

ARTIFICIAL NEURAL NETWORKS AND MIXED-EFFECTS MODELING TO DESCRIBE THE STEM PROFILE OF *Pinus taeda* L.

Izabel Passos Bonete^{1*}, Julio Eduardo Arce², Afonso Figueiredo Filho³, Fabiane Aparecida de Souza Retslaff³, Luciano Rodrigo Lanssanova⁴

^{1*}Universidade Estadual do Centro Oeste, Irati, Paraná, Brazil – ipbonete@unicentro.br*

²Universidade Federal do Paraná, Curitiba, Paraná, Brazil - julioarce001@gmail.com

³Universidade Estadual do Centro Oeste, Irati, Paraná, Brazil – afigfilho@gmail.com; faretslaff@gmail.com

⁴Instituto Federal do Mato Grosso, Juína, Mato Grosso, Brazil - luciano.lanssanova@jna.ifmt.edu.br

Received for publication: 24/09/2018 – Accepted for publication: 28/11/2018

Resumo

Redes neurais artificiais e modelagem de efeitos mistos na descrição do perfil do fuste de Pinus taeda L. Este estudo buscou comparar Redes Neurais Artificiais (RNAs) e Modelagem de Efeitos Mistos (MEM) na descrição do perfil do fuste de *Pinus taeda* L., utilizando-se dados amostrais de 246 árvores. Primeiramente, foram ajustadas três funções de afilamento, de diferentes classes: não segmentada, segmentada e forma variável. Para o ajuste dos modelos, foi utilizada a técnica de regressão não linear (*nls*). Na equação de melhor desempenho para estimativas de diâmetros, ajustada por *nls*, foi aplicada a MEM não linear (*nlme*), em dois níveis, utilizando o fator classe de idade (*ci*) e classe de diâmetros (*cd*). Para tanto, foram compostos três diferentes cenários de estudo, com número de coeficientes com efeitos aleatórios, variando de um a três em cada cenário. Os ajustes foram realizados por meio das funções *nls* e *nlme* do *software* R. As equações selecionadas de efeito misto foram comparadas com RNAs, geradas no *software* Neuro 4.0. Os modelos de funções de afilamento e RNAs foram classificados, conforme critérios estatísticos e análise gráfica de resíduos. A equação de afilamento de Bi (2000) apresentou melhor desempenho para estimativas de diâmetros em relação às equações não segmentada e segmentada. A aplicação da técnica *nlme* na equação de Bi (2000) aumentou a precisão das estimativas dos diâmetros para *Pinus taeda*, em relação ao ajuste realizado pela técnica *nls*. Na comparação das RNAs com as variações da equação de Bi de efeitos mistos, as redes neurais apresentaram melhor desempenho, sendo indicadas na descrição do perfil de fustes de *P. taeda*.

Palavras-chave: Manejo Florestal; afilamento; diâmetros ao longo do tronco.

Abstract

The aim of this study was to compare the effectiveness of artificial neural networks (ANNs) and mixed-effects models (MEMs) in describing the stem profile of *Pinus taeda* L., using sample data from 246 trees. First, three taper functions of different classes were adjusted: non-segmented, segmented, and variable-form. To adjust the models, the nonlinear regression technique (*nls*) was used. In the best performance equation for *nls*-adjusted diameter estimates, the nonlinear MEM (*nlme*) was applied at two levels, using the age class (*ci*) and DBH class (*cd*). For this, three different study scenarios were considered, with the number of coefficients with random effects ranging from one to three in each scenario. The adjustments were made using the *nls* and *nlme* functions in R software. The selected mixed-effect equations were compared with ANNs generated in Neuro 4.0 software. The taper function models and ANNs were classified according to statistical criteria and graphical analysis of residues. The tapering equation of Bi (2000) presented better performance for diameter estimates than the non-segmented and segmented equations. Application of the *nlme* technique in the Bi (2000) equation increased the accuracy of the diameter estimates for *Pinus taeda*, in relation to the adjustment using the *nls* technique. In the comparison of ANNs with the variations of the Bi equation of mixed-effects, the networks performed better, indicated in the description of the *P. taeda* profile.

Keywords: Forest Management; tapering; diameters along of stem.

INTRODUCTION

Accurate estimates of forest production are fundamental for companies in the forest sector. However, the estimation of tree volume is hampered mainly by variation in the shape of the stem profile. Variation in the shape of the stem is related to the decrease in diameter from the bottom to the top of the stem. This is called tapering, thinning, or taper.

The shape and tapering of the trees stem have been the subject of forest research for over a century, considering their relevance and applicability to estimates of diameter and volume by assortment. Many statistical models and multivariate techniques have been tested to describe the shape and tapering of stems. In this context there are three different types of models: non-segmented models that describe the tapering of the stem with a

single function, segmented models that divide the stem into two or more segments by means of different subfunctions, mathematically joined with the aid of inflection points, and variable-form or exponent-form models that describe trunk tapering as a continuous function, using an exponent that varies to compensate for changes in trunk shape in different sections of the tree and still uses the principal components method and cubic *splines* functions.

Although some techniques have proven their effectiveness, none has performed better for all species and conditions to which they have been subjected. However, for the forestry sector, it is essential to obtain reliable information on the stem profile, as the accuracy of stem diameter and volume estimates is crucial for the success of optimizing processes in the forestry sector (CAMPOS *et al.*, 2017).

Thus, in the search for efficient methodologies for estimating stem volume it was found that since the 1990s, the mixed-effects modeling (MEM) technique has been proposed in the forestry literature for data analysis with repeated measures, such as data used in tapering studies. These data are characterized by repeating measurements of the same individual and generally grouping these using one or more grouping factors, so that within each group the observations are likely to be correlated (PINHEIRO; BATES, 2000).

Thus, characterized by using fixed and random effects parameters, linear and nonlinear MEMs are a tool that is used to model correlations within each group and to improve the estimates obtained using regression techniques, as they allow generalizing structures of spatio-temporal correlations and non-constant variance (GOUVEIA *et al.*, 2015).

