligence, ANNs have been applied to empirically capture changes in the trunk profile along the tree height (Soares; Flores; Cabacinha *et al*., 2011; Sandoval; Acuña, 2022).

Most studies have established fixed characteristics and conditions for applying taper models, such as species, clones, spacing, age, and climatic and edaphic conditions. In these evaluations, trees in planted forests were sampled, and error metrics were assessed.

This study was based on the hypothesis that variations in tree form exist under the same experimental conditions, that is among trees of the same clone, age, spacing, and under identical edaphic and climatic conditions, can generate taper modeling errors comparable to those observed across different experimental conditions, such as among distinct clones, ages, spacings, or environments.

In other words, it is assumed that intraconditional variability (within a single combination of factors) can be as significant as interconditional variability (across different combinations of factors), similarly affecting the performance of taper models.

To test this hypothesis, different taper models were applied, including a non-segmented model, segmented models, a derivative of Kozak's model, and two ANNs, to a dataset that integrated diverse experimental conditions. The resulting errors were compared using the most frequently reported error metrics in the literature: the percentage residual standard error (Syx%), and when unavailable, the root mean square error (RMSE). The objective was to determine whether the errors obtained from the integrated datasets were of the same order of magnitude as those obtained under more specific and homogeneous conditions as reported in the literature.

2 MATERIAL AND METHODS

2.1 Data

The materials analyzed are described in Table 1 and were collected through rigorous scaling or using a digital dendrometer. The database comprises seminal eucalyptus and pine (*Corymbia citriodora*, *Pinus sp.*, and *Eucalyptus urophylla x Eucaliptus grandis*); eucalyptus clones in integrated crop-livestock-forest systems arranged in rows, including clones GG100 and i144; *Eucalyptus sp.* in pure plantations from Conselheiro Lafaiete-MG and Fazenda Cidade do Boi in Pompéu-MG, an African mahogany (*Khaya grandifoliola*) stand; and Australian cedar clones from Fazenda Bela Vista in Campo Belo-MG.

Observations regarding tree forms were recorded during the measurements. It is important to note that the dataset included trees with natural imperfections, such as stem sweeps and bottle-shaped forms. These characteristics were not isolated or excluded from the analysis because the objective of the study was to evaluate the model performance under the actual variability found under field conditions. Material 1 consisted of suppressed trees from a plantation approximately 60 years old that exhibited some degree of tortuosity. Material 5 contained a clone with a tortuous profile. Some sampled trees from clone i144 (Material 10) in integrated livestock-forest systems displayed a “bottle” shape, characterized by very gradual tapering in the central section and more pronounced tapering near the top. Material 6 originated from a pine stand that was approximately 60 years old and featured trees with large dimensions.

The diameters along the trunk (*di*) were measured at heights (*hi*) of 0, 0.15, or 0.3 m at the base, and then standardized for all materials at 0.7, 1.3, 2.3, 3.3 m, followed by 1 m intervals up to the tree tip, with *di* recorded down to 5 cm. A total of 237 trees were sampled and divided into a 70%–30% ratio for model estimation and statistical testing.

Table 1 – Cubing is carried out in different materials (species, clones) and locations, range of diameters (dbh) and total heights (ht), number of tree bole sections and number of trees, and respective separation for estimating the models and for testing

Id	Material	dap(cm)	ht(m)	n. sections	n. trees	n. trees (70%)	n. trees. (30%)	Method
1	*Corymbia citriodora*	19.4 - 24.0	22.0 - 25.2	35	3	2	1	Digital dendrometer
2	GG100	14.1 - 21.0	19.0 - 31.5	195	10	7	3	Rigorous Cubing
3	GG100, dates 2009/2011/2013	13.1 - 34.9	14.3 - 42.3	1198	63	44	19	Rigorous Cubing
4	African Mahogany	14.4 - 23.4	22.0 - 25.2	93	13	9	4	Digital dendrometer
5	Australian Cedar Clones 1110/1120/1210/1321	20.8 - 28.2	14.7 - 20.9	457	40	28	12	Digital dendrometer
	Pinus sp.	31.3 - 67.6	26.1 – 46.0	593	30	21	9	Digital dendrometer
7	*Eucalyptus* sp.	14.1 - 32.1	24.4 - 45.9	362	22	15	7	Digital dendrometer
8	*Eucayptus urophylla x Eucalyptus grandis*	13.3 - 32.6	9.5 – 29.0	66	8	6	2	Digital dendrometer
9	i144	21.0 – 32.0	26.5 - 28.9	452	29	20	9	Digital dendrometer
10	*Eucalyptus* sp.	12.3 - 20.5	20.8 - 25.2	260	19	13	6	Rigorous Cubing
Total		12.3 - 67.6	9.5 – 46.0	3711	237	166	71	

Source: Authors (2025)

In where: 1) Guarita 2 CNPMS; 2) Maravilhas-MG; 3) ILPF_CNPMS; 4) Silva Xavier_Sete Lagoas-MG; 5) Faz. Bela Vista_Campo Belo-MG; 6) Vitrine-CNPMS-Sete Lagoas-MG; 7) Conselheiro Lafaiete-MG; 8) Bambuí-MG; 9) ILPF Faz. Lagoa dos Currais, Codisburgo-MG; 10) Faz. Cidade do Boi, Pompéu-MG.

2.2 Models

Six models were used to determine *di* along the trunk. The respective functions were applied to both the training and test data.

