REVISTA FLORESTA Engenharia Florestal
ISSN 1982-4688
Edição 56
Universidade Federal do Paraná
Setor de Ciências Agrárias

ARTIFICIAL NEURAL NETWORK-BASED CORRECTION OF LOG-SCALING MEASUREMENTS COLLECTED WITH OPTICAL DENDROMETER

Ingrid Raphaela Cromwell Pereira[1], Qüinny Soares Rocha[1;2*], Marina Mell Campos Bastos[1], Raylon Pereira Maciel[3], Lina Bufalino[1], Rodrigo Geroni Mendes Nascimento[1]

[1]Universidade Federal Rural da Amazônia, Instituto de Ciências Agrárias, Belém, Pará, Brasil - raphaelacromwell@gmail.com ; marinamell2002@gmail.com ; lina.bufalino@ufra.edu.br ; rodrigo.geroni@ufra.edu.br
[2]Universidade Estadual da Região Tocantina do Maranhão, Centro de Ciências Agrárias, Imperatriz, Maranhão, Brasil – *quinny.rocha@uemasul.edu.br
[3]Universidade Federal Rural da Amazônia, Campus Parauapebas, Parauapebas, Pará, Brasil - raylonmaciel@gmail.com

Received for publication: 11/12/2025 – Accepted for publication: 23/02/2026

Resumo

Correção de medições de cubagem obtidas com um dendrômetro óptico por meio de redes neurais artificiais. O objetivo deste estudo foi corrigir a cubagem de árvores em pé por meio de redes neurais artificiais (RNAs). Os dados foram obtidos em dois inventários florestais realizados em plantios de Tectona grandis (13,8 ha) e de *Schizolobium parahyba* (46 ha), ambos localizados no estado do Pará, Brasil. O volume das árvores foi mensurado com o Criterion RD 1000, pelo método de Smalian, com diâmetros coletados nas alturas de 0,1 m, 0,7 m, 1,3 m e, posteriormente, a cada 2 m até o primeiro galho vivo. As alturas comercial e total, o diâmetro à altura do peito (DAP) medido com o Criterion, a distância do operador até a árvore e a espécie foram utilizadas como variáveis de entrada, enquanto o DAP medido com fita métrica foi empregado como variável de saída. No total, 84 redes *Multilayer Perceptron* foram testadas, variando o número de neurônios nas camadas ocultas e as funções de ativação. As melhores redes foram selecionadas com base no coeficiente de determinação e no erro percentual médio. No geral, 59% das redes convergiram antes de atingir o número máximo de iterações, e a função tangente hiperbólica convergiu em 74% dos testes. As RNAs foram eficazes na correção dos erros não amostrais observados nas medições realizadas com o Criterion. As redes com apenas uma camada oculta apresentaram melhor desempenho, e a função de ativação logística resultou em ajustes mais precisos. Foi identificada evidência de *overfitting* em uma das redes de melhor desempenho, enquanto o *underfitting* ocorreu nos modelos com menor número de iterações.
Palavras-chave: Mensuração Florestal, Afilamento do Fuste, *Multilayer Perceptrons*, *Resilient Propagation*, aprendizado de máquina.

Abstract

The objective of this study was to correct standing-tree volume measurements using artificial neural networks (ANNs). Data were obtained from two forest inventories conducted in *Tectona grandis* (13.8 ha) and *Schizolobium parahyba* (46 ha) plantations in Pará, Brazil. Tree volume was measured with the Criterion RD 1000 using the Smalian method, with diameters collected at 0.1 m, 0.7 m, 1.3 m, and every 2 m up to the first live branch. Commercial and total heights, DBH measured with the Criterion, operator-to-tree distance, and species were used as input variables, whereas DBH measured with a tape was used as the output variable. A total of 84 Multilayer Perceptron networks were tested, varying the number of hidden-layer neurons and activation functions. The best-performing networks were selected based on the correlation coefficient and the percent root-mean-square error. Overall, 59% of the networks converged before reaching the maximum number of iterations, and the hyperbolic tangent function converged in 74% of the tests. The ANNs effectively corrected non-sampling errors present in the Criterion measurements. Networks with a single hidden layer performed best, and the logistic activation function yielded superior fits. Evidence of overfitting was identified in one of the top-ranked networks, while underfitting occurred in models with fewer iterations.
Keywords: Forest Mensuration, Tapering, *Multilayer Perceptrons*, *Resilient Propagation*, Machine Learning.

INTRODUCTION

Forest inventory is a systematic procedure used to describe, both qualitatively and quantitatively, the species composition and structural attributes of a forest area. It serves several purposes, including characterizing forest structure, assessing growth, and estimating timber stock (AMORIM, 2020). Among the variables typically measured, tree volume is among the most important for production planning, as it provides a basis for estimating a stand's productive potential and is essential for sustainable forest management. The individual volume of trees serves as the starting point for evaluating timber stock within a forest stand and supports decisions on silvicultural treatments, harvesting operations, and transportation. Consequently, accurate methods for determining tree volume are fundamental to ensuring reliable forest inventory results (DANTAS *et al*., 2020).

Among the available methods for estimating tree volume, rigorous sectional volume measurement is one of the most widely used and well-established procedures. It involves dividing the stem into sections of fixed or

variable lengths and calculating the volume of each segment using appropriate geometric formulas. The sum of these segment volumes yields the total tree volume. In general, this method is performed on felled trees, making it a destructive, costly, and time-consuming procedure (NICOLETTI *et al*., 2016).

The felling of trees in a stand can compromise the system's overall functionality. Moreover, it may be impractical in plantations of high-value commercial species, where premature cutting does not guarantee allocation to higher-value end uses with greater economic return (CURTO *et al*., 2019). With technological advancements, instruments capable of measuring diameters and heights without removing trees have been developed. As a result, accurate volume estimation can be performed using non-destructive dendrometric equipment. Among these devices, the Criterion RD 1000® stands out for operating on trigonometric principles and using a laser gauge in conjunction with a graduated bar to measure stem diameters along the trunk. The instrument provides a measurement range of up to 254 cm and an accuracy of 6 mm at distances up to 24 m, enabling reliable volume estimation (BIAZATTI *et al*., 2020).