The use of nonlinear MEM in the description of tree stems began with Gregoire and Schabenberger in 1996. Internationally, the technique started to be used, with emphasis on its application to different species in the segmented model of Max and Burkhart (1976) (TRINCADO; BURKHART, 2006; ÖZÇELİK *et al.*, 2011) and in variable-form models (GARBER; MAGUIRE, 2003; DE-MIGUEL *et al.*, 2012). In Brazil, this technique is still novel and little used, with the exception of the studies by Carvalho *et al.* (2014) and Schröder *et al.* (2014).

For Campos *et al.* (2017), the search for better estimates directs the research to the test of new tools and methodologies, with an aim to describing the profile of the tree stem. From this perspective, Schikowski *et al.* (2015), point out that with the advancement of computer science, the use of artificial neural networks (ANNs) or simple neural networks in the field of estimation is gaining attention. Neural networks are defined as parallel computational systems that consist of simple processing units (neurons) and are connected to each other in a specific way, making them able to perform a certain task in the same way as the human brain (LEITE *et al.*, 2016; BINOTI *et al.*, 2013).

Studies using this technique to estimate different variables have been developed internationally, in different locations and situations, generally presenting satisfactory results in tree height modeling (ÖZÇELİK *et al.*, 2013), in forest growth modeling (ASHRAF *et al.*, 2015), and in diameter distribution modeling (DIAMANTOPOULOU *et al.*, 2015).

As an estimation method for Brazilian forest science, the use of neural networks is still novel, but different topics have been investigated, testing the effectiveness of the technique. For example, the use of neural networks for the description of the stem profiles of different species and for different conditions has been studied by Soares *et al.* (2011), Schikowski *et al.* (2015), Martins *et al.* (2016), Silva *et al.* (2016), Campos *et al.* (2017) and, Martins *et al.* (2017).

In this context, the present study investigated the use of nonlinear MEM and ANNs with the aim to compare the performance of these methodologies in the description of stem profiles of *Pinus taeda* L. Such methodologies allow the use of a single model for different categorical classes without the need to make various adjustments. Research on the performance of new techniques that provide more accurate estimates of forest productivity is fundamental for improvement and development of the forest sector.

This study thus seeks to make a relevant contribution to knowledge and to assist forest-based companies by presenting results of advanced techniques that can be used in decision-making in the management of *P. taeda*.

MATERIAL AND METHODS

The study area is located in the region of Telemaco Borba, state of Paraná. We used sample data from 246 *Pinus taeda* L. trees from non-thinned stands with initial planting spacing of 2.5 x 2.5 m. The trees had a diameter at breast height (d) ranging from 5 to 45 cm and ages ranging from 4 to 19 years.

The data were randomly divided into two groups: one consisting of 80% of the total trees to adjust the regression models and conduct network training, and the other, with 20% of the total trees to validate the regression-adjusted equations and generalization of trained ANNs (LEITE *et al.*, 2016; BINOTI *et al.*, 2013).

The diameters with bark along the shaft were measured, being a measure at the base (h_1) and at heights (h_i) of 0.7 m, 1.3 m, 2.0 m and then at distances of 2.0 m, to the full height h . Height at the tree base (h_1) ranged from

0.02 m to 0.20 m and the last section was considered as height h_n . This was the height closest to the total height (h) of each tree in which the last diameter measurement was taken.

The stratification of the database by diameter class or age enables better estimates for tapering (KOHLER *et al.*, 2013). In this sense, we opted for stratification based on age classes (ci) and diameter classes at breast height (cd), considering the 16 existing ages as 16 classes of amplitude equal to one year and 8 classes of diameters, being: below 10 cm (1), 10 to 14.99 cm (2), 15 to 19.99 cm (3), 20 to 24.99 cm (4), 25 to 29.99 cm (5), 30 to 34.99 cm (6), 35 to 39.99 cm (7) and equal to or greater than 40 cm (8).

For nonlinear MEM, three tapering functions were first tested, whose expressions for diameter estimates are presented in a sequence: Model 1 of non-segmented type, Model 2 of segmented type and Model 3 of variable-form type.

Model 1 - 5th degree polynomial or model of Schöepfer (1966)

$$d_i = d. [\beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \beta_4 X^4 + \beta_5 X^5] + \varepsilon_i$$

Model 2 – Max and Burkhardt (1976)

$$d_i = d. (\beta_1(X - 1) + \beta_2(X^2 - 1) + \beta_3(\alpha_1 - X)^2 I_1 + \beta_4(\alpha_2 - X)^2 I_2)^{0.5} + \varepsilon_i$$

Model 3 – Bi (2000)

$$d_i = d. \left[\left(\left(\log \sin \left(\frac{\pi}{2} X \right) \right) / \left(\log \sin \left(\frac{\pi}{2} \cdot \frac{1,3}{h} \right) \right) \right)^{\beta_0 + \beta_1 \sin(\frac{\pi}{2} X) + \beta_2 \cos(\frac{3\pi}{2} X) + \beta_3 \sin(\frac{\pi}{2} X) / X + \beta_4 d + \beta_5 X \sqrt{d} + \beta_6 X \sqrt{h}} \right] + \varepsilon_i$$

Where: d = diameter at breast height with bark (cm); h = total height (m); h_i = height (m) from stem base to diameter d_i ; d_i = diameter (cm) with bark at a height h_i ; $X = h_i/h$; α_i = inflection points of the Max-Burkhardt model; β_i = coefficients of the models; $I_1 = 1$ if $X \leq \alpha_1$ or $I_1 = 0$ if $X > \alpha_1$; $I_2 = 1$ if $X \leq \alpha_2$ or $I_2 = 0$ if $X > \alpha_2$; ε_i = random error.

The adjustments of the proposed taper models were performed using the software R using the nonlinear adjustment technique (*nls*). The significance of the equation coefficients was evaluated using the t-test ($\alpha = 5\%$). Variables related to non-significant coefficients were eliminated and the model was adjusted again.