The model by Kozak *et al*. (1969), known for its simplicity and relative performance, is presented in Equation (1):

$$(\mathrm{d_i}/dap)^2 = \mathrm{b_0} + \mathrm{b_1}\frac{h_i}{ht} + b_2\left(\frac{h_i}{ht}\right)^2 \quad (1)$$

where: *di* is the diameter at a given section of the trunk (cm); *hi* is the height of the section (m); *dap* is the diameter at breast height (*hi* = 1.3 m); *ht* is the total tree height (m); *b0*, *b1*, *b2* are the parameters to be estimated.

The derivative of Kozak's function, Equation (2), was used to compute *di* incrementally, starting from *dap*, subtracting the respective rates (*di'*) towards the base and tip of each tree:

$$d_i' = \frac{dap(\mathrm{b_1}ht+2\mathrm{b_2}h_i)}{2ht(\mathrm{b_0}ht^2+\mathrm{b_1}h_iht+\mathrm{b_2}h_i^2)^{\left(\frac{1}{2}\right)}} \quad (2)$$

The sigmoid model from Guimarães and Leite (1992), which was designed to model a trunk with an inflection point at *dap*, is given by Equation (3):

$$ln\left(\frac{d_i}{dap}\right) = b_0 + b_1\left[1 - e^{(1.3-h_i)}\right] + b_2\ln\left(\frac{ht-h_i}{ht-1.3}\right) \quad (3)$$

where: *ln* is the natural logarithm; *e* is Euler's number.

The segmented Max-Burkhart model, modified to account for *dap* as the contact point between the polynomials, is shown in Equation (4):

$$(\mathrm{d_i}/dap)^2 = \mathrm{b_0} + \mathrm{b_1}\frac{h_i}{ht} + b_2\left(\frac{h_i}{ht}\right)^2 + b_3(1.3 - \frac{h_i}{ht})^2 I \quad (4)$$

where: I = 1 se *hi* <= 1.3 m; I = 0 se *hi* > 1.3 m; b_0, b_1, b_2, b_3, the parameters to be estimated.

Models 5 and 6 were developed using ANNs to estimate the diameters (*di*) at various tree heights based on the independent variables of relative height (*hi*), diameter at breast height (*DBH*), and total tree height (ht). The decision to use ANNs was driven by their ability to capture complex non-linear relationships among variables, which are often present in taper modeling.

Model 5 is implemented in Python using the Keras library with a TensorFlow backend. The Adam optimizer, which promotes fast and stable convergence during training, was selected owing to its robust adaptive weight update mechanism. The neural network architecture consisted of four fully connected ("dense") layers, structured as follows: (1) an input layer with 4 neurons corresponding to *hi*, *DBH*, *ht*, and a bias constant; (2) a first hidden layer with 8 neurons; (3) a second hidden layer with 4 neurons; and (4) an output layer with 1 neuron that returns the estimated di, using a linear activation function. The hidden layers employ a Rectified Linear Unit (ReLU) activation function, which is widely used owing to its computational efficiency.

The numbers of layers and neurons were determined empirically by testing architectures with different depths (two to four layers) and widths (four to sixteen neurons per layer) and comparing their performances based on the residual standard error (Syx%) and RMSE on the validation dataset. The selected architecture achieves a good balance between performance and complexity. Additional hyperparameters defined after the preliminary tests included: (a) training epochs: 300, sufficient for validation error stabilization; (b) batch size: 50, deemed appropriate for the dataset size; and (c) loss function: mean squared error (MSE), the standard for continuous regression tasks.

Model 6 was developed using the proprietary Tiberius software, which was designed for data analysis and mining using neural networks. However, Tiberius imposes some restrictions on architectural customization, unlike the flexibility offered by Keras with TensorFlow. The ANN structure in Tiberius included (1) an input layer receiving the same explanatory variables (*hi, DBH,* and *ht*), (2) a hidden layer with four neurons (with activation automatically configured by the software), and (3) an output layer with one neuron.

Owing to software limitations, it was not possible to manually adjust hyperparameters such as the number of epochs, learning rate, or activation functions. Model selection relied on Tiberius's internal validation function, which uses automatic cross-validation to select the best-performing architectural model among the available options. Despite its lower flexibility, Model 6 was included in this study as a practical and accessible alternative for users with limited programming expertise, in contrast to Model 5, which is more sophisticated and customizable.

The inclusion of two ANN models with different customizations was intended to demonstrate the effects of architecture and flexibility on taper modeling. Although Model 5 offers a greater potential for fine-tuning, Model 6 utilizes a more practical approach.

2.3 Evaluation metrics

The following metrics were used to evaluate the performance of the models applied to the tapered dataset (Eqs. 5–9). The adjusted coefficient of determination (adjusted R^2) is initially included (Eq. 5) because it is a commonly used metric for evaluating regression models. However, its application to ANNs is limited because of the large number of parameters distributed across multiple hidden layers. This characteristic prevents the calculation of the adjusted R^2. Therefore, the R^2 values for the ANNs were retained as supplementary information but were not used as a primary metric for model comparison.

The standard error of the estimation (Eq. 6) does not have the same interpretation as that of the mean-squared error (Eq. 7) because Syx% is generally applied to data that fit the equation, indicating the average dispersion of observed values relative to modeled values, whereas RMSE% is applied to test data to report the average error when applying the model. The only difference between their formulas is the number of parameters (*p*) in the function. For large sample sizes (*n*), both statistics were similar when applied to the same dataset.

To compare the errors in this study with those obtained from other studies, Syx% and RMSE% statistics were used, with RMSE% used in the absence of Syx% information. Through these statistics, it is possible to obtain a ranking of errors among the references resulting from the application of the taper models.

