



ESTIMATIVA DA ALTURA POR MEIO DE REGRESSÃO E REDES NEURAIS ARTIFICIAIS PARA *Pinus taeda* L.

Gabrielle Hambrecht Loureiro¹; Carlos Roberto Sanquetta²; Renato Vinicius Oliveira Castro³; Julio Eduardo Arce²

1 - Doutoranda em Engenharia Florestal pela Universidade Federal do Paraná
(ghloureiro@yahoo.com.br)

2 - Eng. Florestal, Dr., Professor da Universidade Federal do Paraná

3 - Eng. Florestal, Dr., Professor da Universidade Federal de São João Del-Rei
Av. Lothário Meissner, 900 – Jardim Botânico
Curitiba – PR, 80.210-170 – Brasil

Recebido em: 03/10/2016 – Aprovado em: 21/11/2016 – Publicado em: 05/12/2016
DOI: 10.18677/EnciBio_2016B_046

RESUMO

A precisão no ajuste dos modelos hipsométricos é afetada diretamente por fatores ambientais como capacidade produtiva, idade, material genético, variações ambientais, entre outros. Uma alternativa para representar as interações não lineares em relação aos métodos estatísticos tradicionais tem sido a aplicação de redes neurais artificiais (RNA). Por isso, o objetivo deste trabalho foi testar as RNA para estimação da altura em *Pinus taeda*, avaliar sua aplicabilidade e compará-la com os métodos tradicionais de regressão. Os modelos de regressão apresentaram melhores estimativas de altura que as RNA. Ou seja, quando são utilizadas as mesmas variáveis de entrada para a regressão e RNA, empregar modelos hipsométricos estratificados geram estimativas mais precisas.

PALAVRAS-CHAVE: altura, interações não lineares, modelos de regressão.

HYPSONETRIC RELATIONSHIP BY USING REGRESSION MODELS AND ARTIFICIAL NEURAL NETWORKS FOR *Pinus taeda* L.

ABSTRACT

Precision of fitting of hypsometric models is directly affected by environmental factors, such as yield potential, age, genetic material, environmental variations, among others. An alternative method to represent the non-linear interactions in relation to traditional statistical methods is the application of artificial neural networks (ANN). The objective of this study was to test the ANN to estimate *Pinus taeda* L. height, evaluate its applicability and compare it with traditional regression methods. Regression models showed better height estimates that ANN methods. In other words, when the same input variables are used for regression and ANN, the stratified hypsometric models generate more precise estimates.

KEYWORDS: height, regression models, nonlinear interactions.

INTRODUÇÃO

A quantificação das alturas das árvores é fundamental no inventário florestal, pois é utilizada para a estimativa do volume de madeira (THIERSCH et al., 2013) e em determinadas situações, pode servir como indicadora da qualidade produtiva de um local. Essa metodologia é bastante utilizada em florestas equiâneas e apresenta boa precisão em razão da homogeneidade dos plantios (SILVA et al., 2012).

Por ser uma medida indireta, é sujeita a erros, e normalmente implica maior tempo para sua quantificação quando comparada com o diâmetro, fato que torna a operação onerosa (THIERSCH et al., 2013).

A modelagem da altura para povoamentos florestais é bastante avançada e dissipada no Brasil por meio de modelos lineares e não lineares (SANQUETTA et al., 2014). Apesar disso, as empresas florestais enfrentam um grande gargalo no momento do ajuste dos modelos, mais especificamente na estratificação dos dados para a modelagem e precisão da variável dependente desejada, tendo em vista que a relação hipsométrica sofre influência de diversos fatores, como o sítio, a idade, entre outros (MACHADO et al., 2008; VENDRUSCOLO et al., 2015).

A grande dificuldade da escolha do melhor modelo para representar essas relações hipsométricas ocorre devido a não linearidade da relação entre as variáveis envolvidas e as restrições dos parâmetros dos modelos, por razões biológicas. Esses dois fatores aliados tornam o ajuste de modelos estatísticos para representar as relações hipsométricas um problema de regressão não linear, com restrições nos parâmetros (THIERSCH et al., 2013).

Diversos trabalhos foram desenvolvidos para estimação de altura com modelos empíricos, porém, a precisão no ajuste dos modelos hipsométricos é afetada diretamente por fatores ambientais como capacidade produtiva, idade, material genético, variações ambientais, entre outros (ZHANG et al., 1997).