When estimating the volume of standing trees using an optical dendrometer, both sampling and non-sampling errors may occur. Sampling errors are inherent to the sampling design and arise from natural population variability when only a subset of the forest is measured, leading to differences between the true population parameters and their sample-based estimates. Non-sampling errors, in turn, originate from factors related to data acquisition and processing, including reading inaccuracies, operator-to-tree misalignment, instrumental limitations, visibility constraints, and operational inconsistencies. These errors can compromise measurement accuracy and reduce the reliability of derived estimates (DUARTE *et al*., 2016).

One way to minimize non-sampling errors is to apply artificial intelligence techniques. Artificial neural networks (ANNs) are a type of machine learning model inspired by the operation of biological neurons, in which each computational unit processes the incoming information and transmits it to other interconnected units (JHA *et al*., 2023). The main advantage of ANNs lies in their ability to iteratively learn complex patterns and reproduce them with high accuracy across previously unseen datasets, supporting both the identification of underlying phenomena and the evaluation of alternative scenarios (BUENO *et al*., 2020). In forestry science, although research remains relatively recent, ANNs have already been applied to the estimation of volume, height, taper, growth, deforestation, and other biometric variables. Consequently, their use has become increasingly relevant in the forestry domain due to consistent results across different applications (LEAL *et al*., 2020).

Rather than treating artificial neural networks solely as predictive tools, they can also be interpreted as data-driven calibration models that compensate for the systematic measurement biases inherent to indirect dendrometric methods. As highlighted by Dongare et al. (2012), ANNs are particularly suitable when the relationships among variables are complex or poorly understood. In this context, the capacity of Multilayer Perceptron networks to learn nonlinear functional relationships (DREW et al., 2000; MELO; ROCHA, 2023) allows them to operate as transformation functions between measurement systems. Thus, ANN-based approaches can convert optical dendrometer measurements into calibrated estimates aligned with reference measurements, expanding their applicability in forest inventory practices.

Given the possibility of applying artificial intelligence techniques to mitigate non-sampling errors in data acquisition with the Criterion optical dendrometer, the hypotheses guiding this study were: (i) artificial neural networks are capable of correcting non-sampling errors in sectional diameter measurements; and (ii) identifying which ANN configuration yields the best correction performance. Therefore, the objective of this study was to correct diameter data obtained from sectional measurements using artificial neural networks and to systematically model and correct non-sampling errors in stem diameter measurements collected with the Criterion through Multilayer Perceptron (MLP) neural networks.

MATERIAL AND METHODS

Areas of study

The work was conducted in two areas: Area I, a plantation of *Schizolobium parahyba* var. *amazonicum* (Huber ex Ducke) (paricá), and Area II, a plantation of *Tectona grandis* L. f. (teak). The paricá plantation is intercropped with beef cattle and occupies 46 hectares of the Cinco Águas Farm, located in the municipality of Abel Figueiredo, Pará, Brazil (Figure 1a). According to the Köppen–Geiger classification, the local climate is type Am (tropical humid monsoon), with a mean temperature of 26.35 °C. Relative humidity ranges from 52% to 100%, and annual precipitation varies from 1,062.6 mm to 3,183.1 mm. The terrain ranges from flat to gently undulating and is predominantly composed of Yellow Latosols and Yellow Argisols with medium to very clayey textures. The dominant vegetation consists of Subevergreen Equatorial Forest and Hygrophilous Floodplain Equatorial Forest (NASCIMENTO *et al*., 2022).

REVISTA FLORESTA Engenharia Florestal
ISSN 1982-4688
Edição 56
Universidade Federal do Paraná
Setor de Ciências Agrárias