Adjusted R-squared in Equation (5):

$$\bar{R}^2 = 1 - \left(\frac{n-1}{n-p-1}\right) * \left(1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}\right) \tag{5}$$

The residual standard Error (as a percentage) (Syx), in Equation (6):

$$S_{XY}\% = \frac{\sqrt{\frac{\sum(y_i - \hat{y})^2}{n-p}}}{\bar{y}} * 100 \tag{6}$$

The root mean square error (as a percentage) (RMSE), in Equation (7):

$$RMSE\% = \frac{\sqrt{\frac{\sum(y_i - \hat{y})^2}{n}}}{\bar{y}} * 100 \tag{7}$$

Mean Absolute Percentage Error (MAPE), Equation (8):

$$MAPE = \frac{1}{n}\sum\frac{|\hat{y}_i - y_i|}{y_i} \tag{8}$$

Bias, Equation (9):

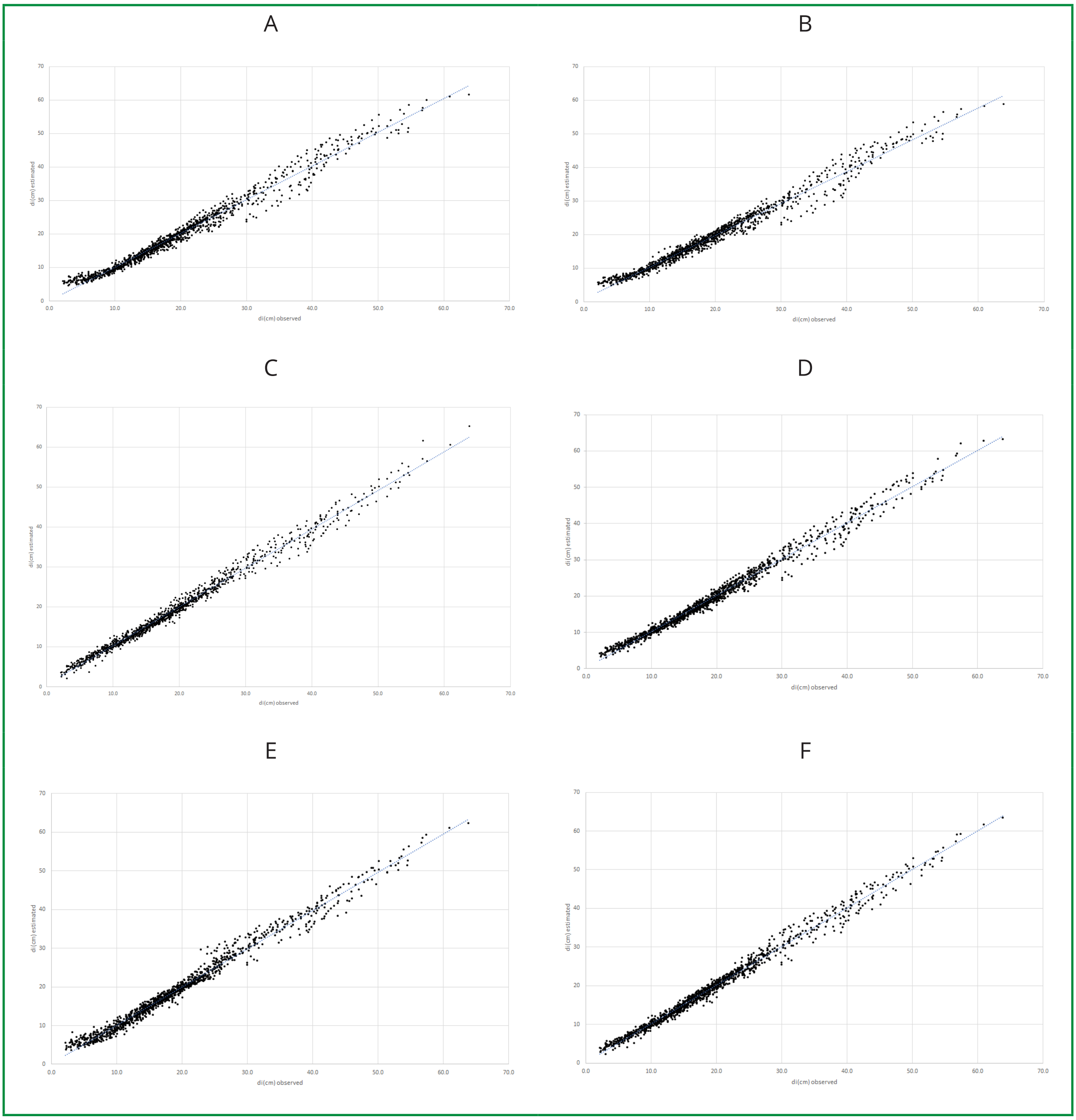
$$Vies = \frac{1}{n}\sum|\hat{y}_i - y_i| \tag{9}$$

The consistency of the results was further evaluated by the frequency distribution of the percentage residuals ($PE = (\hat{y}_i - y_i)/y_i * 100$) in the test data. The percentage standard error (PSE = $\frac{s}{\sqrt{n}} * t_{(5\%,n-1)g.l.}$) was calculated using the diameter class, considering the number of sections per diameter class as the sample (n).

3 RESULTS

Figure 1 shows the dispersion between the observed and estimated values when the six models were applied. Smaller data dispersions were observed in Figures 1c and f, corresponding to the Guimarães model, followed by ANN 6.

Figure 1 – Dispersion between observed and estimated *di* values from the test dataset, Kozak model (A); by rate of variation based on the dbh of each tree (B); by Guimarães model (C); by segmented model (D); by ANN 5 (E) and ANN 6 (F)



Source: Authors (2025)

The evaluation metrics presented in Table 2 indicate the quality of fit. For example, the equation from the Kozak model explained 89.8% of the total variation in the data (R^2), with an RMSE indicating an average error of 8.39%, which represented 6% (MAPE) of the *di* value (diameter measured along the trunk) with a mean tendency to overestimate the true value by 0.094 cm. Based on this set of metrics, the Guimarães and ANN 6 models showed better accuracy, although with a larger bias for ANN 6.