The best performing taper equation, adjusted by *nls* for diameter estimation without stem stratification, was fitted by nonlinear MEM at two grouping levels. According to Pinheiro and Bates (2000), the nonlinear mixed-effects model for a grouping level is formulated as:

$$y_i = f_i(\emptyset_i, \vartheta_i) + \varepsilon_i, i = 1, \dots, M$$

Where: y_i is the vector ($n_i \times 1$) of the response variable for the i -th group, M is the number of groups, f is a differentiable function of a specific parameter vector \emptyset_i , ϑ_i is the covariate vector and ε_i is the vector ($n_i \times 1$) of the random error between the groups, with the assumption of independence, normality and homoscedasticity i.e.: $\varepsilon \sim N(0, \Sigma_i)$, even though the database contains repeated measurements within the group (hierarchical data), with heterogeneous variance and correlation between residues violating the basic assumptions of regression.

The function f is nonlinear in at least one component of the parameter vector \emptyset_i of a specific group, which is modeled in a second stage by:

$$\emptyset_i = X_i \beta + Z_i b_i, b_i \sim N(0, D)$$

Where: X_i is the matrix ($n_i \times p$) of fixed effects covariates; Z_i is the matrix ($n_i \times q$) of random effects covariates; β is the vector ($p \times 1$) of fixed effects; b_i is the vector ($q \times 1$) of random effects, independent and normally distributed with mean 0 and variance-covariance matrix D .

In the case of two grouping levels, the nonlinear mixed-effects model is represented by:

$$y_{ij} = f_{ij}(\emptyset_{ij}, \vartheta_{ij}) + \varepsilon_{ij}, \text{ com } i = 1, \dots, M, j = 1, \dots, M_i$$

Where: y_{ij} is the vector ($n_{ij} \times 1$) of the response variables for the j th group of the second level ($j = 1, \dots, M_i$), nested in the i th group of the first level ($i = 1, \dots, M$); f is a differentiable function of a specific parameter vector \emptyset_{ij} , ϑ_{ij} is the covariate vector and ε_{ij} is a vector ($n_{ij} \times 1$) of the random error within the group, with the assumption of independence, normality and homoscedasticity i.e.: $\varepsilon_{ij} \sim N(0, \Sigma_{ij})$.

The f function is nonlinear in at least one component of the parameter vector \emptyset_{ij} of a specific group, which is modeled in a second stage by:

$$\emptyset_{ij} = X_{ij} \beta + Z_{i,j} b_i + Z_{ij} b_{ij} + \varepsilon_{ij}, \text{ com } b_i \sim N(0, D_1) \text{ e } b_{ij} \sim N(0, D_2)$$

Where: β is the vector ($p \times 1$) of the fixed effects; b_i is the vector ($q_1 \times 1$) of the first level random effects, independent and normally distributed with mean 0 and variance and covariance matrix D_1 ; b_{ij} is the vector ($q_2 \times 1$) of independent and, normally distributed, second-level random effects with mean 0 and variance-covariance matrix D_2 ; X_{ij} is the matrix ($n_{ij} \times p$) of fixed effects covariates; $Z_{i,j}$ is the matrix ($n_{ij} \times q_1$) of covariates of the first level random effects and the index i, j means that j is nested in i ; Z_{ij} is the matrix ($n_{ij} \times q_2$) of second level random effects covariates

Regarding the grouping levels used, the proposal of Carvalho *et al.* (2014) was adopted in which age (ci) and diameter (cd) classes were considered as random effects of the process. It was based on the principle adopted by these authors that variations in diameter classes in the corresponding age classes happen at random and thus should be considered as a random component of the mixed model. Therefore, in the application of the nonlinear MEM technique the data were organized with correlated measurements in diameter classes nested within the age classes, rather than the natural hierarchical structure of nested measurements within the individual trees.

For model adjustment, the *nlme* package implemented in the software R was used, composing different study scenarios, with number of coefficients with random effects, varying from one to three in each scenario, according to De-Miguel *et al.* (2012) and Schröder *et al.* (2014). According to mixed model theory, all coefficients can be randomized, however, this would lead to overparameterization and convergence problems (SCHRÖDER *et al.*, 2014).

In the training and generalization of the ANNs, the Neuro 4.0 system was used, with MLP (Multilayer Perceptron) network configurations and the resilient propagation training algorithm, in the RPROP + variation, as it is an efficient algorithm for stem profile description (SCHIKOWSKI *et al.*, 2015; MARTINS *et al.*, 2016). The sigmoidal activation function (Eq. 1) was used in the hidden layer and in the output layer because it is the most common in neural network tapering studies (SCHIKOWSKI *et al.*, 2015; SILVA *et al.*, 2016).

$$f(u) = \frac{1}{1 + e^{-\beta u}} \quad (1)$$

Where: β is the slope parameter of the sigmoid function; u is the linear combiner resulting from the sum of the input signal products by their weights.

Numerical variables were linearly normalized in the range 0 to 1 and categorical variables were normalized using the 1-of-N methodology (MARTINS *et al.*, 2016; CAMPOS *et al.*, 2017). Each quantitative variable corresponds to a single neuron in the input layer of the network and for each qualitative variable the number of neurons depends on the number of classes of the variable, there being one neuron for each class.

Four different neural network configurations, consisting of four, five or six input variables resulting from the combinations of two qualitative variables, were tested: age classes (ci) with 16 categorical variables and diameter classes (cd) with eight categorical variables, besides the variables quantitative d, h_i, h and their relations. Thus, the four configurations tested were composed using the variables: 1^a) $ci, cd, d, h/h$; 2^a) ci, cd, d, h_i, h ; 3^a) $ci, cd, d, h/h, (h_i/h)^2$; and 4^a) $ci, cd, d, h_i, h/h, (h_i/h)^2$.

