

Estimativa do Afilamento do Fuste de Araucária Utilizando Técnicas de Inteligência Artificial

Ana Paula Marques Martins¹, Aline Bernarda Debastiani¹,
Allan Libanio Pelissari¹, Sebastião do Amaral Machado¹,
Carlos Roberto Sanquetta¹

¹Departamento de Ciências Florestais, Universidade Federal do Paraná – UFPR, Curitiba/PR, Brasil

RESUMO

O objetivo deste estudo foi comparar o desempenho de técnicas de inteligência artificial com funções de afilamento e avaliar a contribuição da idade nessa estimativa. O conjunto de dados foi composto por 135 observações, abrangendo as idades de 6, 12, 18, 24 e 43 anos de um povoamento de *Araucaria angustifolia*. As funções de afilamento ajustadas foram Kozak modificado, Schöepfer, Hradetzky e Garay. Os modelos de inteligência artificial utilizados foram: RNA e árvore modelo (M5P). Os vetores de entrada foram as mesmas variáveis utilizadas nas equações de afilamento e também o mesmo arranjo com a adição da idade. Foram utilizados 70% dos dados para ajuste e 30% para validação. A função de afilamento de Hradetzky forneceu o melhor ajuste. Dentre os modelos avaliados, a RNA propiciou as melhores estimativas, com destaque para a RNA com adição da variável idade. O desempenho da M5P foi satisfatório, porém, inferior às demais técnicas utilizadas.

Palavras-chave: função de afilamento, M5P, *multi layer perceptron*.

Araucaria Stem Taper or Use of Artificial Intelligence Techniques

ABSTRACT

The aim of this study was to compare the performance of artificial intelligence techniques with taper functions and evaluate the effect of age on this estimate. The data set was comprised of 135 observations covering the ages 6, 12, 18, 24 and 43 years of age of a stand of *Araucaria angustifolia*. Adjusted taper functions were Kozak Schöepfer, Hradetzky and Garay modified. The artificial intelligence models used were: ANN and tree model. The input vectors are the same variables used in taper equations and also the same arrangement with the addition of age. 70% of the data was used for adjustment and 30% for validation. The taper function Hradetzky provided the best fit. Among the models evaluated, the ANN provided the best estimates, highlighting the ANN by adding the age variable. The performance of M5P was satisfactory, however, less effective than the other techniques.

Keywords: taper function, M5P, *multi layer perceptron*.

1. INTRODUÇÃO

Tendo em vista a tendência do setor florestal brasileiro em manejar as florestas visando multiprodutos para maximização da renda, se faz necessária a adoção de

técnicas que possibilitem a quantificação e qualificação desses produtos (Mendonça et al., 2015). Nesse contexto, as funções de afilamento têm se mostrado como

uma boa alternativa, pois possibilitam representar matematicamente a variação diamétrica ao longo do fuste e, conseqüentemente, descrever o perfil longitudinal do tronco das árvores (Sanquetta et al., 2014; Soares et al., 2011). Mediante transformações algébricas na equação, estima-se o volume total ou por seções do fuste (Campos & Leite, 2006), sendo este um ponto importante na previsão do sortimento florestal (Mendonça et al., 2015).

Contudo, as funções de afilamento não consideram a variação da idade e, quando aplicadas em povoamentos multiâneos, é indispensável a realização de testes para verificar a necessidade de estratificação dos dados para estimativas consistentes. Por outro lado, técnicas de Inteligência Artificial (IA) têm se mostrado uma alternativa promissora aos métodos estatísticos convencionais, essas técnicas possibilitam a inclusão de mais variáveis explicativas no treinamento, como a idade, fornecendo ao processo de ajuste mais agilidade sem prejuízo na acuracidade das estimativas, uma vez que eliminam a fase de estratificação dos dados em classes etárias.

Uma das técnicas de IA são as Redes Neurais Artificiais (RNAs), essas têm sido empregadas com êxito visando à estimativa de variáveis dendrométricas, como a altura (Ferreira et al., 2014; Binoti et al., 2013), o volume (Gorgens et al., 2009; Binoti et al., 2014; Cordeiro et al., 2015) e o afilamento (Schikowski et al., 2015; Mendonça et al., 2015), assim como o crescimento e a produção (Castro et al., 2013), dentre outras aplicações. Contudo, há dificuldade no entendimento de como as RNAs propiciam tais estimativas, sendo denominadas um sistema “caixa-preta”, no qual não se conhecem as relações intermediárias entre as variáveis de entrada com a variável a ser estimada, sabendo apenas que esta relação consiste no ajuste dos pesos sinápticos.

Porém, há outras técnicas de IA que oferecem mais clareza de como o processamento é realizado e fornecem exatidão equivalente ou até mesmo superior à obtida pelas RNAs, como as árvores de regressão (Witten et al., 2011), porém essas ainda são pouco exploradas no meio florestal. As árvores de regressão representam uma categoria de árvore de decisão, em que nos nós das folhas são criados modelos de regressão. Elas são construídas com base em regras que seguem da divisão da série de dados em função dos valores dos vetores preditivos (Breiman et al., 1984). Uma versão aperfeiçoada dessas árvores de regressão é a árvore modelo, conhecida como M5 ou M5P (Quinlan, 1993).

Diante do exposto, o objetivo do presente estudo foi comparar o desempenho de técnicas de IA (RNA e M5P) com funções de afilamento, bem como avaliar a contribuição da variável idade, para a estimativa dos diâmetros ao longo do fuste de *Araucaria angustifolia*.