The error distribution is the most important evaluation metric because it provides information on the amplitude and probability of errors (Figure 2). The percentage errors in *di* were obtained from the test data. The Guimarães model exhibits greater symmetry. In contrast, the Kozak derivative and ANN 6 were the most asymmetric. In terms of "precision," the models presented similar results, although the amplitude for neural network 6 was the narrowest.

Of the *di* errors, 95.4% fell between –16% and 17% for the Kozak model, –17% and 18% for the Kozak derivative, –13% and 15% for both the Guimarães and segmented models, –17% and 16% for ANN 5, and between –12% and 13% for ANN 6, assuming a normal distribution (calculations not shown).

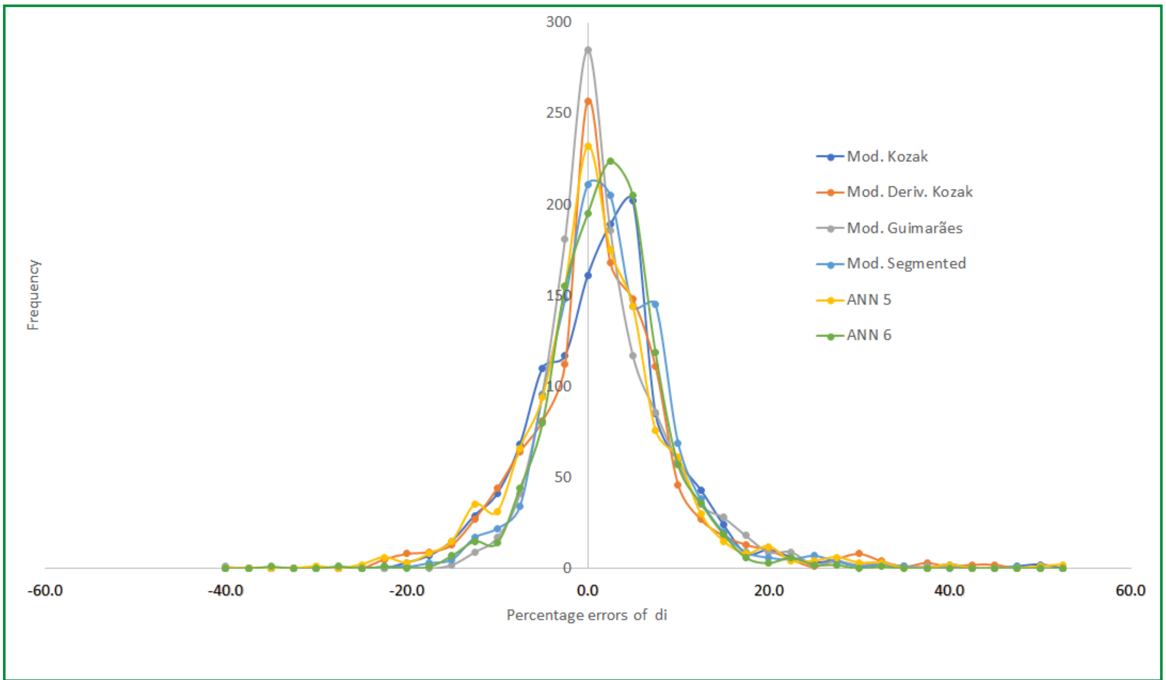
Table 2 – Coefficients, adjusted R^2, root mean square error (RMSE), mean proportional error (MAPE), and bias

Model	b1**	b2**	b3**	b4**	R^2 ajust	Syx %	RMSE % ***	MAPE ***	bias (cm)***
Kozak	1.16945	-2.25325	1.21256		0.898	9.14	8.39	0.059	0.094
Kozak-derived	1.16945	-2.25325	1.21256		0.892	9.49	8.55	0.058	-0.181
Guimarães	-0.01594	-0.07538	0.57745		0.984	6.51	6.05	0.050	-0.019
Segmented	0.99832	-1.49529	0.52343	0.13707	0.980	7.38	6.82	0.051	0.152
ANN 5					0.977*	7.85	7.10	0.057	-0.041
ANN 6					0.986*	6.13	6.06	0.047	0.142

Source: Authors (2025)

In where: *not adjusted; ** significant at 0.001; ***calculated with test data

Figure 2 – Frequency distribution of the percentage residue of *di* in a 2.5% error class interval



Source: Authors (2025)