We introduced variation in network architecture by testing different numbers of hidden layer neurons ranging from 1 to 15 neurons following Schikowski *et al.* (2015) and 10 networks were trained for each architecture variation, making 150 networks for each configuration. As a criterion for stopping the training of the networks the standard format of the Neuro 4.0 software was adopted, in which ANN interrupts the weight adjustment when it reaches the average error of 0.0001 or 3.000 training cycles. In order to select the best ANN in each configuration, the networks were evaluated by the correlation coefficient between the observed and estimated values ($r_{\hat{y}y}$), by the square root of the mean error in percentual (RMSE%) and by the graphical analysis of the residuals for the set training and generalization data provided by the Neuro 4.0 software.

To evaluate the performance in the adjustment / training and validation / generalization of the previously selected models and neural networks, the following statistics were used: the correlation coefficient between observed and estimated diameters ($r_{\hat{y}y}$) (Eq. 2), square root of the mean error in percentual (RMSE%) (Eq. 3), mean absolute percentage error (MAPE) (Eq. 4), Akaike Information Criterion (AIC) (Eq. 5) and graphical analysis of residues (Res%) (Eq. 6).

$$r_{\hat{y}y} = \frac{n(\sum_{i=1}^n y_i \hat{y}_i) - (\sum_{i=1}^n y_i) \cdot (\sum_{i=1}^n \hat{y}_i)}{\sqrt{[n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2] \cdot [n \sum_{i=1}^n \hat{y}_i^2 - (\sum_{i=1}^n \hat{y}_i)^2]}} \quad (2)$$

$$RMSE\% = \frac{100}{\bar{y}} \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (3)$$

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} * 100 \quad (4)$$

$$AIC = -2 \log(\text{likelihood}) + 2p \quad (5)$$

$$\text{Res\%} = \frac{(y_i - \hat{y}_i)}{y_i} \cdot 100 \quad (6)$$

Where: y_i = observed values; \hat{y}_i = estimated values; \bar{y}_i = mean of dependent variable values and n is the total number of observations; \log is the decimal logarithm; likelihood is the model's maximum likelihood estimator and p is the number of model/network parameters.

The choice of the equation adjusted using nls and nlme, as well as the most suitable ANNs for diameter estimates was made from classification, according to the sum of the values that were attributed to the statistics on fit/training and validation/generalization associated with graphical analysis of residues.

For the assignment of the classification values the most accurate equation or network in each evaluated statistic received the grade 1, the one that took second place received a score of 2, and so on. The most accurate equation or ANN for the estimates along the stem was the one that presented the lowest sum of the classification grades and whose residuals did not show a high degree of tendency.

RESULTS

The models adjusted using *nls* and their respective coefficients are presented in Table 1. The proposed models converged and regarding the significance of the coefficients only the variable-form model of Bi did not present all significant coefficients from the t-test ($\alpha = 5\%$).

Table 1. Coefficients obtained in the adjustment of the non-segmented (NS), segmented (S) and variable-form (VF) tapering functions for *Pinus taeda*.

Tabela 1. Coeficientes estimados no ajuste das funções de afilamento do tipo não segmentadas (NS), segmentadas (S) e de forma variável (FV) para *Pinus taeda*.

Type	Model	β_0	β_1	β_2	β_3	β_4	β_5	β_6	α_1	α_2
NS	1	1.210	-4.171	18.162	-38.865	35.828	-12.163			
S	2		-3.999	1.985	143.500	-1.939			0.063	0.715
VF	3	1.853	-0.261	-0.039	-0.979	(-)	0.050	-0.100		

Where: NS = untargeted equation; S = segmented equation; VF = equation of variable-form; β_i = estimated coefficients; α_i = inflection points of the segmented model; (-) represent the coefficients taken from the model as not significant at 5% probability of error by the t-test.

Using the equations obtained by *nls*, the diameters d_i with bark at the heights h_i of the stems were estimated for the fit and validation data set. The equations presented satisfactory results, indicating that they are efficient to adequately explain the diameter estimates along the shaft (Table 2).

Table 2. Statistics of equations adjusted using nonlinear regression (*nls*) to evaluate the estimates of diameters along the stem of *Pinus taeda*.

Tabela 2. Estatísticas das equações ajustadas por regressão não linear (*nls*) para avaliar as estimativas de diâmetros ao longo do fuste para *Pinus taeda*.

	Equation	$r_{\hat{y}y}$	RMSE%	MAPE	AIC	C
Adjustment	1	0.9886(3)	9.24(3)	10.07(3)	9,493.1(3)	3
	2	0.9898(2)	8.72(2)	9.54(2)	9,227.6(2)	2
	3	0.9910(1)	8.13(1)	8.85(1)	8,904.9(1)	1
Validation	1	0.9829(3)	10.96(3)	11.31(3)	2,461.7(3)	3
	2	0.9853(2)	10.19(2)	10.78(2)	2,379.9(2)	2
	3	0.9863(1)	9.74(1)	10.35(1)	2,329.9(1)	1

Where: $r_{\hat{y}y}$ = correlation coefficient between observed and estimated values; RMSE% = square root of the mean error in percentual; MAPE = mean absolute percentage error; AIC = Akaike Information Criterion; C = classification.

The distribution of residues for diameter estimates of the taper models tested is shown in figure 1.

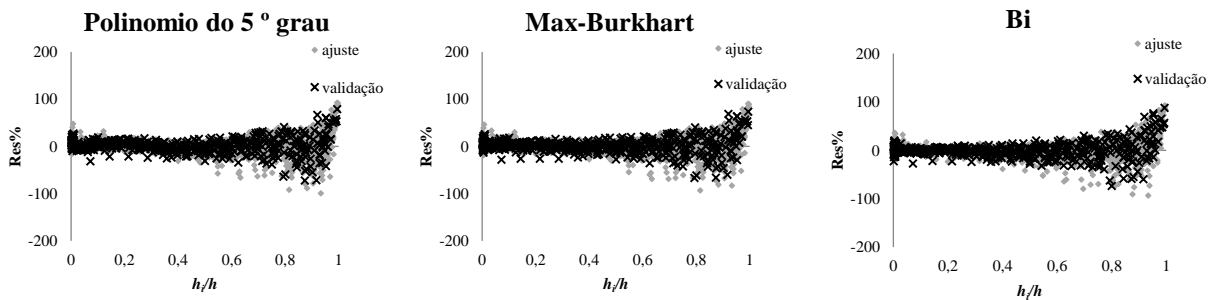


Figure 1. Residues (%) as a function of h_i/h for estimates of diameters along the stem for *Pinus taeda*, referring to the non-segmented, segmented and variable equations.