2. MATERIAL E MÉTODOS

2.1. Área de estudo e obtenção dos dados

O banco de dados utilizados neste estudo provém de um reflorestamento de *A. angustifolia* localizado na região Centro-Sul do Estado do Paraná, com espaçamento de 3 m × 2 m. No povoamento, 30 árvores foram selecionadas na idade de 24 anos, de acordo com a distribuição diamétrica, e submetidas posteriormente à análise de tronco completa (ANATRO). Dessa amostra, foram utilizadas as idades restauradas de 6, 12 e 18 e a idade de corte de 24 anos, totalizando 120 árvores. Adicionalmente, 15 árvores com 43 anos foram cubadas em pé.

Na realização da ANATRO, foram coletados os discos às alturas: 0,10; 0,70; 1,30 e a cada 2 m para 15 árvores, e a cada 2,60 m para as outras 15 árvores até o início da copa, e, a partir deste ponto, a cada 1 m até a altura total da árvore. Os discos foram secos e lixados, e, posteriormente, realizadas a marcação e a leitura dos anéis com o auxílio de lupa e do aparelho LINTAB 6.0 (*Frank Rinn Distributors*, Alemanha). A cubagem rigorosa das 15 árvores em pé foi feita com o dendrômetro Criterion RD1000, em que foram mensurados os diâmetros nas alturas: 0,1; 0,70; 1,30 e a cada 2 m até a altura total do fuste.

2.2. Estimativa dos diâmetros com casca

Para a amostra proveniente da ANATRO, os diâmetros com casca (d_{cc}) ao longo do fuste foram recuperados para as idades anteriores ao de corte (24 anos). Para tal, foi ajustado o modelo de Schneider e Silva modificado (Equação 1), conforme descrito por Figueiredo et al. (2015) para essa mesma espécie.

$$\ln d_{cc} = b_0 + b_1 \ln d_{sc} + b_2 \ln(hiI) + \varepsilon_i \quad (1)$$

em que: d_{cc} : diâmetro com casca (cm) na altura hi (m); d_{sc} : diâmetro sem casca (cm) medido na altura hi (m); I : idade (anos); hi : altura (m) onde se encontra o diâmetro di (cm); \ln : logaritmo neperiano; b_i : coeficientes a serem estimados; e ε_i : erros aleatórios.

Para aumentar a densidade de dados para o ajuste desse modelo, informações referentes a 38 árvores de *A. angustifolia* com idades de 25 a 30 anos, oriundas da FLONA de Três Barras, SC, foram incorporadas na base de dados.

2.3. Ajuste de funções de afilamento

Foram ajustadas quatro funções de afilamento de ampla utilização na ciência florestal, conforme apresentadas na Tabela 1. Como o ajuste dos modelos usa dados de várias idades, buscou-se detectar a possibilidade do agrupamento de idades por meio do teste de identidade de modelos proposto por Graybill (1976) e demonstrado por Regazzi (1993). Esse teste foi aplicado para a função de afilamento de melhor ajuste para o conjunto total de dados (6, 12, 18, 24 e 43 anos).

O teste de identidade de modelos é baseado na diferença entre a soma de quadrado dos resíduos do modelo sem diferenciação por idade (modelo reduzido) e a soma de quadrado dos resíduos do ajuste dos modelos por idade (modelo completo), como

observado na Tabela 2. As hipóteses testadas foram: H_0 : o modelo reduzido ajustado para o conjunto total dos dados das cinco classes de idade é estatisticamente igual aos modelos completos ajustados separadamente para cada uma das classes analisadas; e H_1 : rejeita-se H_0 .

O ajuste dos modelos foi realizado com o auxílio do software Excel 2013 e dos suplementos Solver e Action Stat. Para o modelo de Hradetzky, foram selecionadas as potências do polinômio por *Stepwise*, testando as potências: 0,005; 0,09; 0,08; 0,07; 0,06; 0,05; 0,04; 0,03; 0,02; 0,01; 0,9; 0,8; 0,7; 0,6; 0,5; 0,4; 0,3; 0,2; 0,1; 1; 2; 3; 4; 5; 10; 15; 20 e 25. Ademais, foi realizado o ajuste dos modelos com 70% dos dados, e a validação, com 30% restantes a fim de comparação com as técnicas de IA utilizadas.

2.4. Estimativa do diâmetro ao longo do fuste utilizando IA

As técnicas de IA utilizadas foram: RNA e árvore de regressão. A RNA utilizada foi a *Multi Layer Perceptron* (MLP) aplicada no software Matlab 2014a

Tabela 1. Funções de afilamento ajustadas para povoamento de *A. angustifolia* localizada no Centro-Sul do Estado do Paraná.

Table 1. Stem taper functions fitted to stands of *A. angustifolia* located in Central-Southern of state Parana.

Autor	Modelo
Kozak modificado	$\left(\frac{di}{dap}\right)^2 = b_1 \left[\left(\frac{hi}{ht}\right) - 1\right] + b_2 \left[\left(\frac{hi}{ht}\right)^2 - 1\right]$
Schöepfer	$\frac{di}{dap} = b_0 + b_1 \left(\frac{hi}{ht}\right) + b_2 \left(\frac{hi}{ht}\right)^2 + b_3 \left(\frac{hi}{ht}\right)^3 + b_4 \left(\frac{hi}{ht}\right)^4 + b_5 \left(\frac{hi}{ht}\right)^5$
Hradetzky	$\frac{di}{dap} = b_0 + b_1 \left(\frac{hi}{ht}\right)^{p_1} + b_2 \left(\frac{hi}{ht}\right)^{p_2} + b_3 \left(\frac{hi}{ht}\right)^{p_3} + \dots + b_n \left(\frac{hi}{ht}\right)^{p_n}$
Garay	$\frac{di}{dap} = b_1 \left[1 + b_2 \ln \left(1 - b_3 \left(\frac{hi}{ht}\right)^{b_4}\right)\right]$

Em que: *di*: diâmetro da seção *i* (cm); *dap*: diâmetro à altura do peito, medido a 1,30 m do solo (cm); *hi*: altura da seção *i*; *ht*: altura total (m); *b_i*: coeficientes a serem estimados; *p_i*: potências a serem selecionadas para compor o polinômio de potências fracionárias; e ln: logaritmo neperiano.