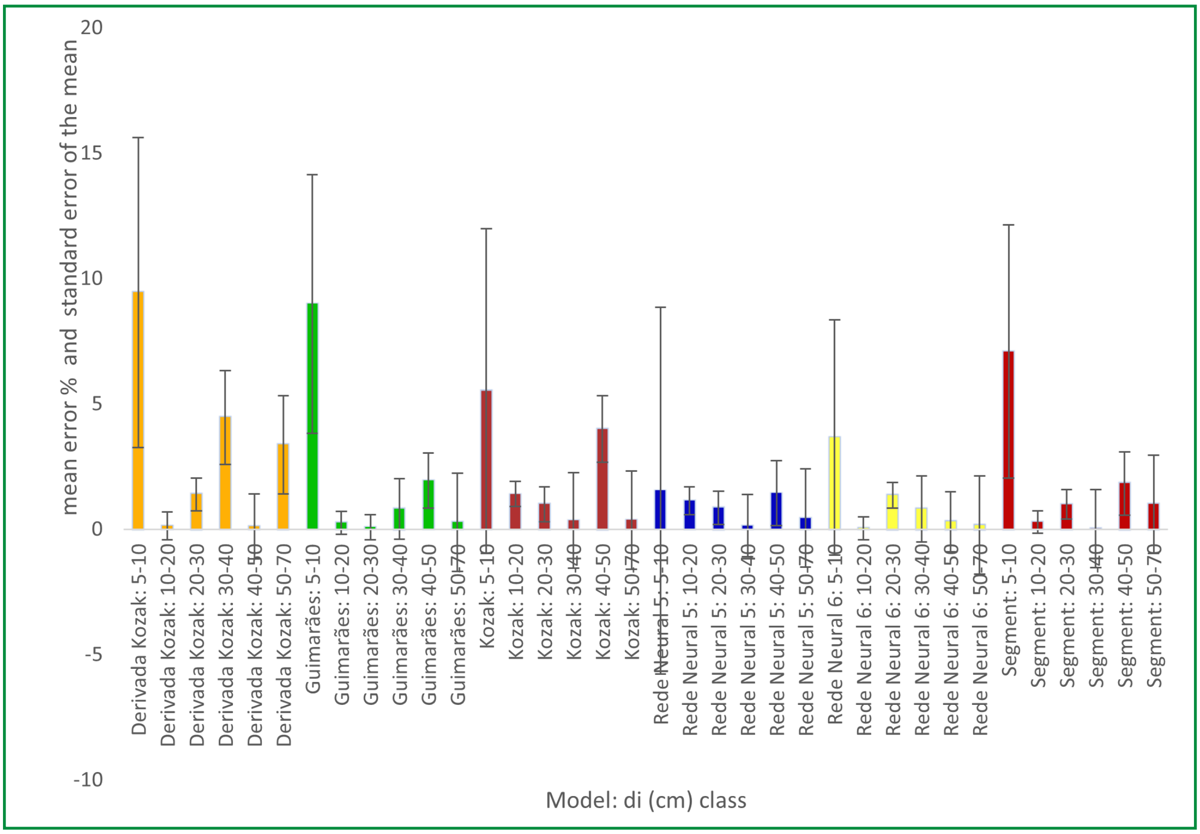
In the diameter class analysis, Figure 3 shows the mean percentage error of di (in absolute terms) and its precision (standard error) for the models applied to the test samples. For smaller diameters (class 5–10), the errors were the highest near the tree tips, except for ANN 5, which had an average error of 1.55% despite having the highest precision (7.3%).

In the intermediate-diameter classes, the error range was smaller, particularly for diameters between 10 and 30 cm. Similar to what was observed for accuracy, the “precision” parameter did not allow us to identify a single model that consistently performed best across all classes.

In the overall analysis (Figure 3), the models by Guimarães and Neural Network 6 showed better performance, except for the 5–10 cm class. In the class-by-class analysis, the Kozak-derived model, generally the lowest performing model, produced more accurate results in the 40–50 cm class. However, it had the highest error in both the largest (50–70 cm) and intermediate (30–40 cm) classes, reinforcing the idea that the model effectiveness can vary depending on the section of the trunk.

ANNs generally have a greater capacity to adapt to non-linear patterns, and lower errors are expected regardless of the diameter class, which was not the case. The sigmoidal model by Guimarães, which features an inflection at breast height (DBH), performed slightly better than the neural networks overall but showed higher errors in the 5–10 cm and 40–50 cm classes. These findings highlight the importance of evaluating models not only through global error metrics but also by selecting those that yield the lowest errors in the most relevant trunk sections.

Figure 3 – Percentual mean error in module per *di* class (5–10 cm, n=20; 10–20, n=547; 20–30, n=300; 30–40, n=99; 40–50, n=60; 50–70 cm, n=168), and respective standard error of the mean, for the models



Source: Authors (2025)

Regarding the joint and specific analyses of materials, the ranking of the residual standard error as a percentage (Syx%) from the references, including the current study, is presented in Table 3. When Syx% was not presented, RMSE% was presented. The discussion includes information on the taper models applied to each species, clone, age, and other factors from each reference.

When comparing the error ranges of the references, it was observed that when combining data from different species, materials, and production systems and obtaining relative residual standard errors (Syx%) ranging from 6.13% to 9.49% with the six models, the joint analysis of materials did not result in significant losses of precision. This places the study in the ninth position out of 17 references when compared to studies that analyzed specific materials from planted forests (Table 3).

Table 3 – Percentage residual standard error (*Syx%*) in the consulted references and the current work