Figura 1. Resíduos (%) em função de h_i/h para estimativas de diâmetros ao longo do fuste para *Pinus taeda*, referentes às equações não segmentada, segmentada e de forma variável.

Considering that the variable-form tapering model proposed by Bi (2000), adjusted and validated using the non-linear regression technique (*nls*) with six significant coefficients ($\beta_0, \beta_1, \beta_2, \beta_3, \beta_5$ and β_6), presented better performance in relation to the non-segmented and segmented models, this model was adjusted using nonlinear MEM (*nlme*) as a function of two factors: age classes (*ci*) and diameter classes (*cd*) in a single adjustment procedure using a multilevel factor.

For this, all possible combinations of random coefficients from one to three parameters were tested. A total of 41 models were used with mixed effects for diameter estimations with bark and without stem stratification. Using the model evaluation statistics, the best performance equations were selected for each scenario. These were later compared with the ANNs.

The selected equations were named, $Bi.\beta_1$, $Bi.\beta_2\beta_3$ and $Bi.\beta_0\beta_1\beta_3$, indicating that they are variations of the Bi equation with the respective mixed coefficients that most improved diameter estimates in each scenario.

To evaluate the diameter estimates for *P. taeda* using ANNs in which four different neural network configurations were tested, composed of four, five or six input variables, 600 neural networks were generated (4 configurations x 15 architectures in each configuration x 10 networks for each architecture). Satisfactory results were obtained with configurations containing at least four hidden layer neurons. In the total set of networks generated inefficient networks were observed when in the presence of few neurons in the intermediate layer, as well as in networks with low values of RMSE% in training (6.39%) but with increasing values of RMSE% in generalization (10.48%) when in the presence of more than twelve neurons in the occult layer.

Based on the correlation coefficient between the observed and estimated diameters (r_{yy}), in the square root of the mean error in percentual (RMSE%) obtained in training and generalization, and the graphical analysis the percents residuals provided by the Neuro 4.0 software, was chosen the best network of each architecture and subsequently the best network of each configuration, reducing to three, the number of networks with characteristics suitable for diameter estimates, with r_{yy} ranging from 0.9939 to 0.9942 and RMSE% from 6.52% to 6.69%, which were named ANN 1, ANN 2 and ANN 3.

The first network, ANN 1, of architecture 5.12.1, was composed of five variables [*ci, cd, d, h_i, h*], totaling 27 neurons in the input layer, twelve neurons in the hidden layer and one neuron in the output layer; ANN 2, of architecture 5.10.1, was composed of five variables [*ci, cd, d, h_i/h, (h_i/h)²*], which is equivalent to 27 neurons in the input layer, ten neurons in the hidden layer and one neuron in the output layer, and ANN 3, of architecture 6.12.1, was composed of six variables [*ci, cd, d, h_i, h_i/h, (h_i/h)²*], which corresponds to 28 neurons in the input layer, twelve neurons in the hidden layer and one neuron in the output layer.

The fit/training and validation/generalization statistics of the nlme-adjusted equations and the selected ANNs for diameter estimates for *P. taeda* are presented in Table 3.

The equation $Bi.\beta_0\beta_1\beta_3$ provided in the adjustment, an improvement in the RMSE% statistic of 16.85% over Bi, which is equivalent to a reduction of 1.37 percentage points in this statistic; 11.17% over $Bi.\beta_1$ (Scenario 1), equivalent to a reduction 0.85 percentage points and 3.15% over $Bi.\beta_2\beta_3$ (Scenario 2), indicating a reduction of 0.22 percentage points. There was also a reduction in MAPE and AIC values of 9.6% and 3.4%, respectively, in relation to $Bi.\beta_1$ (Scenario 1) and of 0.1% and 1.2% in relation to $Bi.\beta_2\beta_3$. (Scenario 2). In general, the residual scatter plots for adjustment and validation data of the three equations tested showed residual homogeneity in the basal and intermediate portion of the shaft.

Table 3. Performance statistics of the equation of Bi and ANNs trained for the estimation of the diameter along the stem for *Pinus taeda*.

Tabela 3. Estatísticas de desempenho das variações da equação de Bi e RNAs treinadas para a estimativa do diâmetro ao longo do fuste para *Pinus taeda*.

	Model	$r_{\hat{y}y}$	RMSE%	MAPE	AIC	C
Adjustment/ Training	Bi. β_1	0.9921(6)	7.61(6)	8.15(6)	8,766.8(6)	6
	Bi. $\beta_2\beta_3$	0.9933(5)	6.98(5)	7.45(4)	8,571.5(2)	5
	Bi. $\beta_0\beta_1\beta_3$	0.9938(4)	6.76(4)	7.37(3)	8,468.2(1)	3
	ANN 1	0.9939(3)	6.69(2)	7.62(5)	8,695.8(5)	4
	ANN 2	0.9938(2)	6.72(3)	7.24(2)	8,595.8(3)	2
	ANN 3	0.9942(1)	6.52(1)	6.94(1)	8,601.5(4)	1
Validation/ Generalization	Bi. β_1	0.9875(1)	9.29(1)	10.32(5)	2,276.3(1)	2
	Bi. $\beta_2\beta_3$	0.9865(4)	9.69(4)	9.71(4)	2,323.5(3)	4
	Bi. $\beta_0\beta_1\beta_3$	0.9867(2)	9.60(2)	9.59(1)	2,313.5(2)	1
	ANN 1	0.9862(5)	9.82(5)	10.39(6)	3,025.1(5)	6
	ANN 2	0.9861(6)	9.84(6)	9.67(3)	2,907.2(4)	5
	ANN 3	0.9866(3)	9.68(3)	9.66(2)	3,032.5(6)	3

Where: $r_{\hat{y}y}$ = correlation coefficient between observed and estimated values; RMSE% = Square Root of the Mean Error in percentual; MAPE = Mean Absolute Percentage Error; AIC = Akaike Information Criterion; C = classification.