Tabela 2. Análise de variância (ANOVA) para o teste de identidade de modelos.

Table 2. Analysis of variance (ANOVA) for the models identity test.

F. V.	G.L.	S.Q.	Q.M.	F
Reg. Completo	ΣGLRegComp (1)	ΣSQRegComp (5)		
Reg. Reduzido	GLRegRed (2)	SQRegRed (6)		
Redução (H_0)	(3) = (1) - (2)	(7) = (5) - (6)	SQRed/GL	QMRRed/QMRes
Resíduo	(4) = (3) - (1)	(8) = (4) - (5)	SQRes/GL	
Total	ΣGLTotComp (3)	ΣSQTotComp (4)		

Em que: F.V.: Fonte de variação; G.L.: graus de liberdade; S.Q.: soma de quadrados; Q.M.: quadrado médio; F: valor de F calculado; RegComp: regressão ajustada por idade; e RegRed: regressão ajustada para o modelo reduzido.

no *Neural Network Toolbox*, algoritmo de aprendizado de retropropagação de *Levenberg-Marquardt* (Hagan & Menhaj, 1994). Para isso, utilizou-se a taxa de aprendizado de 0,01 e a função de ativação tangente hiperbólica sigmoidal. Os vetores de entrada foram padronizados para uma mesma escala, entre 0 e 1, a fim de melhorar a convergência (Fu, 1994).

A MLP foi constituída de uma camada oculta, na qual o número de neurônios variou em ± 5 em torno de um número de neurônios ótimo. O número de neurônios ótimo, de acordo com Heath (2010), é uma relação inicial de 10 vezes mais equações de treinamento do que de pesos. Como critério de parada do treinamento da MLP e para evitar o superajuste, utilizou-se um conjunto de dados distintos para realizar a técnica de validação cruzada.

O algoritmo de árvore de regressão utilizado foi a árvore modelo, disponível no *software* WEKA (*Waikato Environment for Knowledge Analysis*) como M5P, sendo o treinamento realizado em modo *Percentage Split*. As árvores foram geradas a partir de um conjunto de dados de treinamento, sendo que, em cada nó, o algoritmo escolheu os atributos que mais eficientemente subdividiram o conjunto das amostras em subconjuntos homogêneos e caracterizados por sua classe. O critério foi o ganho de informação obtido na escolha do atributo para subdivisão (Quinlan, 1993).

O conjunto total de dados foi dividido em três subconjuntos que contemplaram todas as idades, com proporção de aproximadamente 55%, 15% e 30% para treinamento, validação cruzada e teste, respectivamente. Estas proporções foram as mesmas usadas para a MLP e M5P.

A escolha dos vetores de entrada da RNA e da M5P foi baseada nas variáveis utilizadas nas equações de afilamento e na correlação entre as variáveis independentes dap, ht, hi e idade, com a variável dependente di. Desse modo, foram avaliadas duas configurações de vetores de entrada da RNA e M5P para a estimativa do diâmetro da seção *i*, sendo: RNA 1 e M5P 1: dap, ht, hi; RNA 2 e M5P 2: dap, ht, hi e idade.

2.5. Avaliação dos modelos

A avaliação do desempenho das funções de afilamento se deu a partir da análise das estatísticas de ajuste e seleção de modelos: coeficiente de determinação ajustado (R^2_{aj}) e erro padrão da estimativa (S_{yx} %). Selecionada a função de afilamento de melhor desempenho, esta

foi comparada com as demais técnicas de IA. Para tal foram utilizados os indicadores de qualidade de ajuste: coeficiente de correlação (*r*), viés (V%), acurácia (A%), precisão (P%) e raiz do erro médio quadrático (RMSE%), conforme descritos na Tabela 3. Para a avaliação de todos os modelos foi realizada a distribuição gráfica dos resíduos ordinários, que consiste na diferença absoluta entre os valores estimados e observados.

3. RESULTADOS E DISCUSSÃO

3.1. Estatística descritiva dos dados

Na Tabela 4, é apresentada a estatística descritiva dos dados utilizados em cada idade de avaliação do povoamento de *A. angustifolia*. Notou-se que a idade mais jovem, de seis anos, apresentou a maior dispersão dos dados, tanto para dap quanto para ht; isto pode estar associado a maior irregularidade no crescimento em decorrência da falta de competição.

Tabela 3. Indicadores da qualidade de ajuste utilizados para avaliar o desempenho dos modelos de estimativa do diâmetro ao longo do fuste para *A. angustifolia*.

Table 3. Quality indicators from adjust used to evaluate the performance of the diameter estimation models along stem for *A. angustifolia*.

Estatística	Fórmula
<i>r</i>	$r = \sqrt{1 - \frac{\sum (d_{i_{est}} - d_{i_{obs}})^2}{\sum (d_{i_{obs}} - \bar{d}_{i_{obs}})^2}}$
V%	$V\% = \frac{\left(\frac{\sum (d_{i_{est}} - d_{i_{obs}})}{n} \right)}{\bar{d}_{i_{obs}}} * 100$
A%	$A\% = \sqrt{\frac{\left(\frac{\sum (d_{i_{est}} - d_{i_{obs}})^2}{n-1} \right)}{\bar{d}_{i_{obs}}}} * 100$
P%	$P\% = \sqrt{\frac{\left(\frac{\sum (d_{i_{est}} - V - d_{i_{obs}})^2}{n-1} \right)}{\bar{d}_{i_{obs}}}} * 100$
RMSE%	$RMSE\% = \sqrt{\frac{\frac{1}{n} \sum (d_{i_{obs}} - d_{i_{est}})^2}{\max_{obs} - \min_{obs}}} * 100$

Em que: $d_{i_{est}}$: diâmetro na seção *i* estimado; $d_{i_{obs}}$: diâmetro observado na seção *i*; *n*: número de observações.