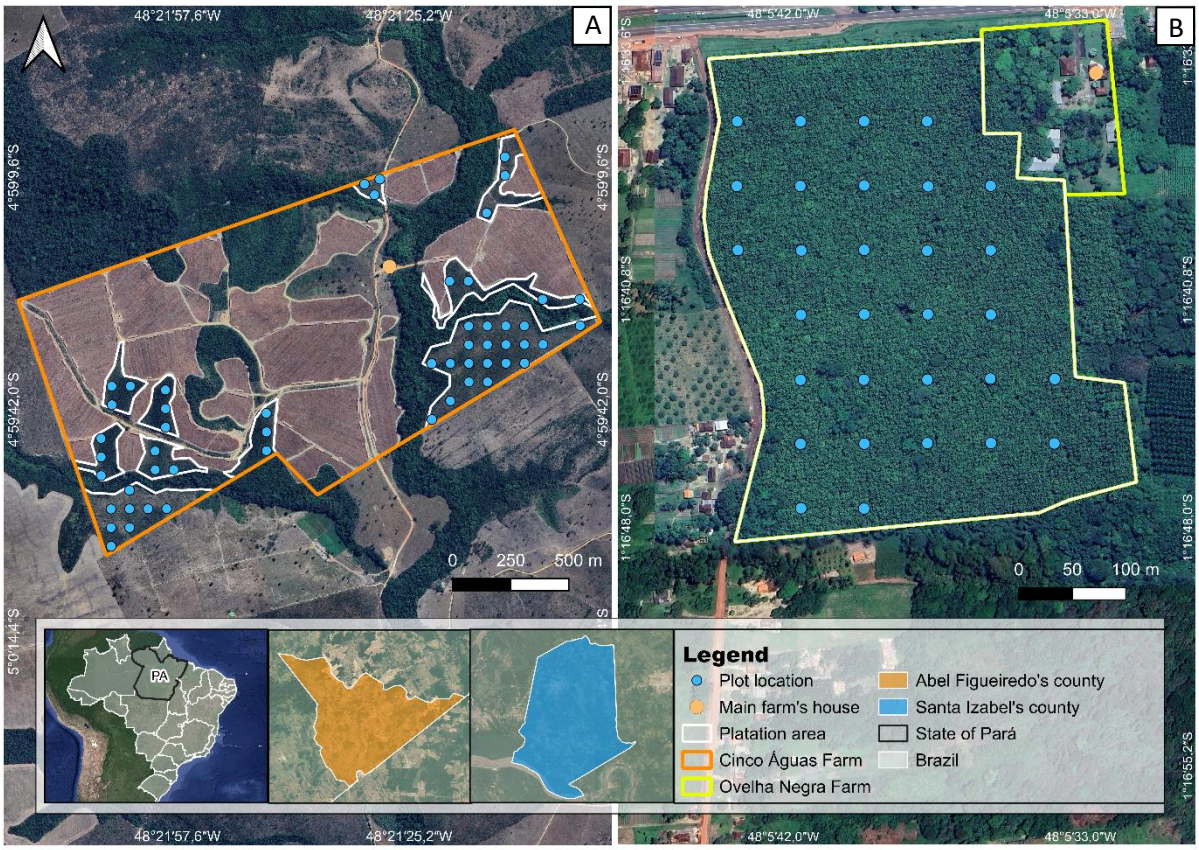

Figure 1. A - Location of Study Area II at Cinco Águas Farm, Abel Figueiredo, Pará State, Brazil. B - Location of Study Area I at Ovelha Negra Farm, Santa Izabel do Pará, Brazil.
Figura 1. A - Localização da área de estudo II na Fazenda Cinco Águas, em Abel Figueiredo, Pará, Brasil. B - Localização da área de estudo I na Fazenda Ovelha Negra, em Santa Izabel do Pará, Brasil.
Information System - Datum: SIRGAS 2000; Database: IBGE 2024.

The teak plantation, covering 13.8 ha, is located at Ovelha Negra's Farm (Figure 1b), in the municipality of Santa Izabel do Pará, Brazil. The local vegetation consists mainly of subevergreen equatorial upland forest, and the predominant soil type is dystrophic Yellow Argisol, occurring on predominantly flat terrain. The region's climate is classified as Af (humid tropical) by the Köppen–Geiger system, characterized by year-round rainfall. The mean monthly temperature is approximately 25 °C, with October and November being the warmest months, and the average annual precipitation reaches 2,350 mm, with relative humidity around 85% (NASCIMENTO *et al.*, 2022).

Data collection

The forest inventory plots in both study areas were systematically established as circular plots of 500 m², spaced 80 m apart. Thirty plots were installed in the teak plantation and 53 in the paricá plantation. Considering the magnetic north–south orientation of the plots, the two trees to the right of the central point were selected for measurement of sectional diameter, totaling 60 measured trees in the teak plantation and 106 in the paricá plantation. All measurements were obtained using the Criterion RD 1000 optical dendrometer, manufactured by Laser Technology, Inc.

In the sectional diameter measurement procedure, the operator stood at a horizontal distance of 10 m or 15 m, selected to approximate the commercial height of the trees in the two study areas, using a measuring tape. Diameter measurements were obtained using the Smalian method at heights of 0.1 m, 0.7 m, and 1.3 m; from this height onward, diameters were measured every 2 m until the insertion of the first live branch. The commercial and total tree heights were also recorded. To assess potential measurement errors with the Criterion, the DBH (diameter at breast height) was also measured with a tape measure.

ANN configuration

Eighty-four Multilayer Perceptron ANN models were tested because this architecture offers greater capacity to model nonlinear relationships than other neural network architectures (MELO; ROCHA, 2023). These networks use hidden layers that capture data nonlinearity. During ANN training, several parameters were kept

fixed: a training error threshold of 0.01, a maximum of 10 million iterations, the Resilient Propagation algorithm with weight backtracking (rprop+), and an input layer composed of DBH measured with the Criterion, operator-to-tree distance, total height, commercial height, and species. Because species is a categorical variable, it was encoded as a binary variable. The output layer variable was the DBH measured with a diameter tape.

The ANN structures varied in the number of hidden layers, the number of neurons in each layer, and the activation function. One- and two-layer architectures were tested, with the number of neurons per layer ranging from two to thirteen. Although no strict rule exists for determining the ideal number of neurons in a hidden layer, these values were chosen based on the recommendation that the maximum number should be twice the number of input variables plus one (RAFIQ *et al.*, 2001). The activation functions evaluated were the logistic and hyperbolic tangent functions. The dataset was randomly split into 50% for training and 50% for validation, using supervised learning (CHAMMA *et al*., 2021).

The selection of the best ANN configuration was based on two evaluation metrics, namely, the Pearson correlation coefficient ($r_{y\hat{y}} = \frac{\sum(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})}{\sqrt{(y_i-\bar{y})^2}\sqrt{(\hat{y}_i-\bar{\hat{y}})^2}}$) and the percentage Root Mean Squared Error ($RMSE_{\%} = \left(\frac{\sqrt{\frac{1}{n}\sum(y_i-\hat{y}_i)^2}}{\bar{y}}\right).100$), where y_i represents observed values and $\hat{y}_i$ predicted values. Computed for both the training and validation datasets, following the methodology of Reis *et al*. (2018) and Araújo *et al*. (2020). To facilitate identification of the best-performing ANNs, a ranking system was created in which each metric was ranked from best to worst, with scores ranging from 1 to 84. The 10 ANNs with the lowest cumulative scores were selected to simulate the correction of the sectional diameter data. Accordingly, the diameter values collected with the Criterion were corrected using each selected ANN. The simulation results were presented as scatter plots, along with stem-and-leaf profiles of the observed and corrected data, to enable a detailed assessment of ANN performance. ANN training was conducted in R software, version 4.1.0 (R Core Team, 2023), using the neuralnet package (FRITSCH *et al*., 2019), and the graphical outputs were produced using the ggplot2 package (WICKHAM *et al*., 2019).

ANN-based diameter correction procedure

The ANN-based correction procedure was trained to learn the functional relationship between diameter measurements obtained with the optical dendrometer (Criterion) and reference measurements obtained with a diameter tape at DBH. Because DBH was measured using both instruments, these paired observations were used to train the ANN to map Criterion-based measurements to reference-equivalent values. Additional independent variables (operator-to-tree distance, species, total height, and commercial height) were included to account for variability associated with measurement conditions and tree characteristics, allowing the model to capture systematic differences between measurement methods.

Once this mapping function was established, the ANN was applied to the sectional diameters measured exclusively with the Criterion along the stem. Although only DBH measurements were simultaneously obtained with both instruments, the learned relationship enabled the estimation of reference-equivalent diameters at other stem positions. This approach assumes that the relationship between the two measurement systems reflects systematic measurement bias rather than geometric variation in stem form. Therefore, the ANN acts as a calibration tool, reducing non-sampling errors associated with indirect optical measurements.