Segundo BINOTI et al. (2013), a inclusão dessas características nos modelos tradicionais, pode resultar em estimativas mais precisas, com maior realismo biológico, além de tornar a equação aplicável em diferentes locais. Contudo, existem grandes dificuldades em aplicar as variáveis categóricas nos modelos tradicionais, pois essas apresentam comportamento não linear, e a aplicação de variáveis binárias na regressão pode aumentar a complexidade do modelo.

Uma alternativa para representar as interações não lineares em relação aos métodos estatísticos tradicionais tem sido a aplicação de redes neurais artificiais (RNA). No manejo florestal, a aplicação dessa ferramenta computacional leva ao aperfeiçoamento das técnicas de biometria e inventário florestal, permitindo assim maior suporte à tomada de decisão (PENG & WEN, 1999).

A RNA tem uma ou mais camadas de entrada e saída e pode ter nenhuma, uma ou mais camadas intermediárias. A camada de entrada recebe as informações do problema e as transmite para as demais camadas, a camada intermediária tem a função de extrair as características dos dados e a camada de saída irá gerar a solução do problema. As conexões entre os neurônios, representados por pesos, têm a função de ponderar as informações recebidas, bem como armazenar o conhecimento adquirido pela rede no final do treinamento (BINOTI & BINOTI, 2014).

Na Engenharia Florestal a superioridade das redes neurais, quando comparadas com modelos tradicionais, pode ser verificada em BINOTI (2014) no estudo para estimar volume individual de *Eucalyptus* sp.; BINOTI et al. (2013), trabalhando com a mesmo gênero para estimar a altura em nível de povoamento, CASTRO et al. (2013) para a modelagem do crescimento e produção, também para

Eucalyptus sp., SCHIKOWSKI et al. (2015), estudando a forma do fuste em *Eucalyptus* sp., CASTRO et al. (2015) na modelagem do crescimento e produção para um povoamento de *Eucalyptus* utilizando dois métodos para quantificação do índice de local. Assim, pode-se notar que o gênero *Eucalyptus* tem sido amplamente estudado utilizando redes neurais.

Segundo o anuário IBA (2014), o gênero *Pinus* é o segundo mais plantado no país, destacando-se pela alta produtividade para a produção de madeira serrada, e pelas características de suas fibras para a produção de celulose e papel. Uma vez que no Brasil ainda não foram realizados estudos de RNA com o gênero *Pinus* sp., o objetivo deste trabalho foi testar as RNA para estimação da altura em *Pinus taeda* L., avaliar sua aplicabilidade e compará-la com os métodos tradicionais de regressão.

MATERIAL E MÉTODOS

Descrição dos plantios

Os dados do estudo são provenientes de inventários contínuos de plantios de *Pinus taeda* L. não desbastados, localizados no município de Telêmaco Borba, região centro oriental do estado do Paraná. A área do estudo abrange cerca de 3.000 pertencentes à empresa Klabin.

Foram mensuradas 304 parcelas, totalizando 2.454 remedições. As parcelas utilizadas apresentam forma retangular com área variável, sendo a área mínima de 446,6 m² e máxima de 979,4 m². A idade dos plantios mensurados no inventário contínuo variou entre 3,3 anos e 20 anos, com maior frequência nas idades 4 a 8 anos (Figura 1).

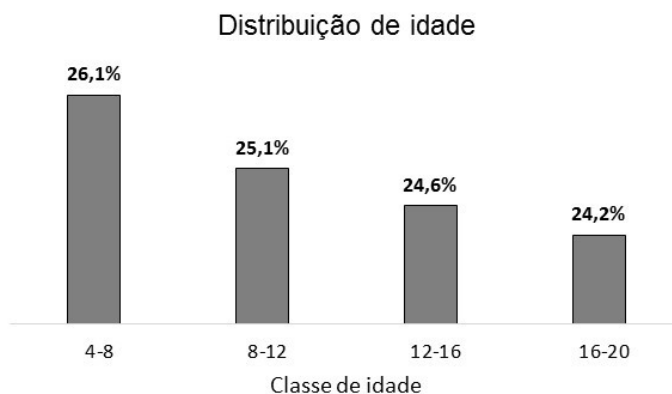


FIGURA 1. Distribuição de idade dos dados em análise

Ajuste dos modelos de relação hipsométricas

Para o ajuste das relações hipsométricas ($h = f(dap)$) foram selecionadas as árvores com diâmetros a 1,3 m (*dap*) e alturas mensuradas, excluindo aquelas com código de qualidade morta e quebrada. As parcelas foram estratificadas conforme a espécie e ano de plantio, gerando 11 estratos.