A correlação linear de Pearson entre as variáveis foi calculada a fim de determinar as relações existentes entre elas (Tabela 5). Com isso, notou-se que o dap foi a variável mais correlacionada com a maioria das variáveis, inclusive para a variável di, indicando a importância do dap na estimativa do afilamento do fuste.

Ainda analisando as correlações com a variável di, nota-se que subsequente ao dap, a idade é a variável que mais se correlaciona (r = 0,535), isso demonstra o seu potencial como vetor de entrada para a estimativa do di por meio das técnicas de IA.

3.2. Estimativa dos diâmetros com casca

Os resultados propiciados pela equação de Schneider e Silva (modificado) ajustada (Equação 2) foram satisfatórios para a estimativa dos diâmetros com casca ao longo do fuste ($R^2_{aj.} = 0,995$ e $S_{yx}\% = 2,66\%$), além de boa distribuição gráfica dos resíduos (Figura 1).

$$\ln dcc = 0,31185 + 0,94946 \ln dsc - 0,0102 \ln(hiI) + \varepsilon_i \quad (2)$$

3.3. Estimativa dos diâmetros ao longo do fuste (di)

Os coeficientes de ajuste das equações de afilamento e arquitetura da RNA utilizada estão apresentados na Tabela 6, em que as potências selecionadas pelo método

Stepwise para o polinômio de Hradetzky foram: 0,005; 0,3; 3; 4 e 10.

A arquitetura da RNA 1 foi composta por três neurônios na camada de entrada, compreendendo os vetores referentes à ht, hi e dap, em que a melhor iteração da RNA foi constituída por 13 neurônios na camada oculta, sendo ela a camada responsável pelo processamento das informações, e, ainda, um neurônio na camada de saída referente ao di. Já a arquitetura da RNA 2 foi composta por quatro neurônios na camada de entrada, na qual adicionou-se a variável idade às demais anteriormente utilizadas na RNA 1, cuja iteração que produziu as melhores estimativas foi composta por 10 neurônios na camada oculta e 1 neurônio na camada de saída.

Pela análise das estatísticas de ajuste e seleção de modelos, a equação com o melhor desempenho foi Hradetzky, uma vez que ela propiciou o maior $R^2_{aj.}$ e o menor $S_{xy}\%$, sendo iguais a 0,979 e 9,798%, respectivamente. A qualidade dos melhores ajustes foi seguida pelos modelos de Schöepfer, Garay e Kozak modificado, os quais proporcionaram resultados muito semelhantes entre si, com ligeira desvantagem para a função de Kozak modificado, indicando que elas podem ser utilizadas para o conjunto total de dados sem que haja maiores perdas na acurácia das estimativas.

Nesse sentido, diversos autores confirmaram a superioridade das funções de Hradetzky e do polinômio de quinto grau de Schöepfer na avaliação do afilamento para algumas espécies do gênero *Pinus* no Brasil (Araújo et al., 2012; Yoshitani et al., 2012; David et al., 2014). Além disso, Loureiro et al. (2012), ao avaliarem equações de afilamento não lineares para um plantio de *A. angustifolia* com 39 anos de idade no município

Tabela 4. Valores médios e Coeficiente de Variação (CV %) para as variáveis dendrométricas por idade.

Table 4. Mean values and coefficient of variation (CV %) for tree measurement variables by age.

Idade (anos)	n	dap (cm) (CV %)	ht (m) (CV %)
6	30	10,78 (33,98)	8,36 (26,54)
12	30	20,62 (21,04)	15,22 (12,58)
18	30	26,46 (21,31)	18,11 (10,30)
24	30	30,66 (23,84)	19,57 (2,14)
43	15	45,67 (23,33)	22,09 (11,76)

Tabela 5. Correlação linear de Pearson entre as variáveis dendrométricas utilizadas.

Table 5. Pearson linear correlation between the tree measurement variables used.

	Idade (anos)	dap (cm)	ht (m)	hi (m)	di (cm)
Idade	1,000	-	-	-	-
dap (cm)	0,816	1,000	-	-	-
ht (m)	0,754	0,836	1,000	-	-
hi (m)	0,226	0,262	0,331	1,000	-
di (cm)	0,535	0,626	0,515	-0,524	1,000

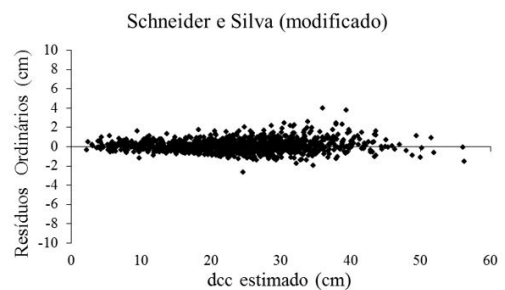


Figura 1. Distribuição gráfica dos resíduos da equação de Schneider e Silva (modificado) ajustada.

Figure 1. Graphical distribution of residuals from Schneider and Silva adjusted equation (modified).

de Rio Negro - PR, concluíram que a equação de Garay foi a que melhor estimou os diâmetros ao longo do fuste ($R^2_{aj.} = 0,9896$ e $S_{yx} = 6,93\%$), confirmando o bom desempenho dessa equação para a espécie.