RESULTS

The performance metrics presented in Table 1 indicate that the ANN-based correction substantially improved the agreement between Criterion-derived measurements and reference values. The best-performing models achieved Pearson correlation coefficients above 0.84 in validation and reduced RMSE values to approximately 15%, demonstrating a meaningful reduction in measurement discrepancies. The differences between direct DBH measurements obtained with a diameter tape and indirect measurements collected using the Criterion optical dendrometer are mainly related to the indirect geometric measurement principle of the equipment. The Criterion estimates diameter based on distance and angular measurements, which may introduce deviations associated with operator positioning, measurement distance, stem irregularities, and visual alignment. In the present study, these differences were reflected in systematic deviations between observed and reference diameters, particularly along specific portions of the stem profile (Figure 2). The application of ANN-based correction reduced these discrepancies, thereby improving agreement between optical measurements and reference DBH values, as demonstrated by the correlation and RMSE metrics (Reis et al., 2018; Araújo et al., 2020).

Table 1. Ranking of the 84 artificial neural network configurations for DBH correction evaluated using Pearson's correlation coefficient and RMSE%.

Tabela 1. Ordenamento das 84 configurações de redes neurais artificiais para correção do DAP, avaliadas por meio do coeficiente de correlação de Pearson e do RMSE%.

ANN	Hidden	Function	Time	ANN error	Function error	Iterations	Adjustment		Validation		Pt.
							$r_{y\hat{y}}$	RMSE%	$r_{y\hat{y}}$	RMSE%	
1	2	log	02:06	80.1757	0.0112	1,255,044	0.9459	9.68	0.8107	16.53	72
2*	3	log	00:03	78.9452	0.0108	56,322	0.9468	9.60	0.8269	15.53	52
3*	4	log	00:03	79.1583	0.0115	53,880	0.9466	9.62	0.8492	15.13	41
4*	5	log	12:19	62.0933	0.0100	5,549,918	0.9584	8.52	0.8194	15.95	56
5	6	log	03:34	22.3034	0.0100	1,625,578	0.9853	5.11	0.7514	21.69	63
6*	7	log	02:25	18.3406	0.0100	1,083,833	0.9879	4.63	0.7687	16.21	45
7*	8	log	03:07	38.2697	0.0106	1,301,707	0.9746	6.69	0.7951	17.87	59
8*	9	log	12:04	7.0135	0.0100	4,475,671	0.9954	2.86	0.2321	16.59	59
9	10	log	12:22	10.8856	0.0103	4,647,901	0.9928	3.57	0.6781	24.69	67
10*	11	log	13:51	2.8915	0.0104	4,855,696	0.9981	1.84	0.7116	22.58	54
11	12	log	05:04	5.3877	0.0103	1,617,239	0.9965	2.51	0.3266	21.12	62
12*	13	log	09:53	3.1150	0.0105	3,966,681	0.9980	1.91	0.6955	24.07	59
13	2 - 2	log	00:01	761.7258	0.0100	80	0.0000	29.84	0.0000	28.34	137
14	3 - 3	log	15:12	95.9199	0.0104	6,497,193	0.9349	10.59	0.8268	15.98	76
15	4 - 4	log	-	-	0.1511	10,000,000	-	-	-	-	145
16	5 - 5	log	-	-	0.1621	10,000,000	-	-	-	-	145
17	6 - 6	log	-	-	0.1558	10,000,000	-	-	-	-	145
18	2 - 3	log	14:02	95.9189	0.0105	6,709,598	0.9349	10.59	0.8268	15.98	76
19	2 - 4	log	13:05	94.4223	0.0106	4,779,330	0.9360	10.50	0.0813	15.81	92
20	2 - 5	log	-	-	0.0697	10,000,000	-	-	-	-	145
21	2 - 6	log	-	-	0.0451	10,000,000	-	-	-	-	145
22	2 - 7	log	-	-	0.0972	10,000,000	-	-	-	-	145
23	2 - 8	log	-	-	0.3889	10,000,000	-	-	-	-	145
24	2 - 9	log	-	-	0.3109	10,000,000	-	-	-	-	145
25	2 - 10	log	-	-	0.0862	10,000,000	-	-	-	-	145
26	2 - 11	log	00:01	761.7258	0.0100	53	0.0000	29.84	0.0000	28.34	137
27	3 - 4	log	-	-	0.0326	10,000,000	-	-	-	-	145
28	3 - 5	log	-	-	0.1405	10,000,000	-	-	-	-	145
29	3 - 6	log	-	-	0.0571	10,000,000	-	-	-	-	145
30	3 - 7	log	-	-	0.0973	10,000,000	-	-	-	-	145
31	3 - 8	log	-	-	0.2253	10,000,000	-	-	-	-	145
32	3 - 9	log	-	-	0.0997	10,000,000	-	-	-	-	145
33	3 - 10	log	-	-	0.3273	10,000,000	-	-	-	-	145
34	4 - 5	log	10:39	20.6671	0.0100	3,850,345	0.9863	4.91	0.4898	29.39	77
35	4 - 6	log	-	-	0.1084	10,000,000	-	-	-	-	145
36	4 - 7	log	-	-	0.0792	10,000,000	-	-	-	-	145
37	4 - 8	log	-	-	0.0803	10,000,000	-	-	-	-	145
38	4 - 9	log	-	-	0.2015	10,000,000	-	-	-	-	145
39	5 - 6	log	-	-	0.1565	10,000,000	-	-	-	-	145
40	5 - 7	log	-	-	0.1715	10,000,000	-	-	-	-	145
41	5 - 8	log	-	-	0.0635	10,000,000	-	-	-	-	145
42	6 - 7	log	-	-	0.0756	10,000,000	-	-	-	-	145
43	2	tan	00:01	761.7258	0.0100	80	0.0000	29.84	0.0000	28.34	137
44	3	tan	08:26	95.8519	0.0102	97,718	0.9350	10.58	0.8391	15.49	60
45	4	tan	17:51	105.4484	0.0110	9,376,072	0.9282	11.10	0.8449	15.05	67
46	5	tan	16:13	105.4480	0.0110	9,747,114	0.9282	11.10	0.8449	15.05	67
47	6	tan	8:43	96.5721	0.0101	4,818,898	0.9345	10.62	0.8411	15.49	63
48*	7	tan	00:26	92.1136	0.0105	188,638	0.9376	10.38	0.8515	14.91	46
49	8	tan	09:43	35.5843	0.0107	3,707,999	0.9764	6.45	0.7165	20.62	67