Foram utilizados 70% dos dados para o ajuste dos modelos, e os outros 30% utilizados para a validação das equações. A separação dos dados foi realizada de forma aleatória.

Foram ajustados 4 modelos hipsométricos, 2 relacionados a variável dependente altura com a variável independente diâmetro à 1,30m (*dap*) e os outros

dois com as variáveis independentes *dap* e idade. Os modelos utilizados estão apresentados na Tabela 1.

TABELA 1. Modelos hipsométricos para *Pinus taeda*

Modelo	Modelo
1	$Ln h = b_0 + b_1 \cdot \frac{1}{dap}$
2	$Ln h = b_0 + b_1 \cdot Ln dap$
3	$Ln h = b_0 + b_1 \cdot \frac{1}{dap} + b_2 \cdot \frac{1}{I}$
4	$Ln h = b_0 + b_1 \cdot Ln dap + b_2 \cdot \frac{1}{I}$

Em que: *h* = altura total (m); *dap* = diâmetro a 1,30 m (cm); *I* = idade (anos); *b_{is}* = parâmetros a serem estimados.

O ajuste dos modelos foi realizado por meio do software Statistica 7.0 (STATSOFT, Inc, 2009), sendo posteriormente selecionados dois modelos, os quais foram aplicados aos dados de validação. Para selecionar o modelo por estrato, foram analisadas as estatísticas: coeficiente de determinação ajustado e erro padrão da estimativa em porcentagem, sendo o mesmo recalculado utilizando Fator de Meyer de modo a permitir a correção da discrepância logarítmica uma vez que as variáveis dependentes sofreram transformação logarítmica (Tabela 2).

TABELA 2. Estatísticas utilizadas para a seleção de modelos de relação hipsométrica para *Pinus taeda*

Estatísticas	Fórmulas
Coeficiente de determinação ajustado	$R^2_{ajust} = 1 - \left\{ (1 - R^2) \cdot \left(\frac{n-p}{n-p-1} \right) \right\}$ $R^2 = \frac{SQReg}{SQT}$
Erro padrão da estimativa em porcentagem	$S_{yx}\% = \frac{S_{yx}}{\bar{y}} \cdot 100$ $S_{yx} = \sqrt{\frac{\sum_1^n (y_i - \hat{y}_i)^2}{(n-p)}}$
Fator de Meyer (FM)	$FM = e^{0,5(S_{yx}^2)}$
Análise gráfica de resíduos	$Resíduo \% = \left(\frac{y_i - \hat{y}_i}{y_i} \right) \cdot 100$

Em que: R^2 = coeficiente de determinação ajustado; $SQReg$ = soma dos quadrados da regressão; SQT = soma dos quadrados total; n = número de observações; p = número de coeficientes incluindo b_0 ; R = coeficiente de determinação; S_{yx} = erro padrão da estimativa; y_i = valor real de cada observação; \hat{y}_i = valor estimado de cada observação; e = 2,718281828;

Além disso foram utilizadas, também, as estatísticas complementares Viés (V), Média das Diferenças Absolutas (MD) e Desvio Padrão das Diferenças (DPD), que possibilitam uma análise mais detalhada do desempenho das estimativas (Tabela 3).

TABELA 3. Estatísticas complementares para análise dos modelos hipsométricos *Pinus taeda* L.

Estatística	Estimador
Viés (V)	$V = \frac{\sum_{i=1}^n Y_i - \sum_{i=1}^n \hat{Y}_i}{n}$
Média das diferenças absolutas (MD)	$MD = \frac{\sum_{i=1}^n Y_i - \hat{Y}_i }{n}$
Desvio padrão das diferenças (DPD)	$DPD = \sqrt{\left(\frac{\sum_{i=1}^n d_i^2}{n} - \left(\frac{\sum_{i=1}^n d_i}{n} \right)^2 \right) / (n-1)}$

Em que: Y_i = valor observado e \hat{Y}_i = valor estimado; n = número de observações; e $d_i = (Y_i - \hat{Y}_i)$.
 Fonte: Figueiredo Filho *et al.* (1996)

O viés indica a existência ou não de tendências entre os resíduos, porém não mede o quão afastado do eixo zero estão os valores residuais. Já a média das diferenças fornece justamente a ideia da amplitude dos erros, e o desvio padrão das diferenças mostra a homogeneidade entre os resíduos (SOUZA, 2013).