Adicionalmente, o teste de identidade de modelos foi aplicado ao polinômio de Hradetzky para os ajustes por idade (modelo completo) e para o agrupamento das idades (modelo geral ou reduzido), cuja ANOVA é apresentada na Tabela 7. Com isso, foi observada diferença estatística significativa, ao nível de 5% de probabilidade de erro, entre os modelos completo e reduzido, o que indica a necessidade dos ajustes da função de Hradetzky por classes de idade (Tabela 8).

De acordo com a Tabela 8, foi possível detectar a melhoria na homogeneidade da forma do fuste com o avanço da idade, pois o $S_{yx} \%$ diminuiu e o $R^2_{aj.}$ aumentou para os dados provenientes do mesmo plantio entre as idades de 6 a 24 anos, sugerindo que a estratificação por idade resultou em melhoria na acuracidade das estimativas. Contudo, o polinômio de Hradetzky não apresentou bom ajuste apenas para a classe de idade de 6 anos, com o menor $R^2_{aj.}$ entre as classes de idade, sendo igual a 0,947; $S_{yx} \%$ superior a 14%, respectivamente.

De acordo com Araújo (2015), a forma do fuste e, conseqüentemente, sua altura e volume são afetados pela

Tabela 6. Coeficientes, estatísticas de ajuste e seleção das funções de afilamento e arquitetura da RNA.

Table 6. Coefficients, fit statistics and selection of the stem taper functions and RNA architecture.

Modelo	b_0	b_1	b_2	b_3	b_4	b_5	$R^2_{aj.}$	$S_{yx} \%$
Kozak mod.	-	-1,662*	0,380*	-	-	-	0,970	11,673
Schöepfer	1,177*	-2,822*	12,688*	-30,588*	32,911*	-13,368*	0,979	9,867
Hradetzky	1,178*	0,113*	-0,567*	-0,903*	0,534*	-0,354*	0,979	9,798
Garay	-	1,264*	0,221*	0,989*	0,224*	-	0,975	10,550
Arquitetura (entrada/oculta/saída)								
RNA 1	3/13/1							
RNA 2	4/10/1							

*Significativo a 5% de probabilidade de erro.

Tabela 7. ANOVA para o teste de identidade de modelos para o polinômio de Hradetzky aplicada à estimativa de diâmetros ao longo do fuste de *A. angustifolia*.

Table 7. ANOVA for the model identity test for the Hradetzky polynomial applied to estimate diameters along the stem of *A. angustifolia*.

F.V.	G.L.	S.Q.	Q.M.	F
Mod. Completo	30	150,66		
Mod. Reduzido	6	149,28		
Diferença	24	1,37	0,0571	8,922*
Resíduos	1,128	7,26	0,0064	
Total	1,158	157,9		

*Significância a 5% de probabilidade de erro.

Tabela 8. Coeficientes e estatísticas de seleção para o polinômio de Hradetzky ajustado para classes de idade de *A. angustifolia*.

Table 8. Coefficients and statistics of selection for Hradetzky polynomial adjusted for age classes of *A. angustifolia*.

Idade (anos)	b_0	b_1	b_2	b_3	b_4	b_5	$R^2_{aj.}$	$S_{yx} \%$
6	1,2258	0,1626	-0,5990	-4,2519	4,2725	-0,8093	0,947	14,144
12	1,1632	0,1157	-0,5455	-0,7698	0,0058	-0,0345	0,981	8,191
18	1,1713	0,0992	-0,5444	-0,6917	0,2714	-0,3024	0,982	7,580
24	1,1740	0,1083	-0,5900	0,1924	-0,6125	-0,2724	0,984	6,904
43	1,1283	0,1085	-0,5460	-0,1854	-0,2574	-0,2448	0,966	8,987
Geral	1,1777	0,1128	-0,5674	-0,9033	0,5343	-0,3543	0,979	9,798

idade do povoamento. Diante disso, não é apropriado utilizar uma equação geral para descrever o perfil do fuste de árvores em várias idades, pois provavelmente as estimativas obtidas serão enviesadas.

A função de afilamento ajustada sem estratificação dos dados (modelo geral) de melhor desempenho foi utilizada para a comparação com as técnicas de IA, uma vez que o ajuste de tais técnicas foi realizado utilizando-se o banco de dados que contempla todas as classes de idades, cujas estatísticas de desempenho realizadas sobre os modelos, tanto no ajuste ou treinamento quanto para validação ou teste, são apresentadas na Tabela 9.

Na validação ou teste, a correlação (r) entre o valor observado e estimado foi superior a 0,987 (M5P 2); o RMSE, usado para representar a amplitude do erro, foi inferior a 3,561% (M5P2); o viés expressa o erro sistemático ou tendenciosidade do ajuste e foi inferior a 2,330% (RNA 1); considerando a acurácia como sendo a medida de afastamento entre o valor observado e o estimado, essa foi inferior a 10,785% (M5P 2) e a precisão que representa a dispersão do valor estimado, foi inferior a 10,556% (M5P 2).

Ademais, pelo ranking realizado com os valores das estatísticas para o conjunto de dados de teste, a RNA 2, que inclui a variável idade, apresentou-se como o melhor modelo, seguido da RNA 1, a qual utilizou as mesmas variáveis das equações de afilamento. Isso demonstrou a contribuição da idade para a estimativa do diâmetro ao longo do fuste. Comportamento contrário foi observado pela M5P, na qual a ausência da variável idade resultou no melhor desempenho.

Os resultados propiciados pelas estimativas da M5P foram satisfatórios, ainda que inferiores aos da RNA

e da função de afilamento selecionada. Contudo, não foram encontrados estudos na literatura especializada que utilizaram essa técnica para estimativa de variáveis dendrométricas, porém, quando aplicadas em outras áreas, como modelagem do processo chuva-vazão (Solomatine & Dulal, 2003; Etemad-Shahidi & Mahjoobi, 2009), de sedimentos (Bhattacharya & Solomatine, 2006) e evapotranspiração (Rahimikhoob, 2014), o desempenho foi similar ao das RNAs.