ANN	Hidden	Function	Time	ANN error	Function error	Iterations	Adjustment		Validation		Pt.
							$r_{y\hat{y}}$	RMSE%	$r_{y\hat{y}}$	RMSE%	
50*	9	tan	00:10	75.9705	0.0103	79,598	0.9488	9.42	0.8378	15.31	42
51	10	tan	8:57	79.7064	0.0121	3,095,779	0.9462	9.65	0.6083	41.75	101
52	11	tan	00:41	48.8384	0.0100	249,269	0.9674	7.55	0.7201	25.65	81
53	12	tan	03:58	47.2807	0.0104	1,235,374	0.9685	7.43	0.7831	19.11	64
54	13	tan	03:09	47.3433	0.0100	851,810	0.9684	7.44	0.7267	21.56	73
55	2 - 2	tan	11:02	101.0425	0.0100	5,557,858	0.9313	10.87	0.8337	15.60	76
56	3 - 3	tan	04:23	101.0469	0.0113	1,531,435	0.9313	10.87	0.8338	15.60	75
57	4 - 4	tan	01:30	100.5505	0.0105	522,292	0.9317	10.84	0.8306	16.35	81
58	5 - 5	tan	-	-	0.1335	10,000,000	-	-	-	-	145
59	6 - 6	tan	-	-	0.1538	10,000,000	-	-	-	-	145
60	2 - 3	tan	05:51	96.8017	0.0107	2,943,899	0.9343	10.64	0.8284	15.89	76
61	2 - 4	tan	07:36	101.0415	0.0104	3,930,086	0.9313	10.87	0.8337	15.60	76
62	2 - 5	tan	14:57	94.1450	0.0102	5,375,495	0.9362	10.49	0.8225	16.51	74
63	2 - 6	tan	06:14	101.0418	0.0101	2,265,346	0.9313	10.87	0.8337	15.60	76
64	2 - 7	tan	10:07	96.8001	0.0103	4,288,754	0.9343	10.64	0.8284	15.89	76
65	2 - 8	tan	01:22	373.9183	0.0220	610,829	0.7135	20.90	0.7098	19.79	113
66	2 - 9	tan	00:01	761.7258	0.0100	45	0.0000	29.84	0.0000	28.34	137
67	2 - 10	tan	10:01	101.0417	0.0124	3,982,268	0.9313	10.87	0.8337	15.60	76
68	2 - 11	tan	06:50	29.3671	0.0100	2,551,572	0.9805	5.86	0.4436	42.34	84
69	3 - 4	tan	12:25	82.7778	0.0103	5,502,890	0.9441	9.84	0.8091	18.92	79
70	3 - 5	tan	05:48	101.0420	0.0101	2,621,349	0.9313	10.87	0.8337	15.60	76
71	3 - 6	tan	09:37	96.8002	0.0110	3,558,488	0.9343	10.64	0.8284	15.89	76
72	3 - 7	tan	00:28	100.5509	0.0101	271,322	0.9317	10.84	0.8344	15.89	74
73	3 - 8	tan	00:01	761.7258	0.0100	39	0.0000	28.84	0.0000	28.34	135
74	3 - 9	tan	-	-	1.3874	10,000,000	-	-	-	-	145
75	3 - 10	tan	-	-	0.0736	10,000,000	-	-	-	-	145
76	4 - 5	tan	-	-	0.0979	10,000,000	-	-	-	-	145
77	4 - 6	tan	-	-	0.0252	10,000,000	-	-	-	-	145
78	4 - 7	tan	18:55	94.13724	0.0105	7,277,868	0.8764	10.49	0.8224	16.55	86
79	4 - 8	tan	-	-	0.4903	10,000,000	-	-	-	-	145
80	4 - 9	tan	00:12	744.6578	0.0100	139,966	0.1497	29.50	0.1354	28.83	134
81	5 - 6	tan	-	-	0.0432	10,000,000	-	-	-	-	145
82	5 - 7	tan	-	-	0.2307	10,000,000	-	-	-	-	145
83	5 - 8	tan	-	-	0.0914	10,000,000	-	-	-	-	145
84	6 - 7	tan	-	-	0.9428	10,000,000	-	-	-	-	145

Legend: * ten best ANNs; Hidden, number of hidden layers and number of neurons in each layer; log, logistic activation function; tan, hyperbolic tangent activation function; rŷy, Pearson correlation coefficient; RMSE%, Root Mean Squared Error percent; Pt., ANN ranking score.
Evaluation metrics were calculated based on the agreement between ANN-predicted DBH values and reference DBH measurements obtained with a diameter tape.

The best scores were obtained by ANNs 3, 50, 6, 48, 2, 10, 4, 7, 8, and 12, with ANNs 7, 8, and 12 presenting identical rankings (Table 1). All selected configurations contained a single hidden layer, and eight of the ten best-performing ANNs used the logistic activation function. More than half of all models (59%) converged before reaching the maximum number of iterations. Across the full set of evaluated ANNs, the hyperbolic tangent function trained 74% of the networks, whereas the logistic function trained 43%. The training error threshold of 0.01 was not a reliable performance indicator, as illustrated by ANNs 13, 26, 43, 66, 73, and 80. In contrast, the final ANN error was more informative, as four of the ten best configurations had some of the smallest final errors.

Most of the more complex ANNs, those with two hidden layers and a larger number of neurons, did not reach the minimum error threshold, particularly when using the logistic activation function. ANNs trained with 50,000 to 100,000 iterations yielded the best evaluation metrics, as observed for ANNs 2, 3, 44, and 50. Conversely, increasing the number of iterations beyond this range did not improve performance, as shown by ANNs 14, 18, 45, 46, and 78. The lowest iteration counts produced the poorest results, as seen in ANNs 13, 26, 43, 66, and 73. ANNs 8, 9, 10, 11, and 12 achieved the best results on the training data; however, only ANN 10 also performed well on the validation dataset, whereas ANNs 8 and 11 ranked among the worst.

In the simulation of the corrected sectional diameter data using the ten best ANNs (Figure 2), the curves produced by ANNs 3 and 5, the two top-ranked models, were highly similar. For *Schizolobium parahyba*, the observed and simulated curves followed the same trend between 0.4 and 1.0 hi/ht, whereas for *Tectona grandis*, this similarity occurred between 0.0 and 0.3 hi/ht. Overall, except for ANN 8, the 10 best networks produced comparable patterns between the observed and ANN-corrected values. The teak dataset showed stronger agreement among the ANN-generated curves, while the paricá dataset exhibited greater variation in the patterns produced by different networks.