Com essas três estatísticas complementares foram estabelecidas notas para os modelos e foram classificados como melhores aqueles que resultaram na menor soma de notas.

Treinamento das redes neurais

Os mesmos dados utilizados para o ajuste de modelos por meio de regressão foram selecionados para o treinamento das redes; no entanto, neste caso os dados não foram estratificados. A estratificação foi realizada posteriormente ao treinamento das redes com o intuito de gerar as estatísticas de comparação com os modelos de regressão. A mesma proporção de dados utilizada para o ajuste das relações hipsométricas foi utilizada para o treinamento das RNA, sendo 70%, e os outros 30% restantes foram utilizados para validação das redes selecionadas.

As redes foram treinadas para estimar a variável dependente altura total, utilizando as diferentes combinações entre as variáveis independentes quantitativas *dap* e idade. As redes treinadas foram do tipo *perceptrons*, conhecidas como MLP (*Multilayer Perceptron*) com uma camada de entrada para receber os dados (variáveis independentes), três ou quatro camadas intermediárias para processamento e uma camada de saída (variável dependente - altura estimada). Nestas redes, as funções de ativação testadas foram a linear e a logística. Também foram treinadas redes do tipo RBF (*Radial Basis Function*) onde os neurônios da camada de saída foram sempre lineares (função identidade) e os neurônios da camada intermediária, funções de base radial (gaussianas) como função de ativação e redes lineares, que possuem funções de ativação sempre lineares (HAYKIN, 2001).

Foram limitadas para ambas as redes um número máximo de 60 neurônios por camada, uma vez que o número de neurônios deve ser limitado para evitar o “*overtraining*”.

Na construção e treinamento, utilizou-se o software Statistica 7.0 (STATSOFT, Inc, 2009). As 60 melhores redes foram retidas. Dessas, foram

selecionadas as 20 redes com maior correlação entre os dados estimados e observados. As etapas de treinamento (pré-processamento, treinamento propriamente dito e o pós-processamento) foram realizadas pela ferramenta Automated Neural Network. A própria ferramenta selecionou, em cada arquitetura de RNA, as variáveis de entrada de importância para estimativa das variáveis de saída.

As redes selecionadas foram analisadas pelas mesmas estatísticas dos modelos hipsométricos, com exceção do coeficiente de determinação, fazendo-se uma nova seleção considerando apenas as quatro melhores redes. Essas redes foram aplicadas à base de dados de validação, sendo finalmente selecionadas as duas melhores redes e aplicadas aos dados de validação por estrato.

TABELA 4. Estratificação para a relação hipsométrica de *Pinus taeda*

Estrato	Ano de plantio	N° de observações
1	1993	1.661
2	1995	500
3	1997	1.027
4	1998	400
5	2001	2.313
6	2002	2.012
7	2003	27.565
8	2004	72.681
9	2005	98.110
10	2006	39.561
11	2007	3.300
Total		249.130

Avaliação da performance dos métodos

Para avaliar a performance da regressão e redes neurais artificiais, a média das estimativas das alturas geradas por cada método foi comparada entre si e entre a média das alturas reais por meio de Análise de Variância (ANOVA), seguida pelo teste de Tukey a 95% de probabilidade de confiança. Os pressupostos da ANOVA (normalidade e independência dos resíduos e homogeneidade das variâncias) foram previamente verificados e atendidos.

RESULTADOS E DISCUSSÃO

Ajuste dos Modelos de Regressão

Nas análises estatísticas os modelos de regressão 1 e 2, que relacionam a variável dependente altura com o diâmetro apresentaram, em geral, estatísticas inferiores quando comparados aos modelos que relacionam a altura com as variáveis independentes diâmetro e idade. Para os modelos 1 e 2 foram gerados coeficientes de determinação ajustado (R^2 ajust) de no máximo 70% e erro padrão da estimativa (S_{yx}) variando entre 8% e 20%. Como apresentaram estatísticas pouco satisfatórias tiveram como consequência as piores notas em viés (V), média das diferenças absolutas (MD) e desvio padrão das diferenças (DPD) (Tabela 5).