Em um estudo, Araújo (2015), ao comparar o desempenho de uma RNA com as equações de afilamento para pau-mulato (*Calicophyllum spruceanum* Benth.), observou que o modelo de Garay ($R^2 = 0,943$, RMSE = 8,8%) resultou no melhor ajuste, sendo, contudo, inferior ao obtido pela RNA ($R^2 = 0,952$, RMSE = 7,6%) com a variável idade nos vetores de entrada. Esse autor afirmou, ainda, que as RNAs têm como vantagem a flexibilidade de inserção de mais variáveis explicativas simultaneamente, como a idade do povoamento.

Outros estudos, como os de Schikowski et al. (2015), obtiveram desempenho superior ao utilizar RNAs ($R^2 = 0,988$, RMSE = 5,302%), ao compará-lo com as equações de Hradetzky ($R^2 = 0,990$, RMSE = 6,230%) e de Garay ($R^2 = 0,988$, RMSE = 5,483%) para a estimativa do diâmetro relativo de *Eucalyptus* sp. Além disso, Mendonça et al. (2015) compararam o desempenho de uma RNA com o modelo de Schöepfer para a estimativa dos diâmetros ao longo do fuste para *Eucalyptus dunni*, e concluíram que o desempenho dos dois foi satisfatório, com uma pequena vantagem para a RNA.

Neste estudo, a dispersão dos resíduos ordinários foi homogênea e sem tendenciosidades na estimativa dos diâmetros ao longo do fuste, como observado na

Tabela 9. Indicadores de qualidade de ajuste dos modelos na fase de ajuste ou treinamento e validação ou teste. **Table 9.** Quality indicators of adjustment of the models in the adjustment phase or training and validation or testing.

Fase	Modelo	r	RMSE (%)	V (%)	A (%)	P (%)	Ranking
Ajuste / Treino	Hradetzky - Geral	0,989	2,493	0,393	9,777	9,769	-
	RNA 1	0,992	1,641	0,181	7,361	7,358	-
	RNA 2	0,994	1,919	-0,428	6,920	6,907	-
	M5P 1	0,951	5,976	-12,828	24,647	21,041	-
	M5P 2	0,975	4,109	7,397	16,948	15,246	-
Validação / Teste	Hradetzky - Geral	0,991	2,895	0,692	8,745	8,718	3
	RNA 1	0,997	2,860	2,330	8,569	8,246	2
	RNA 2	0,997	2,811	0,903	8,458	8,410	1
	M5P 1	0,988	3,354	1,621	10,158	10,027	4
	M5P 2	0,987	3,561	2,207	10,785	10,556	5

Sendo r: coeficiente de correlação; RMSE (%): raiz do erro médio quadrático; V (%): Viés; A (%): acurácia; P (%): precisão.

Figura 2. Estudos como os de Schikowski et al. (2015), Mendonça et al. (2015) e Araújo (2015), utilizando resíduos relativos, observaram maiores erros na porção superior do fuste, ou seja, nos menores diâmetros. Já nos resultados encontrados por Souza (2013), as estimativas pela RNA e equação de Garay apresentaram

erros percentuais acentuados na base dos fustes, nos maiores diâmetros.

Com isso, de maneira geral, as estimativas geradas pela RNA com a adição da idade propiciaram as melhores estatísticas de ajuste, contudo, comparado aos demais modelos, esta superioridade não foi muito