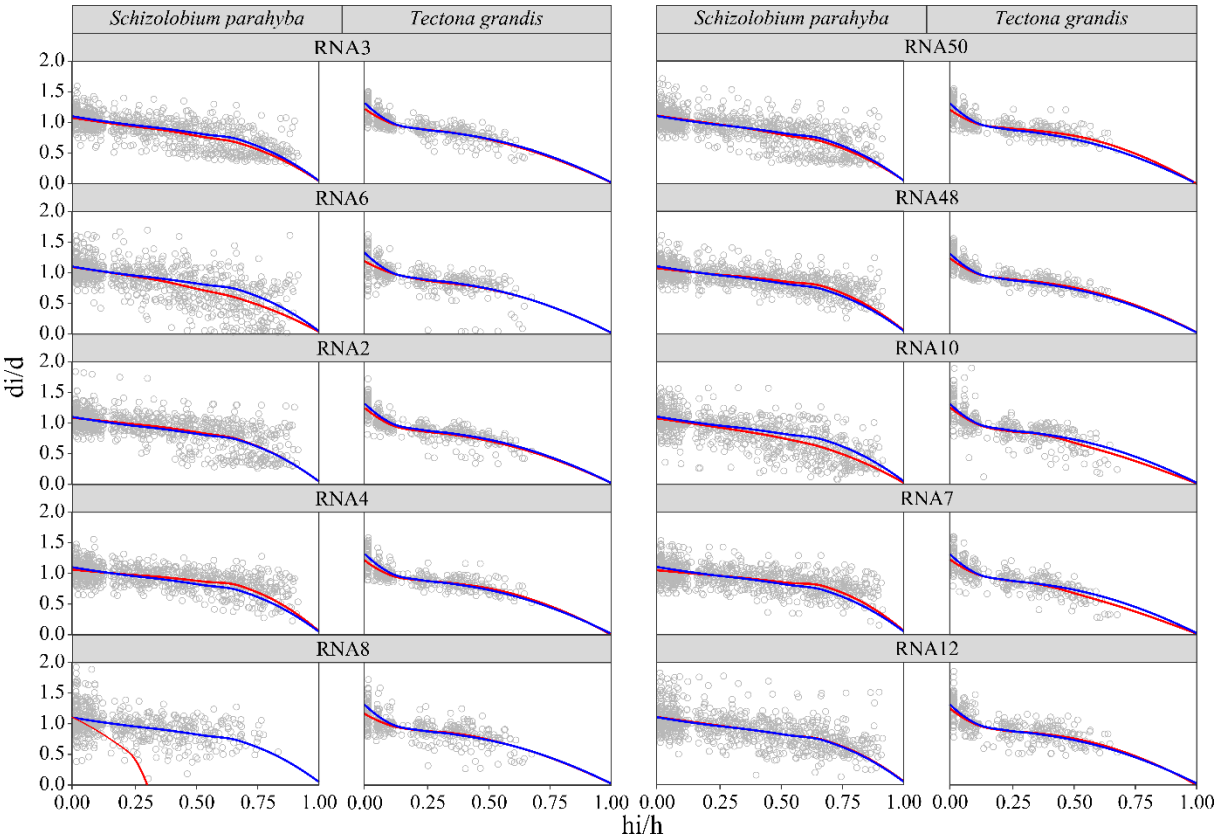

Figure 2. Observed and ANN-corrected taper curves produced by the 10 best artificial neural networks, shown separately for *Tectona grandis* L. f. and *Schizolobium parahyba* var. *amazonicum* (Huber ex Ducke) stands in Pará State, Brazil. Gray points represent the corrected log-scaling data, while the blue line corresponds to the observed taper profile and the red line to the ANN-corrected profile.

Figura 2. Perfis de afilamento observados e corrigidos pelas 10 melhores redes neurais artificiais, apresentados separadamente para plantios de *Tectona grandis* L. f. e *Schizolobium parahyba var. amazonicum* (Huber ex Ducke) no estado do Pará, Brasil. Os pontos em cinza representam os dados cubados corrigidos, a linha azul corresponde ao perfil observado e a linha vermelha ao perfil corrigido pelas RNAs.

The graphical comparison presented in Figure 2 indicates that the ANN-based correction did not merely smooth the data but systematically adjusted taper profiles, thereby reducing measurement bias in specific stem regions. The convergence of corrected curves toward the central tendency of the diameter distribution suggests that the neural networks effectively learned the relationship between Criterion-based measurements and reference diameters, enabling biologically consistent correction across the stem profile.

DISCUSSION

The results suggest that the ANN models operated not merely as predictive algorithms but as calibration functions capable of compensating for systematic measurement deviations associated with indirect optical instruments. As described by Wang (2003), the nonlinear activation functions enable neural networks to model complex transformation relationships, while the adaptive learning process (KROGH, 2008) allows the network to adjust internal weights to minimize systematic discrepancies. Therefore, the ANN effectively learned a transfer function between measurement systems, consistent with the distributed processing characteristics described by Wu and Feng (2018), enabling the propagation of calibration from DBH to sectional diameters along the stem.

According to Dongare *et al*. (2012), ANNs are an appropriate choice when large datasets are available, and the underlying problem is not well understood enough to allow the derivation of an explicit analytical model, which justifies their application in correcting sectional diameter data obtained with the Criterion optical dendrometer. These authors also emphasize that although the researcher defines the ANN architecture, the system automatically adjusts its internal parameters. Due to this stochasticity, it may not be easy to refine the solution incrementally when the model does not converge adequately. Nevertheless, this limitation is offset by the high efficiency of ANN-based solutions, which often achieve performance difficult to match by other technological approaches.

Multilayer Perceptrons ANNs models learn through iterative optimization. During training, the network adjusts the connection weights, assigning larger weights to features that contribute more to minimizing error and naturally reducing the influence of those that contribute less. This architecture also exhibits generalization ability (DREW *et al*., 2000). These characteristics may explain why some ANNs converged quickly during training yet still produced unsatisfactory results, as observed for ANNs 26, 66, and 73.