In ANNs, ANN 3 provided in training, an improvement in the RMSE% statistic of 2.5% over ANN 1, which is equivalent a 0.17 percentage points reduction, and of 2.98% over ANN 2, indicating a reduction of 0.20 percentage points. There was also a reduction in MAPE values of 5.9% compared to ANN 1 and 4.1% compared to ANN 2. However, although the network AIC values were very close, with variations compared to ANN 3, of a maximum of 1%, there was an increase in the AIC values in relation to the variations of the Bi equation (Scenarios 1, 2 and 3), supposedly due to the large number of parameters of the networks ANN 1, ANN 2, and ANN 3, which totaled 349, 291, and 361 parameters, respectively.

Residues distributions of variations of the Bi equation and ANNs selected for stem diameter estimates for *P. taeda* are shown in figure 2.

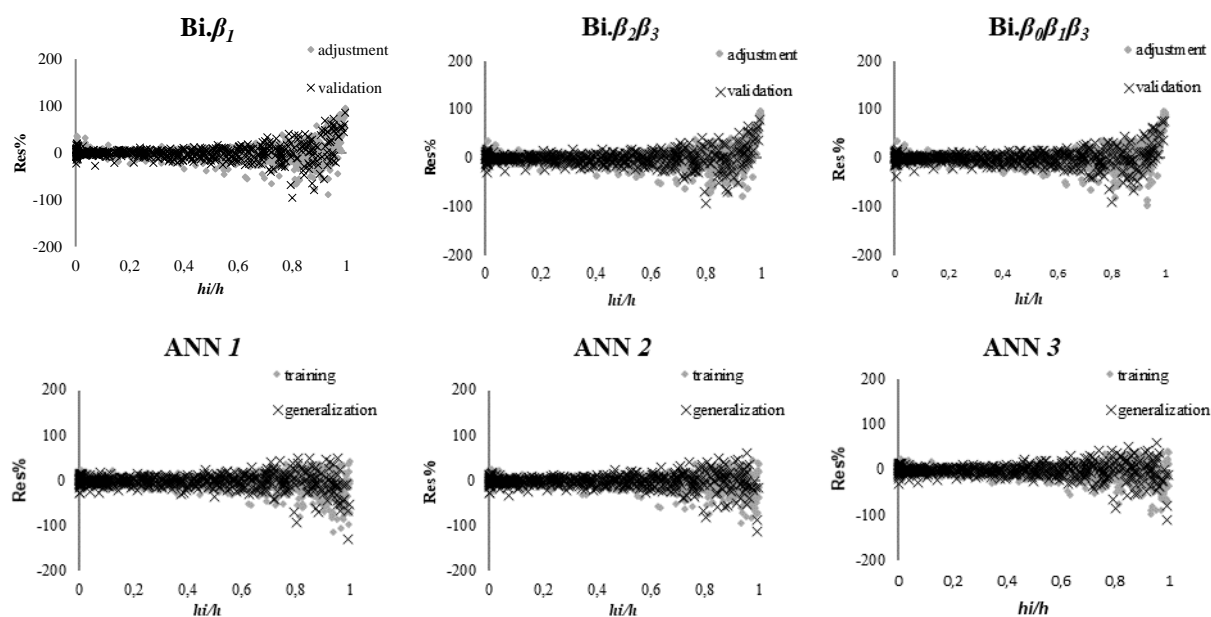


Figure 2. Residues (%) as a function of h_i/h for diameter estimates without stratification of *Pinus taeda* stems for the variations of the Bi equation and for the selected ANNs.

Figure 2. Resíduos (%) em função de h_i/h para estimativas de diâmetros sem estratificação do fuste de *Pinus taeda* para as variações da equação de Bi e para as RNAs selecionadas.

DISCUSSION

Accurate estimates with the variable-form function proposed by Bi were obtained by De-Miguel *et al.* (2012) and Schröder *et al.* (2014). This function was sometimes referred to as one of the best, when compared with other functions of variable-form, such as those of Kozak (1988, 1994, 1995 and 2004) and Riemer *et al.* (1995) (SCHRÖDER *et al.*, 2014).

Although the 5th degree polynomial is one of the most widely used non-segmented models to describe the profiles of *P. taeda* and *P. elliottii* in southern Brazil (KÖHLER *et al.*, 2013; TEO *et al.*, 2013), the equation occupied the third position in terms of performance in the present study.

In relation to Max and Burkhart's segmented model, it presented better performance compared to the non-segmented model (5th degree polynomial), supposedly because the Max and Burkhart segmented model uses three functions to represent tree profiles and it considers two points of morphological change in the shaft, which makes it more flexible than the 5th degree polynomial. Assis *et al.* (2001) tested segmented and non-segmented models for *P. taeda* and also found better estimates in segmented versus untargeted models.

The graphs of the equations showed great similarity, with larger error amplitudes as the diameters approach the apex (Figures 1 and 2) where the smallest diameters occur, both for fit and validation data. This trend has already been observed in other studies on tapering in *P. taeda* such as those conducted by Kohler *et al.* (2013) and Téo *et al.* (2013) using different statistical models and with under and overestimation of diameters, even at heights greater than 40% of the shaft. Kohler *et al.* (2013) clarified that, as these diameters have little influence on the estimation of commercial volume, trends for underestimation or overestimation of these diameters can be considered irrelevant.