BARTOSZECK et al. (2002), em um trabalho com bracatingais no Estado do Paraná também destacaram a inferioridade do modelo de Curtis (modelo 1) que utiliza apenas a variável diâmetro quando ajustado em nível de povoamento. No trabalho de GUIMARÃES et al. (2009) em plantios de *Eucalyptus* sp. foi observado que a utilização do material genético e idade nos modelos hipsométricos gerou estimativas melhores do que nos modelos que utilizaram apenas a variável *dap*.

Os modelos 3 e 4, apresentaram coeficientes de determinação de no mínimo 72%, chegando a 93% no estrato 4. Já o erro padrão da estimativa variou entre 4%

e 12%. Nas análises estatísticas complementares apresentaram as melhores notas, sendo selecionados para a etapa de validação.

TABELA 5. Análises estatísticas dos modelos hipsométricos para *Pinus taeda*

Estrato	Modelo	R ² ajust	Syx%	V	MD	DPD	Notas	Classificação
1	1	58%	9%	(1)-0,01	(3) 1,71	(3) 2,17	7	3
	2	54%	9%	(3)-0,01	(4) 1,80	(4) 2,27	11	4
	3	76%	7%	(2)-0,01	(1) 1,26	(1) 1,63	4	1
	4	72%	7%	(4)-0,01	(2) 1,40	(2) 1,77	8	2
2	1	35%	9%	(4) 0,10	(4) 1,96	(4) 2,39	12	4
	2	35%	7%	(3) 0,03	(3) 1,93	(3) 2,38	9	3
	3	86%	4%	(1)-0,00	(1) 0,78	(1) 1,08	3	1
	4	87%	4%	(2)-0,00	(2) 1,12	(2) 1,12	6	2
3	1	67%	7%	(4)-0,03	(4) 1,54	(3) 1,54	11	4
	2	64%	6%	(3)-0,04	(3) 1,26	(4) 1,62	10	3
	3	87%	5%	(1)-0,00	(1) 0,93	(1) 1,18	3	1
	4	77%	6%	(2)-0,01	(2) 1,04	(2) 1,30	6	2
4	1	45%	9%	(3)-0,01	(4) 1,50	(4) 1,78	11	4
	2	45%	8%	(4) 0,02	(3) 1,49	(3) 1,77	10	3
	3	93%	3%	(1)-0,00	(1) 0,51	(1) 0,66	3	1
	4	91%	4%	(2)-0,00	(2) 0,57	(2) 0,70	6	2
5	1	60%	12%	(2) 0,00	(4) 1,31	(4) 1,68	10	4
	2	66%	12%	(3)-0,01	(3) 1,27	(3) 1,62	9	3
	3	86%	7%	(1)-0,00	(1) 0,80	(1) 1,01	3	1
	4	84%	8%	(4)-0,01	(2) 0,87	(2) 1,08	8	2
6	1	33%	14%	(4) 0,09	(4) 1,59	(4) 2,12	12	4
	2	40%	14%	(3) 0,06	(3) 1,23	(3) 1,84	9	3
	3	70%	10%	(2) 0,01	(2) 1,03	(2) 1,42	6	2
	4	71%	10%	(1) 0,00	(1) 0,99	(1) 1,40	3	1
7	1	51%	17%	(3) 0,04	(4) 1,58	(4) 1,99	11	4
	2	49%	17%	(4) 0,16	(3) 1,48	(3) 1,90	10	3
	3	78%	12%	(1) 0,00	(1) 1,04	(1) 1,33	3	1
	4	77%	12%	(2)-0,01	(2) 1,04	(2) 1,34	6	2
8	1	50%	20%	(4) 0,26	(4) 1,71	(4) 2,09	12	4
	2	67%	17%	(1) 0,00	(3) 1,42	(3) 1,79	7	3
	3	81%	12%	(2) 0,01	(2) 1,01	(2) 1,29	6	2
	4	81%	12%	(3)-0,01	(1) 0,99	(1) 1,28	5	1
9	1	58%	20%	(4) 0,04	(4) 1,56	(4) 1,97	12	3
	2	70%	18%	(1) 0,01	(3) 1,33	(3) 1,71	7	2
	3	82%	13%	(3) 0,01	(2) 1,02	(2) 1,30	7	2
	4	83%	13%	(2)-0,01	(1) 1,00	(1) 1,27	4	1
10	1	42%	17%	(4) 0,02	(4) 1,41	(4) 1,75	12	4
	2	43%	16%	(3)-0,02	(3) 1,27	(3) 1,60	9	3
	3	70%	12%	(2) 0,01	(2) 1,00	(2) 1,26	6	2
	4	72%	12%	(1)-0,01	(1) 0,96	(1) 1,23	3	1
11	1	47%	20%	(4) 0,28	(4) 2,23	(4) 2,86	12	4
	2	60%	18%	(1) 0,06	(1) 1,19	(1) 2,10	3	1
	3	73%	11%	(3) 0,13	(2) 1,41	(2) 2,35	7	2
	4	75%	11%	(2) 0,10	(3) 1,49	(3) 2,46	8	3

Nota 1: As classificações para cada estatística complementar estão apresentadas entre parênteses.