The resilient propagation algorithm with weight backtracking trains ANNs by iteratively updating connection weights based on the sign of the gradient of the activation function, with all weights initialized randomly. During training, a local error value is computed at each iteration for each hidden-layer neuron, and the corresponding weight updates are applied to progressively reduce the error until stabilization or the maximum number of iterations is reached. The sum of these local errors across all training iterations represents the total ANN error; therefore, smaller values indicate better overall performance (KROGH, 2008). This explains why the per-iteration function error was not a reliable indicator of the best configuration, as it reflects only a partial error calculated independently for each neuron at each iteration.

Wang (2003) reports that the activation function, in addition to introducing nonlinearity into the neural network, also constrains the range of neuronal outputs, preventing the ANN from exhibiting divergent or unbounded behavior. The logistic, arctan, and hyperbolic tangent functions are common activation functions; all exhibit a sigmoidal response, though they differ in their output ranges. In the present study, both tested functions yielded satisfactory results, as illustrated by the two best-performing networks: ANN 3 (logistic function) and ANN 50 (hyperbolic tangent function). However, despite the logistic function showing superior training performance, it may have been more susceptible to saturation effects. In contrast, the hyperbolic tangent function, whose output is symmetric around zero, enabled the convergence of 13 additional networks compared to the logistic function. The number of hidden neurons directly affects an ANN's performance, and determining the optimal size of this layer is not straightforward. Because no universal rule exists for defining this structure, both the number of hidden layers and the number of neurons per layer are typically determined through trial-and-error (SYDENHAM; THORN, 2005). In the present study, the simplest ANNs, those containing a single hidden layer, yielded the best results. However, it was not possible to precisely identify the optimal number of neurons, as the best-performing configurations contained 4 neurons (ANN 3) and 9 neurons (ANN 50) in the hidden layer.

According to Simon and Aliferis (2024), a large number of iterations may improve the learning process of an ANN and enable the model to fit the training data more closely; however, its generalization ability can deteriorate when evaluated on an independent dataset, indicating overfitting. In this situation, the ANN tends to memorize training examples rather than capture the underlying functional relationships. Conversely, when the number of iterations is insufficient, the network may fail to learn the data patterns, preventing the error from falling below an acceptable threshold and leading to underfitting. Both issues are often difficult to detect before model evaluation because they typically manifest subtly. This behavior was evident in ANN 8, whose overfitting became apparent only when the simulated taper profiles were examined. Underfitting was more easily identifiable in ANNs 73 and 66, where fewer iterations led to poor performance on both the training and validation datasets.

ANNs exhibit stochastic, adaptive, and dynamic behavior due to their distributed processing architecture. Although the theoretical functioning of the hidden layers in Multilayer Perceptron networks is well established, the internal configuration of weights and the individual contributions of each input variable cannot be interpreted directly (WU; FENG, 2018). This intrinsic variability helps explain why some networks converged after only a few iterations (ANNs 66 and 73), whereas others required nearly 100,000 iterations to reach the prescribed error threshold (ANNs 45 and 46). It also accounts for cases in which ANNs with more hidden neurons successfully converged (ANNs 26 and 80), whereas networks with fewer neurons failed to meet the error criterion (ANNs 20 and 76). This behavior is further illustrated by the fact that some complex networks converged relatively quickly (ANNs 34 and 72), whereas certain simpler architectures required substantially more iterations to converge (ANNs 1 and 45).

The training time of neural networks is directly influenced by the number of iterations and by the computational capacity of the hardware employed. Yang and Wang (2020) emphasize that training neural networks with hidden layers has traditionally been computationally demanding, particularly in deep architectures that require dozens or even hundreds of processing layers. This limitation has been substantially mitigated by advances across

multiple technological domains, most notably the rapid development of specialized hardware, such as graphics processing units and other general-purpose parallel computing devices.

Correcting non-sampling errors in sectional diameter measurements obtained with optical dendrometers is essential, as such errors are often difficult to detect and quantify. These inaccuracies can propagate through volume estimation procedures and taper-based assortment calculations, affecting the reliability of forest production estimates. In this context, ANNs have demonstrated promising performance, improving the accuracy of diameter measurements from optical dendrometers.

CONCLUSIONS

At the end of this study, the following conclusions were reached:

- The artificial neural networks tested were able to reduce the non-sampling errors present in the sectional diameter measurements obtained with the Criterion optical dendrometer, fulfilling the primary objective of the study and supporting the hypothesis that ANNs can improve the accuracy of indirectly measured diameters under the evaluated conditions.
- The best-performing architectures were Multilayer Perceptron ANNs models with a single hidden layer, in which the logistic activation function yielded superior results compared to the hyperbolic tangent for the species, variables, and training configuration assessed.
- The identification of clear cases of overfitting and underfitting highlighted the importance of selecting an appropriate number of iterations and a suitable architectural complexity to ensure adequate model generalization, particularly when the models are applied to datasets distinct from those used for training.
- For future research, it is recommended to expand the analysis to additional species, explore alternative ANN architectures, and evaluate regularization strategies that may mitigate overfitting and improve model robustness in different operational contexts.

ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), Finance Code 88887.510262/2020-00, under the Graduate Program Development Initiative (PDPG) in the Legal Amazon. We also thank Professor Décio José de Figueiredo, curator of the Laboratory of Dendrometry (LADEN) at the Federal University of Paraná (UFPR), for facilitating the loan of the Criterion RD 1000 optical dendrometer to the Laboratory of Forest Resource Measurement and Management (LabFor) at the Federal Rural University of the Amazon (UFRA).

REFERÊNCIAS

AMORIM, D. C. R. Quantificação de biomassa e altura das árvores em uma floresta de mata atlântica em Belo Horizonte - MG, comparação entre levantamento Laser Scanner e inventário florestal por área fixa. **Revista Gestão & Sustentabilidade Ambiental**, Palhoça, v. 9, n. 3, p. 711, 2020.

ARAÚJO, E. F.; FERNANDES, M. M.; SILVA, J. P. M.; KUNZ, S. H.; FERNANDES, M. R. de M. Estimativa da altura de espécies florestais em regeneração natural utilizando redes neurais artificiais. **Revista de Ciências Ambientais**, Canoas, v. 14, n. 3, p. 27–37, 2020.