The application of nonlinear MEM in the Bi equation, with random effects on age and diameter classes, improved the diameter estimates for *P. taeda* in relation to the fixed effect model in the three studied scenarios, as the number of coefficients with random effects increased. Thus, in the $Bi.\beta_0\beta_1\beta_3$ equation, evaluated as the best performer, the coefficients β_0 , β_1 and β_3 were the ones that most influenced the accuracy of the estimates. Similar results regarding the improvement of MEM in relation to the fixed effects equation were also achieved by Garber and Maguire (2003), Trincado and Burkhart (2006), Özçelik *et al.* (2011) and Schröder *et al.* (2014), with the development of different tapering equations for different forest species.

In practice, the MEM technique allows the use of a single equation with specific random effects for different categorical classes, without the need for model adjustment for each class and, mainly, without the need for model calibration for a new data set. For Schröder *et al.* (2014), the effect on diameter classes, allows the application of MEM in data from forest inventories, since models developed with random effects on the individual tree may have academic value, but are not practical due to the need for calibration processes for independent data which are required when the effect is on the individual tree.

Regarding ANNs, it was found that the number of architecture-trained networks met expectations for predictive performance. In the presence of more than 12 neurons in the occult layer, overtraining was found, a problem in which neural networks fit the data set very well, but are inefficient to predict new results. Given the possibility of inclusion of categorical variables in the ANNs, only one network can explain the longitudinal profile of the stems, without the need for stratification as in conventional regression, allowing, as well as in MEM, the application of a single neural network in variable estimates for forest resource planning purposes.

Among the three networks selected in the study, there was an improvement in the diameter estimates for *P. taeda*, being ANN 3, with six input variables [ci , cd , d , h_i , h_i/h , $(h_i/h)^2$], the best performing neural network. Campos *et al.* (2017) obtained even better results in training ($r_{\hat{y}} = 0.9950$ and $RMSE\% = 5.3\%$) by analyzing the ability of the ANN to describe the stem profile of trees of different genera and species, under varying growing conditions. In addition, they obtained scatter plots of non-biased residues, which allowed these authors to conclude that the neural network has the ability to reliably describe the stem profile. Schikowski *et al.* (2015) obtained, on average, for the three best selected networks, higher $r_{\hat{y}}$ values (0.9940) and lower $RMSE\%$ (5.73%) than those obtained in the present study.

Nonlinear MEMs and ANNs were ranked and it was found that, for adjustment data, neural networks performed better than Bi equation variations, with ANN 3 showing better performance, followed by ANN 2, $Bi.\beta_0\beta_1\beta_3$, ANN 1, $Bi.\beta_2\beta_3$ and $Bi.\beta_1$ (Table 3). Although, for validation data, it was found that the ANNs performed more modestly than the Bi variations, presenting small proportion differences, similar to the results found by Schikowski *et al.* (2015).

Graphical analysis of the residuals (Figure 2) shows that ANN 3 was less biased towards underestimating the diameters in the upper portion of the shaft, when compared with the variation of the Bi equation ($Bi.\beta_0\beta_1\beta_3$), estimating the diameters in this portion, with greater accuracy.

Özçelik *et al.* (2013) conducted a study for height estimates of juniper trees in the Crimea region of Turkey, using various methods, including ANNs and nonlinear MEM in six growth functions. For the authors, ANN models offered a number of advantages compared to the other methods examined and stood out as a reliable alternative.

One of the great advantages of using MEM and ANNs is that, in addition to providing more accurate estimates, such methodologies allow using data that do not necessarily meet the basic assumptions for regression (OZÇELIK *et al.*, 2011; SCHIKOWSKI *et al.*, 2015), such as taper data that are repeated measurements and correlated within the same individual. Another factor that justifies the use of these techniques is the possibility of reducing the number of sample trees and the number of actual diameter measurements required from each tree. This is of great value to forest companies. Soares *et al.* (2011) emphasized that results obtained with ANNs are quite satisfactory for predicting tree diameter and that it is enough to know only three measurements of real diameters of each tree and to have data from approximately 10% of selected trees of the population.

Although such methodologies are more sophisticated than ordinary least squares modeling, which is commonly used in Brazil, with the advancement of technology, the description of the stem profile using these innovative techniques is no longer a difficulty. However, their application is still very incipient, requiring more studies to make them more widely accepted.

Thus, MEM and ANNs can be considered useful tools for estimating diameters along the stem of *P. taeda* trees to support decision making in the management of the species.

CONCLUSIONS

- The application of the nlme technique in the Bi (2000) equation, as a function of the factors age classes and diameter classes, increased the accuracy of the diameter estimates for *P. taeda* in relation to the adjustment performed using the *nls* technique.
- ANNs generated for stem diameter estimates performed better than the Bi equation variations resulting from the application of mixed-effects modeling indicating that ANNs are efficient for describing the profile of *P. taeda* trees for the area of study.