Nota 2: Os modelos selecionados estão apresentados em negrito

Na análise gráfica de resíduos observa-se uma tendência em superestimar a altura das árvores menores e subestimar com o aumento desta variável. Essas características são ainda mais visíveis nos modelos que apresentam como variável independente apenas o dap.

As subestimativas e superestimativas também são bem visíveis nos estratos que possuem alta variabilidade em altura como os estratos 7, 8 e 9 (Figura 2). Isso pode ser explicado pelo fato das árvores mensuradas possuírem alturas diferentes para um mesmo diâmetro, fazendo com que o modelo estime para um determinado diâmetro a altura mais frequente nas medições.

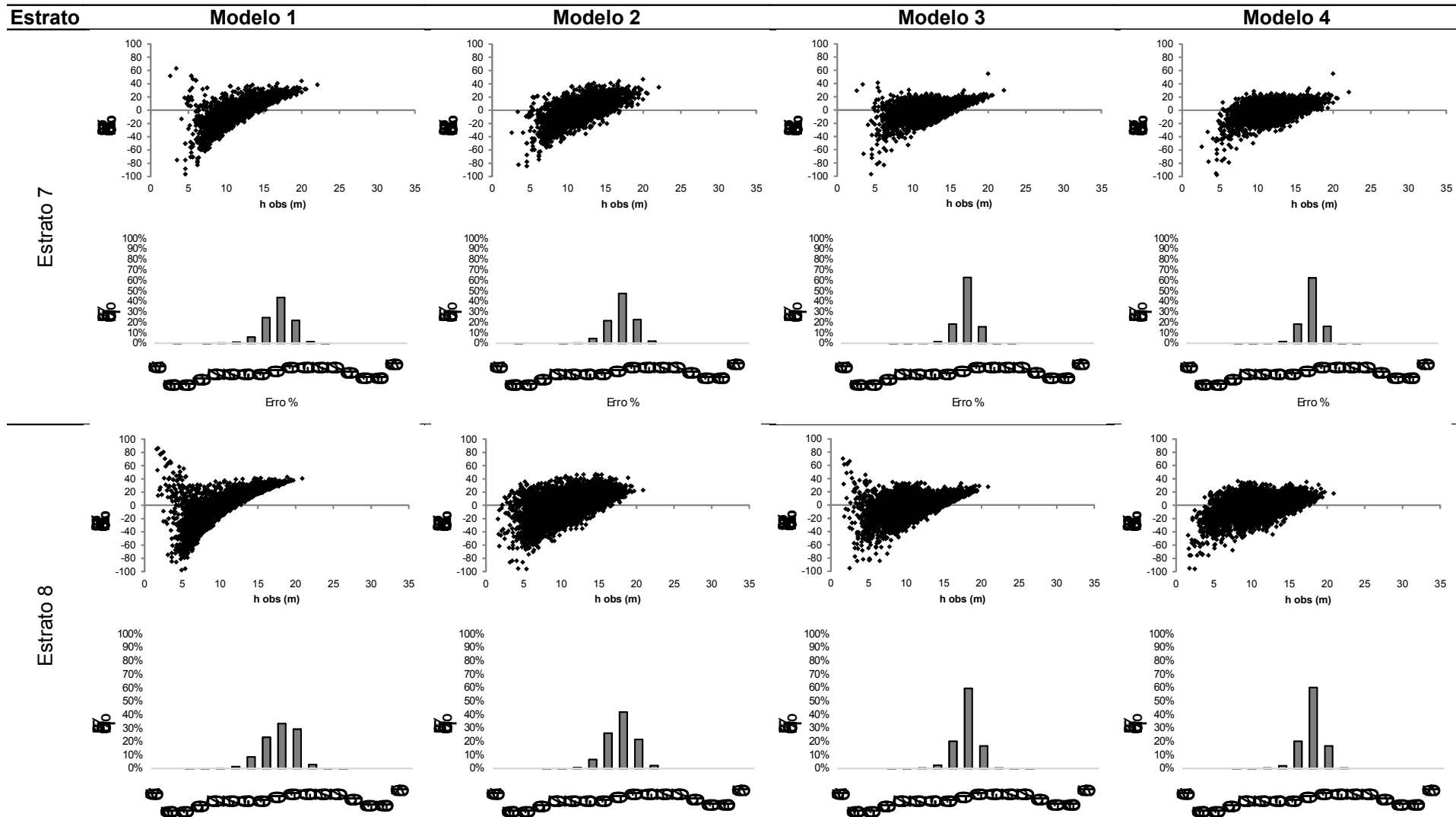


FIGURA 2. Gráficos de dispersão de resíduos do ajuste de modelos de regressão para os estratos 7, 8 e 9. Continua...