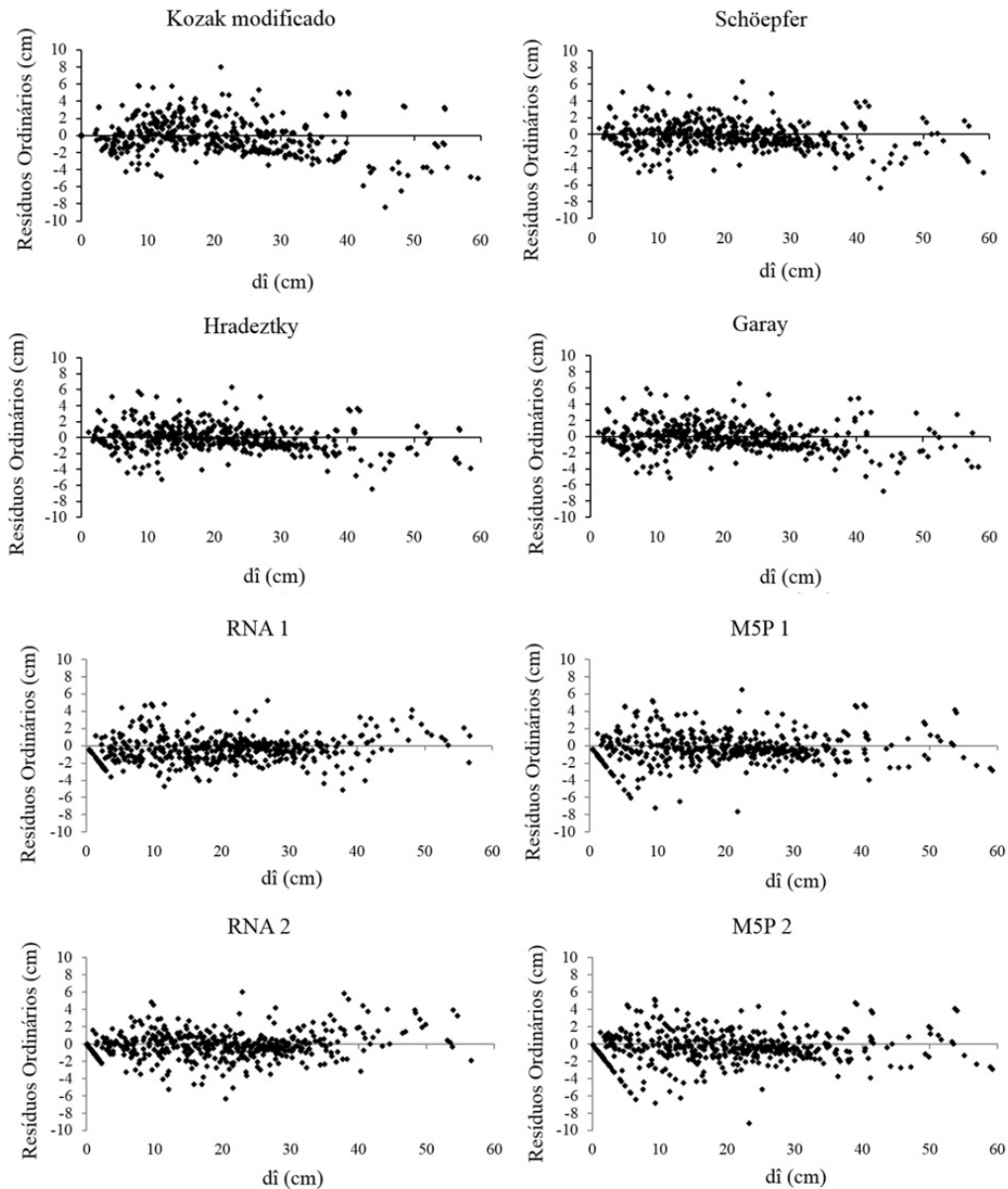


Figura 2. Dispersão gráfica dos resíduos gerados pelo ajuste na fase de teste das funções de afilamento e pelas técnicas de IA.

Figure 2. Graphical dispersion of residuals generated by adjusting in the phase of test for the taper functions and the AI techniques.

acentuada. Porém, vale ressaltar que a maior vantagem do uso dessa técnica de IA é a possibilidade de inserção de variáveis explicativas, como a idade que neste caso eliminou a necessidade de estratificação dos dados em classes etárias. Conforme destacado por Sanquetta et al. (2015), que afirmam que as técnicas de IA de modo geral têm a vantagem de serem versáteis e flexíveis, além de não precisar atender a alguns pressupostos da regressão, tais como normalidade dos dados e homocedasticidade da variância.

Dessa forma, essas técnicas de IA são ferramentas úteis para a estimativa do diâmetro ao longo do fuste e o posterior volume acurado das árvores, desempenhando, assim, suporte à tomada de decisão no manejo da *A. angustifolia*.

4. CONCLUSÕES

A RNA 2 (com adição da variável idade) propiciou resultados superiores aos demais modelos avaliados. Ao utilizar as mesmas variáveis de entrada que a função de afilamento, a RNA ainda apresentou resultado superior às demais técnicas.

Não foi observada contribuição da variável idade no desempenho obtido com a M5P. O desempenho da M5P foi satisfatório, porém, inferior às demais técnicas utilizadas, sendo recomendado maiores estudos e aplicações com essa técnica.

Como a idade é um fator que afeta a forma das árvores, recomenda-se o emprego de redes neurais artificiais com a inclusão dessa variável de entrada, visando ao estudo do afilamento do fuste e à estimativa da produção volumétrica no manejo dos plantios comerciais de *A. angustifolia*.

STATUS DA SUBMISSÃO

Recebido: 10 maio, 2016

Aceito: 19 out., 2016

AUTOR(ES) PARA CORRESPONDÊNCIA

Ana Paula Marques Martins

Departamento de Ciências Florestais,
Universidade Federal do Paraná – UFPR, Av.
Prof. Lothário Meissner, 632, Jardim Botânico,
Campus III, CEP 80210-170, Curitiba, PR, Brasil
e-mail: anapaula_marquesm@yahoo.com.br

REFERÊNCIAS

- Araújo BHP. *Modelagem da altura, volume e afilamento do fuste de Calycophyllum spruceum Benth. empregando regressão e redes neurais artificiais* [dissertação] Manaus: Instituto Nacional de Pesquisas da Amazônia; 2015.
- Araújo EJG, Pelissari AL, David HC, Miranda ROV, Péllico S No, Morais VA et al. Relações dendrométricas em fragmentos de povoamentos de pinus em Minas Gerais. *Pesquisa Florestal Brasileira de 2012*; 32(72): 355-366. <http://dx.doi.org/10.4336/2012.pfb.32.72.355>.
- Bhattacharya B, Solomatine DP. Machine learning in sedimentation modelling. *Neural Networks* 2006; 19(2): 208-214. PMID:16530383. <http://dx.doi.org/10.1016/j.neunet.2006.01.007>.
- Binoti DHB, Binoti MLMS, Leite HG. Configuração de redes neurais artificiais para estimação do volume de árvores. *Revista Ciência da Madeira* 2014; 5(1): 1-6.
- Binoti MLMS, Binoti DHB, Leite HG. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. *Revista Árvore* 2013; 37(4): 639-645. <http://dx.doi.org/10.1590/S0100-67622013000400007>.
- Breiman L, Friedman JH, Olshen RA, Pedra CJ. *Classification and regression trees*. Boca Raton: Chapman & Hall/CRC; 1984. 358 p.
- Campos JCC, Leite HG. *Mensuração florestal: perguntas e respostas*. 2. ed. Viçosa: Universidade Federal de Viçosa; 2006.
- Castro RVO, Soares CPB, Martins FB, Leite HG. Crescimento e produção de plantios comerciais de eucalipto estimados por duas categorias de modelos. *Pesquisa Agropecuária Brasileira* 2013; 48(3): 287-295. <http://dx.doi.org/10.1590/S0100-204X2013000300007>.
- Cordeiro MA, Pereira NNJ, Binoti DHB, Binoti MLMS, Leite HG. Estimativa do volume de *Acacia mangium* utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. *Pesquisa Florestal Brasileira* 2015; 35(83): 255-261. <http://dx.doi.org/10.4336/2015.pfb.35.83.596>.
- David HC, Marinheski A Fo, Pelissari A, Péllico S No, Araújo E, Baum L. Critérios de estratificação para o ajuste de funções de afilamento em fustes de pinus. *Pesquisa Florestal Brasileira de 2014*; 34(79): 197-206. <http://dx.doi.org/10.4336/2014.pfb.34.79.659>.
- Etemad-Shahidi A, Mahjoobi J. Comparison between M5' model tree and neural networks for prediction of significant wave height in lake superior. *Ocean Engineering* 2009; 36(15-16): 1175-1181. <http://dx.doi.org/10.1016/j.oceaneng.2009.08.008>.
- Ferreira JCB, Lafetá BO, Penido TMA, Campos PM, Castro PM. Altura de mudas de *Tibouchina granulosa* cogn. (melastomataceae) estimada por redes neurais artificiais. *REVESBAU* 2014; 9(1): 151-160.