Ranking	Autor(es)	Syx %	Itens
1	Lopes; Rode; Pauleto *et al*. (2018)	*2.02 – 3.15	3 models (African mahogany, 7 years old, 35 trees cubed with digital dendrometer)
2	Assis; Scolforo; Mello *et al.* (2002)	3.59 – 6.08	5th-degree polynomial (*Pinus taeda*)
3	Souza; Cosenza; Araújo *et al*. (2018)	3.93 – 6.14	4 models (3 genotypes of *Eucalytus* sp., 10 years old, 70 trees)
4	Andrade; Terra; Carvalho (2022)	4.02 – 4.78	16 models selected from 46 (*Corymbia citriodora*, 5.3 years old, 24 trees)
5	Andrade (2014)	4.29 – 6.98	18 models (*Eucalyptus. urophylla and Eucalyptus grandis,* 5-7 years old, 270 trees)
6	Soares; Flores; Cabacinha *et al.* (2011)	*4.35 - 7.41	ANN (*Eucalyptus* sp. 6,5 years old, 15 trees)
7	Souza; Silva; Xavier *et al.* (2008)	**6.02 – 7.07	3 segmented models (*Eucalytpus* sp. 16 years old, 41 trees)
8	Costa; Finger; Schneider *et al.* (2016)	6.80 - 7.50	4 models (*Araucaria angustifólia*, 85 trees cubed with a digital dendrometer)
9	**This work**	**6.13 - 9.49**	**6 models (miscellaneous materials, 237 trees)**
10	Machado; Urbano; Conceição *et al*. (2004)	6.15 – 9.99	3 models (*Pinus oocarpa*, 1100 trees)
11	Kohler; Koehler; Figueiredo filho *et al.* (2013)	7.57 – 8.45	2 models (*Pinus taeda* 11-23 years old, 120 trees)
12	Bernardi; Thiersch; Arteaga *et al.* (2021)	5.31 – 10.88	6 models (*Eucalyptus* sp. 6,3 years old, 163 trees)
13	Sandoval; Acuña (2022)	*7.7 – 8.50	ANN (3 *Nothofagus* sp. species analyzed separately, 1380 trees)
14	Yoshitani Junior; Nakajima; Arce *et al*. (2012)	7.92-11.59	3 models (*Pinus taeda*, 11-26 years old, 320 trees)
15	Mora; Silva; Gonçalves *et al.* (2014)	8.97 – 10.68	4 models (Hybrid *Eucalyptus. urophylla and Eucalyptus grandis*, 70 trees)
16	Sandoval; Acuña (2022)	*9.0 - 10.8	6 models (3 Nothofagus sp. species analyzed separately, 1,380 trees)
17	Queiroz; Machado; Figueiredo *et al.* (2008)	10.65 – 13.34	3 models (Bracatinga, 6-19 years old. 121 trees)

Source: Authors (2025)

In where: * RMSE; ** the error obtained by the Parresol model was excluded

4 DISCUSSIONS

Most studies on tree modeling aim to choose the best model for a dataset, as presented in the Results section of this study, where the Guimarães model and ANN 6 showed better performance. However, these analyses do not guarantee that these models can be recommended for other datasets or all trunk sections.

In references where taper models were applied to the same material under fixed spacing, age, and edaphic and climatic conditions, the conclusions were not unanimous, as other studies reported different models with better results. This suggests that there is a combination between the data and models, making it inconclusive to assert that model *A* is better than model *B* in all cases, regardless of whether it is non-segmented, segmented, form-exponent, "biomathematical," or ANN-based. In this study, we verified that unanimity did not occur in the diameter classes along the trunk (Figure 3) by jointly analyzing the materials listed in Table 1 jointly.

By disregarding the differences between materials and comparing the errors with those obtained in references that evaluated distinct materials, the hypothesis was tested that errors between the estimated and observed di between materials under fixed conditions would be smaller than the errors obtained by jointly analyzing materials under varied conditions. The studies presented in Table 3 are as follows:

Lopes; Rode; Pauleto *et al*. (2018) obtained the lowest errors in the ranking of African Mahogany in an agroforestry system by applying the models of Kozak; Munro; Smith *et al*. (1969), Demaerschalk, and Ormerod.

Assis; Scolforo; Mello *et al*. (2002), who applied the 5th-degree polynomial and integer and fractional power polynomials to 58 *Pinus taeda* trees, mostly obtained Syx values between 3.59% and 6.08%, ranking second. In contrast Kohler, Koehler; Figueiredo Filho *et al*. (2013) applied the same models to 120 *Pinus taeda* trees and ranked 11th with Syx values between 7.57% and 8.45%. Yoshitani Junior; Nakajima; Arce *et al*. (2012), also on *Pinus taeda*, ranked 14th (Syx = 7.92% to 11.59%).

Souza; Cosenza; Araújo *et al*. (2018), ranking third, evaluated four models in three *Eucalyptus* sp. genotypes compared to Bernardi; Thiersch; Arteaga *et al*. (2021) and Mora; Silva; Gonçalves *et al*. (2014), who also evaluated *Eucalyptus* sp. and ranked 12th and 15th, respectively. Bernardi; Thiersch; Arteaga *et al*. (2021) analyzed the Ormerod, Max-Burkhart, Muhairwe, and modified Methol models, Kozak 2004, and Schoepfer (or 5th degree-polynomial), in 163 *Eucalyptus* sp. trees at 6.3 years of age, obtaining Syx values between 5.31% and 10.88%. Mora; Silva; Gonçalves *et al*. (2014) cubed 70 trees of the urograndis hybrid at 8 years of age, with a spacing of 3 × 2 m, and applied the Baldwin, Demaerschalk, Kozak, and Ormerod models to obtain standard errors of estimate (Syx) that varied between 9.06% and 11.26% for a planted forest.

Andrade; Terra; Carvalho (2022) and Andrade (2014) evaluated 46 and 18 models, which ranked fourth and fifth, respectively. The first study, despite using the widest selection of models (46), cubed a small sample of 24 *Corymbia citriodora* trees, which may be related to the narrow error range. The study by Andrade (2014) focused on the hybrids of *Eucalyptus urophylla* and *Eucalyptus grandis* and found percentage standard errors between 4.3% and 6.98%.

In applications using neural networks, Soares; Flores; Cabacinha *et al*. (2011), ranking sixth, rigorously analyzed 615 trees of a eucalyptus clone at 6.5 years of age, built a multilayer network, and obtained RMSE values between 4.35% and 7.41% for di estimates. Sandoval; Acuña (2022) studied a *Nothofagus* species from cold regions and compared an ANN model, obtaining RMSE values between 7.7% and 10.1% (ranking 13th), with six taper models (Bruce 1 and 2, Demaerschalk, Biging, Lee, and Kozak's last model published in 2004), which produced RMSE values between 9% and 10.8% (ranking 16th).