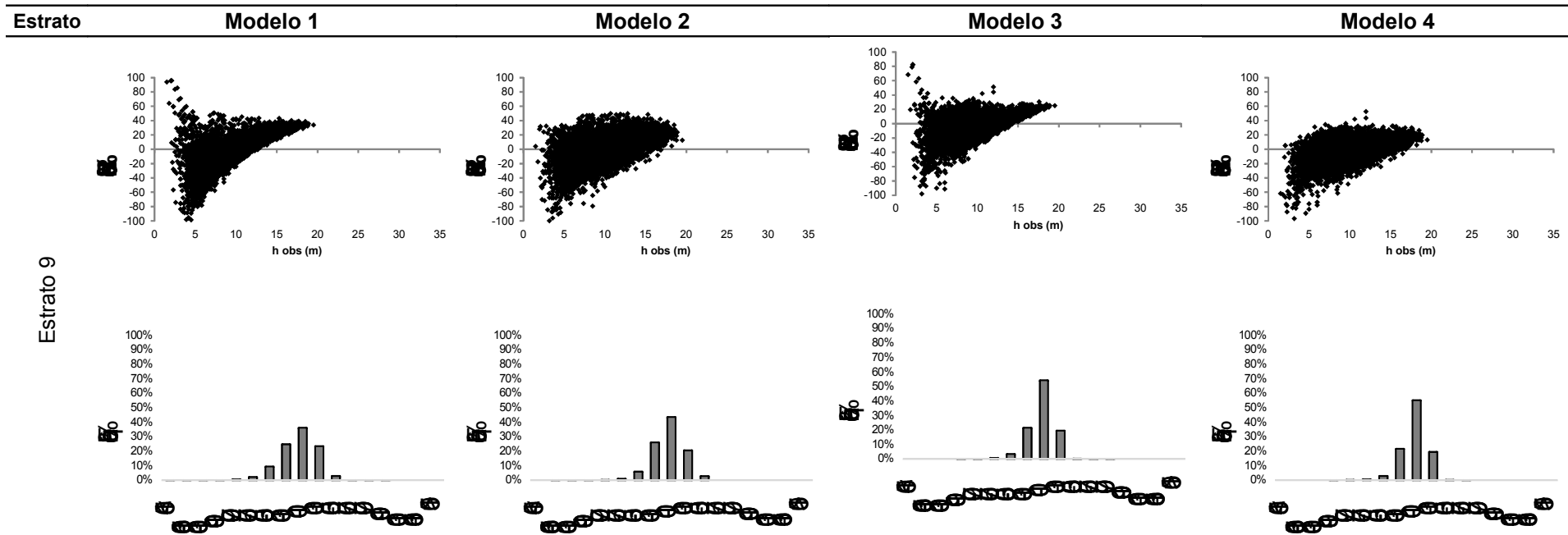


FIGURA 2. Gráficos de dispersão de resíduos do ajuste de modelos de regressão para os estratos 7, 8 e 9. Continuação

Treinamento das Redes Neurais

Foram selecionadas 60 redes neurais, 9 lineares, 25 RBF, 21 MLP com 3 camadas intermediárias e 5 MLP com 4 camadas intermediárias. Essas redes foram nomeadas, Rede 1, Rede 2, ..., Rede 60. Dessas, 20 redes com maior coeficiente de correlação (R) foram selecionadas, sendo calculadas as estatísticas principais e complementares (Tabela 6). As redes 21, 23, 55, 57 além de apresentar os maiores coeficientes de correlação e menores erros, tiveram as melhores notas nas estatísticas complementares, sendo selecionadas para a etapa de validação com a base de dados total.