BIAZATTI, S. C.; MÔRA, R.; SCCOTI, M. S. V.; BRITO JÚNIOR, J. F. de; QUEIROZ, N. dos S.; CURTO, R. de A. Criterion dendrometer as a non-destructive method for dendrometric estimations of native species in western Amazon. **Revista Ibero-Americana de Ciências Ambientais**, Aracaju, v. 11, n. 6, p. 59–70, 2020.

BUENO, G. F.; COSTA, E. A.; CRISTINA, A. N.; SOARES, A. A. V.; MIRANDA, R. O. V. de; SCHONS, C. T. Efeito do número de neurônios na camada oculta para relações hipsométricas de eucalipto usando redes neurais artificiais. **BIOFIX Scientific Journal**, Curitiba, v. 5, n. 2, p. 222–230, 2020.

CHAMMA, W. D. S.; BATISTELLA, D.; CRISIGIOVANNI, E. L.; VICTORINO, H. da S.; LIMA, V. A. Aprendizado de máquina aplicado em imagens de satélite para classificação de telhados. **Brazilian Journal of Development**, São José dos Pinhais, v. 7, n. 7, p. 72558–72576, 2021.

CURTO, R. D. A.; LAURO, A. C.; TONINI, H.; KOHLER, S. V.; ARAÚJO, E. J. G. de; BIAZATTI, S. C. Cubagem de árvores em pé com dendrômetro óptico em sistema de integração lavoura-pecuária-floresta. **Pesquisa Florestal Brasileira**, Colombo, v. 39, n. 1, p. 1–11, 2019.

DANTAS, D.; PINTO, L. O. R.; GONÇALVES, A. F. A.; TERRA, M. de C. N. S.; CALEGÁRIO, N. Predição volumétrica por meio da Krigagem pontual reduz o esforço de amostragem em inventários florestais pré-corte. **Caderno de Ciências Agrárias**, Montes Claros, v. 12, p. 1–9, 2020.

DONGARE, A. D.; KHARDE, R. R.; KACHARE, A. D. Introduction to artificial neural network. **International Journal of Engineering and Innovative Technology**, Huwei, v. 2, n. 1, p. 189–194, 2012.

DREW, P. J.; MONSON, J. R. T.; KINGDOM, U. Artificial Neural Networks. **Surgery**, Philadelphia, v. 127, p. 3–11, 2000.

DUARTE, L. T.; SILVA, D. B. do N.; BRITO, J. A. de M. Análise de paradados do censo demográfico 2010: uma investigação de fatores associados a erros não amostrais do levantamento de dados. **Revista Brasileira de Estudos de População**, Rio de Janeiro, v. 33, n. 3, p. 679–701, 2016.

FRITSCH, S.; GUENTHER, F.; WRIGHT, M. N.; SULING, M.; MUELLER, S. M. **Package "neuralnet": Training of neural networks**, 2019.

JHA, S. YANG, S.; BRANDEIS, T. J.; KUEGLER, O.; MARCANO-VEGA, H. Evaluation of regression methods and competition indices in characterizing height-diameter relationships for temperate and pantropical tree species. **Frontiers in Forests and Global Change**, Lausanne, v. 6, p. e1282297, 2023.

KROGH, A. What are artificial neural networks? **Nature Biotechnology**, Pittsburgh, v. 26, n. 2, p. 195–197, 2008.

LEAL, F. A.; LEAL, G. da S. A.; SILVA, T. C. da. Redes Neurais Artificiais e Modelos Alométricos Aplicados Para Estimativa de Volume e Altura Em Eucalyptus Urophylla S.T.Blacke. **Advances in Forestry Science**, Cuiabá, v. 7, n. 3, p. 1181–1188, 2020.

MELO, J. A. F. de; ROCHA, K. da S. Aplicação de redes neurais artificiais na classificação e zoneamento de focos de calor na reserva extrativista chico mendes - Acre, Brasil. **Observatório de la Economía Latinoamericana**, Curitiba, v. 21, n. 10, p. 16740–16763, 2023.

NASCIMENTO, R. L. X.; SOUZA, C. C. de; GRASSI, G. OLIVEIRA, M. A. das N. de. **Caderno de caracterização do estado do Pará**. Brasília: Codevasf, 2022, 146 p.

NICOLETTI, M. F.; MACHADO, S. do A.; CARVALHO, S. de P. C. e; FIGUEIREDO FILHO, A.; SANQUETTA, C. R. Descrição do perfil do fuste de árvores por meio das funções de afilamento. **Scientia Agraria Paranaensis**, Marechal Cândido Rondon, v. 15, n. 4, p. 365–372, 2016.

RAFIQ, M. Y.; BUGMANN, G.; EASTERBROOK, D. J. Neural network design for engineering applications. **Computers and Structures**, England, v. 79, n. 17, p. 1541–1552, 2001.

REIS, L. P.; SOUZA, A. L. de; REIS, P. C. M. dos R.; MAZZEI, L. BINOTI, D. H. B.; LEITE, H. G. Diameter structure in a community of shrub-tree caatinga, municipality of Floresta, state of Pernambuco, Brazil. **Floresta**, Curitiba, v. 48, n. 1, p. 133–142, 2018.

SIMON, G.; ALIFERIS, C. **Artificial intelligence and machine learning in health care and medical sciences: best practices and pitfalls**. Cham: Springer, 2024, 810 P.

SYDENHAM, P. H.; THORN, R. **Handbook of measuring system design**. 3rd. ed. Stillwater: Wiley, 2005, 1648 p.

WANG, S.-C. **Interdisciplinary computing in Java programming language**. Amsterdam: Kluwer Academic Publishers, 2003, 266 p.

WICKHAM, H.; CHANG, W.; HENRY, L.; PEDERSEN T. L.; TAKAHASHI, K. WILKE, C.; WOO, K.; YUTANI, H.; DUNNINGTON, D.; BRAND, T. van den. **Package "ggplot2": create elegant data visualisations using the grammar of graphics**, 2019.

WU, Y. C.; FENG, J. W. Development and application of artificial neural network. **Wireless Personal Communications**, Berlin, v. 102, n. 2, p. 1645–1656, 2018.

YANG, G. R.; WANG, X. J. Artificial neural networks for neuroscientists: a primer. **Neuron**, Maryland Heights, v. 107, n. 6, p. 1048–1070, 2020.