Souza; Silva; Xavier (2008) analyzed cubing data from *Eucalyptus* sp. at a greater age (16 years) from seed propagation and applied the Max-Burkhart, Demaerschalk, and Kozak models to obtain Syx values ranging from 6.02% to 7.07%. Costa *et al*. (2016) obtained Syx values between 6.8% and 7.7%, ranking eighth with the Schoepfer(5th-degree polynomial), Kozak, Lee, and Sharma and Zhang models in 85 *Araucaria angustifolia* trees.

A study with a native species from the southern region of Brazil in commercial planting was conducted by Queiroz; Machado; Figueiredo *et al*. (2008), who cubed 121 *Mimosa scabrella* (Bracatinga) with ages ranging from 6 to 19 years and applied the Schoepfer (5th-degree polynomial), Hradetzky (fractional power polynomial), and Kozak models, obtaining Syx values between 10.65% and 13.34%; larger errors justified by being a material without genetic improvement (ranking 17th).

In the current study (ranking 9th), ANN 5 and 6 performed better than those reported by Sandoval; Acuña (2022). Although the same optimizer, Adam, and ReLU functions were used, the data were different, and there were other configurations for the network, such as the number of layers and neurons, which may have provided better efficiency, as well as in ANN 6, obtained using Tyberius software.

Comparing the error ranges in the literature, it was observed that when combining the data from different species, materials, and production systems, relative residual standard errors (Syx) ranged from 6.13% to 9.49% across the six models. The joint analysis of the materials did not result in significant losses of precision. This ranks the current study in the 9th position out of 17 references when compared to studies that analyzed specific materials from planted forests (Table 3). It is likely that there are variations in the form between trees of the same material under the same conditions that limit precision, even when using models that are more adherent to the analyzed profile. The hypothesis that errors will be systematically smaller in the taper modeling of distinct materials than in the joint analysis of materials from planted forests was not confirmed in this study.

5 CONCLUSIONS

Taper models fit different sections of the stem profile with varying levels of accuracy, and no single model consistently delivered the best performance across all diameter classes along the trunk.

Internal variations in tree form, fixing species, clone, age, spacing, edaphic and climatic conditions were not systematically smaller than the variation in the set of tree forms from different planted forests, species, clone, age, spacing, edaphic and climatic conditions.

ACKNOWLEDGMENTS

This study was supported by project No. 20.18.03.015.00.00 “ILPF Strategies for agricultural innovation in the Cerrado Mineiro region and surrounding areas - SisGerais.” We thank Walfrido Machado Albernaz and Emanuel da Silva Pinto Júnior (Emater) for their support in the cubing process in Maravilhas-MG, Marília Queiroz de Rezende for her support in the cubing of African Mahogany, Ricardo Vilela from Fazenda Bela Vista for his support in the cubing of Australian Cedar clones, Gustavo Pitangui de Salvo from Fazenda Lagoa dos Currais, and João Carlos Pereira from CJR for their support in the cubing of eucalyptus clones, Miguel Gontijo Neto for his cooperation in the cubing activities at the ILPF trial installed at Fazenda CNPMS, Lucas Ramos, who executed the cubing of *Pinus* sp. as a scholarship holder, Persival Ferreira Costa Junior from CALFLOR for his support in the cubing in Conselheiro Lafayete-MG, and Wendel Mendes Lima from SD Madeiras for his support in the rigorous cubing at Fazenda Cidade do Boi.

REFERENCES

ANDRADE, V. C. L. Modelos de taper do tipo expoente-forma para descrever o perfil do fuste de árvores. **Pesquisa Florestal Brasileira**, Colombo, v. 34, n. 80, p. 271-283, 2014. http://doi.org/10.4336/2014.pfb.34.80.614.

ANDRADE, V. C. L.; TERRA, D. L. C.; CARVALHO, S. P. C. Análise de regressão do perfil do fuste de *Corymbia citriodora* formado em área de Cerrado. **Ciência Florestal**, Santa Maria, v. 32, n. 3, p. 1500-1527, 2022. https://doi.org/10.5902/1980509866721.

ASSIS, A. A. L.; SCOLFORO, J. R. S.; MELLO, J. M.; OLIVEIRA, A. D. Avaliação de modelos polinomiais não-segmentados na estimativa de diâmetros e volumes comerciais de *Pinus taeda*. **Ciência Florestal**, Santa Maria, v. 12, n. 1, p. 89-107, 2002. https://doi.org/10.5902/198050981704.

BERNARDI, L. K.; THIERSCH, M. F. B. M.; ARTEAGA, A. J. M.; ALMEIDA, A. A. A.; PÁDUA, F. A.; THIERSCH, C. R. Diferentes modelos para o afilamento do tronco de *Eucalyptus* sp. para o cenário florestal brasileiro. **Ciência Florestal**, Santa Maria, v. 31, n. 3 p. 1364-1382, 2021. https://doi.org/10.5902/1980509840376.

COSTA, E. A.; FINGER, C. A. G.; SCHNEIDER, P. R.; HESS, A. F. Função de afilamento e sortimentos de madeira para *Araucaria angustifólia*. **Ciência Florestal**, Santa Maria, v. 26 n. 2, p. 523-533, 2016. https://doi.org/10.5902/1980509822753.

GUIMARÃES, D. P.; LEITE, H. G. Um novo modelo para descrever o perfil do tronco. **Revista Árvore**, Viçosa, MG, v. 16, n. 2, p. 170-180, 1992.

KOHLER, S. V.; KOEHLER, H. S.; FIGUEIREDO FILHO, A. Modelos de afilamento para *Pinus taeda* por classes de idade. **Floresta e Ambiente**, Seropédica, v. 20, n. 4, p. 470-479, 2013. http://dx.doi.org/10.4322/floram.2013.039.