- Figueiredo A Fo, Retslaff FAS, Kohler SV, Becker M, Brandes D. Efeito da idade no afilamento e sortimento em povoamentos de *Araucaria angustifolia*. *Floresta e Ambiente* 2015; 22(1): 50-59. <http://dx.doi.org/10.1590/2179-8087.080114>.
- Fu L. *Neural networks in computer intelligence*. New York: McGraw-Hill; 1994.
- Gorgens EB, Leite HG, Santos HN, Gleriani JM. Estimativa do volume de árvores utilizando redes neurais artificiais. *Revista Árvore* 2009; 33(6): 1141-1147. <http://dx.doi.org/10.1590/S0100-67622009000600016>.
- Graybill FA. *Theory and application of the linear model*. Belmont: Duxbury Press; 1976.
- Hagan MT, Menhaj M. Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* 1994; 5(6): 989-993. PMID:18267874. <http://dx.doi.org/10.1109/72.329697>.
- Heath GE. *Training, testing and validating data set in Neural Network*. USA: MathWorks; 2010. [citado em 2013 jul. 4]. Disponível em: http://www.mathworks.com/matlabcentral/newsreader/view_thread/295781#917734
- Loureiro GH, Curto RA, Rosot NC, Marangon GP. Avaliação de equações de afilamento em um plantio de *Araucaria angustifolia* (Bertol.) Kuntze utilizando o diâmetro como variável dependente. In: *Anais do 4º Congresso Florestal Paranaense: Manejo de Florestas Plantadas*. Curitiba; 2012.
- Mendonça NP, Carvalho MC, Gomide LR, Ferraz AC Fo, Ferreira MA. Previsão de diâmetros ao longo do fuste de eucalipto via redes neurais artificiais. *Enciclopédia Biosfera de 2015*; 11(22): 2419-2429. http://dx.doi.org/10.18677/Enciclopedia_Biosfera_2015_212.
- Quinlan JR. *C4.5: Programs for machine learning*. São Francisco: Morgan Kaufmann; 1993.
- Rahimikhoob A. Comparison between M5 model tree and neural networks for estimating reference evapotranspiration in an arid environment. *Water Resources Management* 2014; 28(3): 657-669. <http://dx.doi.org/10.1007/s11269-013-0506-x>.
- Regazzi AJ. Teste para verificar a identidade de modelos de regressão e a igualdade de alguns parâmetros num modelo polinomial ortogonal. *Ceres* 1993; 40: 176-195.
- Sanquetta CR, Corte APD, Rodrigues AL, Watzlawick LF. *Inventários florestais: planejamento e execução*. 3. ed. revista e ampliada. Curitiba: Multi-Gráfico Gráfica e Editora; 2014.
- Sanquetta CR, Wojciechowski J, Corte APD, Behling A, Péllico S No, Rodrigues AL et al. Comparison of data mining and allometric model in estimation of tree biomass. *BMC Bioinformatics* 2015; 16(1): 247. PMID:26250142. <http://dx.doi.org/10.1186/s12859-015-0662-5>.
- Schikowski AB, Corte APD, Sanquetta CR. Estudo da forma do fuste utilizando redes neurais artificiais e funções de afilamento. *Pesquisa Florestal Brasileira* de 2015; 35(82): 119-127. <http://dx.doi.org/10.4336/2015.pfb.35.82.867>.
- Soares CPB, Paula F No, Souza AL. *Dendrometria e Inventário Florestal*. 2. ed. Viçosa: Editora UFV; 2011.
- Solomatine DP, Dulal KN. Model trees as an alternative to neural networks in rainfall-runoff modeling. *Hydrological Sciences Journal* 2003; 48(3): 399-411. <http://dx.doi.org/10.1623/hysj.48.3.399.45291>.
- Souza RR. *Estudo da forma do fuste de árvores de eucalipto em diferentes espaçamentos* [dissertação]. Diamantina: Universidade Federal dos Vales do Jequitinhonha e Mucuri; 2013.
- Witten IH, Frank E, Hall MA. *Data mining practical machine learning tool and techniques*. 3. ed. Elsevier. 2011.
- Yoshitani M Jr, Nakajima NY, Arce JE, Machado AS, Druszcz JP, Hosokawa RT et al. Funções de afilamento para plantios desbastados de *Pinus taeda*. *Floresta de 2012*; 42(1): 169-176. <http://dx.doi.org/10.5380/ufv.v42i1.26315>.