REFERENCES

- ASHRAF, M. I.; MENG, F. R.; BOURQUE, C. P. A.; MACLEAN, D. A. A Novel Modelling Approach for Predicting Forest Growth and Yield under Climate Change. **PLOS ONE**, United States, v.10, n.7, p. 1 - 18, 2015.
- ASSIS, A. L.; SOARES, J. R. S.; MELLO, J. M.; WEIMAR, F. A. OLIVEIRA, A. D. Comparação de modelos polinomiais segmentados e não segmentados na estimativa de diâmetros e volumes ao longo do fuste de *Pinus taeda*. **Cerne**, Lavras, v. 7, n. 1, p. 20 - 42, 2001.
- BI, H. Trigonometric variable-form taper equations for Australian Eucalypts. **Forest Science**, Washington, v. 46, n. 3, p. 397 - 407, 2000.
- BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. L. Aplicação de Redes Neurais Artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Árvore**, Viçosa, v. 37, n. 4, p. 639 - 645, 2013.
- CAMPOS, B. P. F.; SILVA, G. F.; BINOTI, D. H. B.; MENDONÇA, A. R.; LEITE, H. G. Descrição do perfil do tronco de árvores em plantios de diferentes espécies por meio de Redes Neurais Artificiais. **Pesquisa Florestal Brasileira**, Colombo, v. 37, n. 90, p. 99 - 107, 2017.
- CARVALHO, S. P. C.; RODRIGUEZ, L. C. E.; CALEGARIO, N.; SAVIAN, T. V.; LIMA, M. P. de; SILVA, C. A.; MENDONÇA, A. R.; NICOLETTI, M. F. Modelagem não linear mista para descrever o afilamento de árvores clonais de *Eucalyptus* sp. **Scientia Forestalis**, Piracicaba, v. 42, n. 104, p. 605 - 614, 2014.
- DE-MIGUEL, S.; MEHTÄTALO, L.; SHATER, Z.; KRAID, B.; PUKKALA, T. Evaluating marginal and conditional predictions of taper models in the absence of calibration data. **Canadian Journal of Forest Research**, Ottawa, v. 42, n. 7, p. 1383 - 1394, 2012.
- DIAMANTOPOULOU, M. J. ÖZÇELIK, R.; CRESCENTE-CAMPO, F.; ELER, U. Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. **Biosystems Engineering**, v. 133, p. 33 - 45, 2015.
- GARBER, S. M.; MAGUIRE, D. A. Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. **Forest Ecology and Management**, Amsterdã, v. 179, p. 507 - 522, 2003.
- GOUVEIA, J. F. F.; SILVA, J. A. A.; FERREIRA, R. L. C.; GADELHA, F. H. L.; LIMA FILHO, L. M. A. Modelos volumétricos mistos em clones de Eucalyptus no polo gesso do Araripe, Pernambuco. **Floresta**, Curitiba, PR, v. 45, n. 3, p. 587 - 598, 2015.

- KOHLER, S. V.; KOEHLER, H. S.; FIGUEIREDO FILHO, A. Modelos de Afilamento para *Pinus taeda* por Classes de Idade. **Floresta e Ambiente**, Rio de Janeiro, Seropédica, v. 20, n. 4, p. 470 - 479, 2013.
- LEITE, H. G.; BINOTI, D. H. B.; OLIVEIRA NETO, R. R.; LOPES, P. F.; CASTRO, R. R.; PAULINO, E. J.; BINOTI, M. L. M. S.; COLODETTE, J. L. Redes Neurais Artificiais para a estimação da densidade básica da madeira. **Scientia Forestalis**, Piracicaba, v. 44, n. 109, p. 149 - 154, 2016.
- MARTINS, A. P. M.; MACHADO, S. A.; FIGUEIREDO FILHO, A.; DALLA CORTE, A. P.; GORENSTEIN, M. R. Efeito da idade na forma do fuste de *Araucaria angustifolia* na região Centro-Sul do Paraná. **Pesquisa Florestal Brasileira**, Colombo, v. 37, n. 90, p. 109 - 117, 2017.
- MARTINS, E. R.; BINOTI, M. L. M. S.; LEITE, H. G.; BINOTI, D. H. B.; DUTRA, G. C. Configuração de Redes Neurais Artificiais para estimação do afilamento do fuste de árvores de eucalipto. **Agrária**, Recife, v.11, n.1, p. 33 - 38, 2016.
- MAX, T. A.; BURKHART, H. E. Segmented polynomial regression applied to taper equations. **Forest Science**, Washington, v. 22, n. 3, p. 283 - 289, 1976.
- ÖZÇELİK, R.; DIAMANTOPOULOU, M. J.; CRECENTE-CAMPO, F.; ELER, U. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology and Management**, v. 306, p. 52 - 60, 2013.
- ÖZÇELİK, R.; BROOKS, J. R.; JIANG, L. Modeling stem profile of Lebanon cedar, Brutian pine, and Cilicica fir in Southern Turkey using nonlinear mixed-effects models. **European Journal of Forest Research**, Dordrecht, v. 130, n. 4, p. 613 - 621, 2011.
- PINHEIRO, J. C.; BATES, D. M. **Mixed-effects models in S and S-PLUS**. New York: Springer, 2000, 537p.
- SCHIKOWSKI, A. B.; DALLA CORTE, A. P.; SANQUETA, C. R. Estudo da forma do fuste utilizando Redes Neurais Artificiais e funções de afilamento. **Pesquisa Florestal Brasileira**, Colombo. v. 35, n. 82, p. 119 - 127, 2015.
- SCHÖEPFER, W. **Automatisierung des massen, sorten and wertberechnung stenender waaldbestande schriftenreihe bad**. [S.l.]: Wurttt-Forstl., 1966.
- SCHRÖDER, T.; ARNONI, E. C.; VALÉRIO, A. F.; LISBOA, G. L. Taper Equations for *Pinus elliottii* Engelm. in Southern Paraná, Brazil. **Forest Science**, v. 61, n. 2, p. 311 - 319, 2014.
- SILVA, S.; OLIVEIRA NETO, S. N.; LEITE, H. G.; OBOLARI, A. M. M.; SCHETTINI, B. L. S. Avaliação do uso de regressão e rede neural artificial para modelagem do afilamento do fuste de eucalipto em sistema silvipastoril. **Enciclopédia Biosfera**, Goiânia, v. 13 n. 23; p. 189 - 199, 2016.
- 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**, v. 78, n. 1, p. 19 - 27, 2011.
- TEO, S. J.; MARCON, A.; EHLERS, T.; BIANCHI, J. C.; PELOSO, A.; NAVA, P. R.; COSTA, R. H. Modelos de afilamento para *Pinus elliottii* em diferentes idades, na região de Caçador, SC. **Floresta**, Curitiba, v. 43, n. 3, p. 439 - 452, 2013.
- TRINCADO, G.; BURKHART, H. E. A generalized approach for modeling and localizing stem profiles curves. **Forest Science**, Washington, v. 52, n. 6, p. 670 - 682, 2006.