KOZAK, A.; MUNRO, D. D.; SMITH, J. G. H. Taper functions and their applications in forest inventory. **Forest Chronicle**, v.45, n.4, p.278-283, 1969.

LOPES, L. S. S.; RODE, R.; PAULETTO, D.; BALONEQUE, D. D.; SILVA, A. R.; SANTOS, K. N. F. Ajuste de modelos de taper e sortimento de toras de Mogno africano em sistemas Agroflorestais em Belterra, Pará. **Agroecossistemas**, v. 10, n. 1, p. 18 - 28, 2018, ISSN online 2318-0188.

MACHADO, S. A.; URBANO, E.; CONCEIÇÃO, M. B.; FIGUEIREDO FILHO, A.; FIGUEIREDO, D. J. de. Comparação de modelos de afilamento do tronco para diferentes idades e regimes de desbaste em plantações de *Pinus oocarpa* Schiede. **Boletim de Pesquisa Florestal**, Colombo, n. 48, p. 41-64, 2004.

MÔRA, R.; SILVA, G. F.; GONÇALVES, F. G.; SOARES, C. P. B.; CHICHORRO, J. F.; CURTO, R. D. Análise de diferentes formas de ajuste de funções de afilamento. **Scientia Florestalis**, Piracicaba, v. 42, n. 102, p. 237-249, 2014.

QUEIROZ, D.; MACHADO, S. A.; FIGUEIREDO FILHO, A.; ARCE, J. E.; KOEHLER, H. S. Identidade de modelos em funções de afilamento para *Mimosa scabrella* Bentham em povoamentos nativos da região metropolitana de Curitiba/PR. **Floresta**, Curitiba, v. 38, n. 2, p. 339-349, 2008.

SAKICI, O. E.; OZDEMIR, G. Stem taper estimations with artificial neural networks for mixed Oriental beech and Kazdagi fir stands in Karabuk region, Turkey. **CERNE**, Lavras, v.24, p. 439-451, 2018.

SALEKIN, S.; CATALÁN, C. H.; BOCZNIEWICZ, D.; PHIRI, D.; MORGENROTH, J.; MEASON, D. F.; MASON, E. G. Global tree taper modelling: a review of applications, methods, functions, and their parameters. **Forests**, Basel, v. 12, n. 7, 913, 2021. https://doi.org/10.3390/f12070913.

SANDOVAL, S.; ACUÑA, E. Stem taper estimation using artificial neural networks for Nothofagus trees in natural forest. **Forests**, Basel, v. 13, n. 12, 2143, 2022. https://doi.org/10.3390/f13122143.

SEKI, M. Predicting stem taper using artificial neural network and regression models for Scots pine (*Pinus sylvestris* L.) in northwestern Türkiye. **Scandinavian Journal of Forest Research**, v. 38, n. 1–2, p. 97–104, 2023. https://doi.org/10.1080/02827581.2023.2189297.

SOARES, F. A. A. M. N.; FLÔRES, 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**, New York, v. 78, n. 1, p. 19-27, 2011. https://doi.org/10.1016/j.compag.2011.05.008.

SOCHA, J.; NETZEL, P.; CYWICKA, D. Stem taper approximation by artificial neural network and a regression set models. **Forests**, Basel, v. 11, n. 79, 2143, 2020.

SOUZA, C. A. M.; SILVA, G. F.; XAVIER, A. C.; CHICHORRO, J. F.; SOARES, C. P. B.; SOUZA, A. L. Avaliação de modelos de afilamento segmentados na estimação da altura e volume comercial de fustes de *Eucalyptus* sp. **Revista Árvore**, Viçosa, MG, v. 32, n. 3, p. 453-463, 2008. https://doi.org/10.1590/S0100-67622008000300008.

SOUZA, G. S. A.; COSENZA, D. N.; ARAÚJO, A. C. S. C.; PIMENTA, L. V. A.; SOUZA, R. B.; ALMEIDA, F. M.; LEITE, H. G. Evaluation of non-linear taper equations for predicting the diameter of eucalyptus trees. **Revista Árvore**, Viçosa, MG, v. 42, n. 1, e420102, 2018. https://doi.org/10.1590/1806-90882018000100002.

YOSHITANI JUNIOR, M.; NAKAJIMA, N. Y.; ARCE, J. E.; MACHADO, S. A.; DRUSZCZ, J. P.; HOSOKAWA, R. T.; MELLO, A. A. Funções de afilamento para plantios desbastados de *Pinus taeda*. **Floresta**, Curitiba, PR, v. 42, n. 1, p. 169-176, 2012.

Authorship Contribution

1 Thomaz Corrêa e Castro da Costa

Forestry Engineering, PhD. in Forestry Science

https://orcid.org/0000-0002-7602-6042 • thomaz.costa@embrapa.br

Contribution: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Resources; Validation; Visualization; Writing – original draft; Writing – review & editing

2 Henrique Coelho Mendes

Graduated in Computer Science, employee at Banco Inter

https://orcid.org/0009-0000-1200-7312 • h8mendes@gmail.com

Contribution: Investigation; Methodology; Writing – review & editing

How to quote this article

COSTA, T. C. C.; MENDES, H. C. Accuracy of tree profile estimation in planted forests. **Ciência Florestal**, Santa Maria, v. 35, e90612, p. 1-22, 2025. DOI 10.5902/1980509890612. Available from: https://doi.org/10.5902/1980509890612. Accessed in: day month abbr. year.

Data Availability Statement:

Datasets related to this article will be available upon request to the corresponding author.