Essas redes são do tipo MLP, onde apenas a de número 23 apresenta quatro camadas intermediárias, enquanto as demais possuem 3 camadas intermediárias. Além disso, as quatro redes possuem 2 variáveis na camada de entrada e diferentes números de neurônios nas camadas intermediárias.

TABELA 6. Análises estatísticas das redes neurais selecionadas para estimar a altura de *Pinus taeda* L.

Rede	Tipo	Arquitetura	R	Syx%	V	MD	DPD	Total	Classificação
3	RBF	5:60-60-1:1	0,91	16%	(10) 0,14	(9) 1,08	(10) 1,39	29	7
5	RBF	5:60-60-1:1	0,91	19%	(20)-0,74	(19) 1,25	(18) 1,46	57	17
10	RBF	4:53-60-1:1	0,91	17%	(1)-0,01	(17) 1,15	(19) 1,49	37	10
20	MLP	2:2-1-1:1	0,91	18%	(19)-0,50	(20) 1,25	(20) 1,50	59	18
21	MLP	2:2-2-1:1	0,92	15%	(5) 0,08	(1) 1,01	(1) 1,31	7	1
22	MLP	2:2-20-1:1	0,92	15%	(8)-0,12	(3) 1,03	(3) 1,31	14	4
23	MLP	2:2-27-27-1:1	0,92	15%	(2)-0,03	(5) 1,04	(6) 1,35	13	3
24	MLP	3:45-3-1:1	0,92	15%	(7)-0,11	(6) 1,05	(5) 1,34	18	5
25	MLP	3:45-4-1:1	0,91	16%	(13)-0,22	(11) 1,11	(14) 1,41	38	11
26	MLP	3:45-5-1:1	0,91	16%	(14)-0,24	(12) 1,11	(13) 1,41	39	12
29	MLP	4:53-10-1:1	0,91	17%	(18)-0,50	(18) 1,16	(12) 1,41	48	15
42	RBF	4:18-27-1:1	0,91	16%	(3)-0,03	(13) 1,12	(17) 1,43	33	9
47	Linear	2:2-1:1	0,92	16%	(9)-0,12	(8) 1,06	(8) 1,36	25	6
51	MLP	2:2-6-1:1	0,92	16%	(12)-0,20	(10) 1,09	(9) 1,38	31	8
52	MLP	2:2-7-1:1	0,91	16%	(15)-0,25	(14) 1,12	(15) 1,42	44	14
53	MLP	2:2-8-1:1	0,91	17%	(17)-0,32	(16) 1,15	(16) 1,42	49	16
54	MLP	2:2-9-1:1	0,92	16%	(11)-0,16	(7) 1,07	(7) 1,35	25	6
55	MLP	2:2-10-1:1	0,92	15%	(4)-0,04	(2) 1,02	(2) 1,31	8	2
56	MLP	2:2-28-1:1	0,92	16%	(16)-0,30	(15) 1,13	(11) 1,40	42	13
57	MLP	2:2-41-1:1	0,92	15%	(6) 0,11	(4) 1,04	(4) 1,32	14	4

Nota 1: As classificações para cada estatística complementar estão apresentadas entre parênteses.

Nota 2: As redes selecionadas estão apresentadas em negrito

Validação das Redes Neurais

As redes 21, 23, 55, 57 quando aplicadas na base de dados total de validação apresentaram comportamentos e resultados estatísticos muito similares. Neste caso a utilização de estatísticas complementares (Tabela 7) e a análise gráfica de resíduos (Figura 4) foram fundamentais para a escolha das redes a serem aplicadas nos dados de validação estratificados.

Por meio das notas atribuídas para as estatísticas complementares viés (V), média das diferenças absolutas (MD) e desvio padrão das diferenças (DPD) a rede 21, apresentou a melhor classificação (Tabela 7).

TABELA 7. Estatísticas das redes neurais selecionadas aplicadas aos dados de validação total para estimar a altura de *Pinus taeda*

Rede	R	Syx%	V	MD	DPD	Total	Classificação
21	0,93	12%	(3) 0,07	(1) 1,011	(1) 1,299	5	1
22	0,92	13%	(1)-0,03	(4) 1,043	(4) 1,346	9	3
23	0,93	12%	(2)-0,05	(2) 1,014	(2) 1,302	6	2
24	0,92	13%	(4) 0,11	(3) 1,037	(3) 1,314	10	4

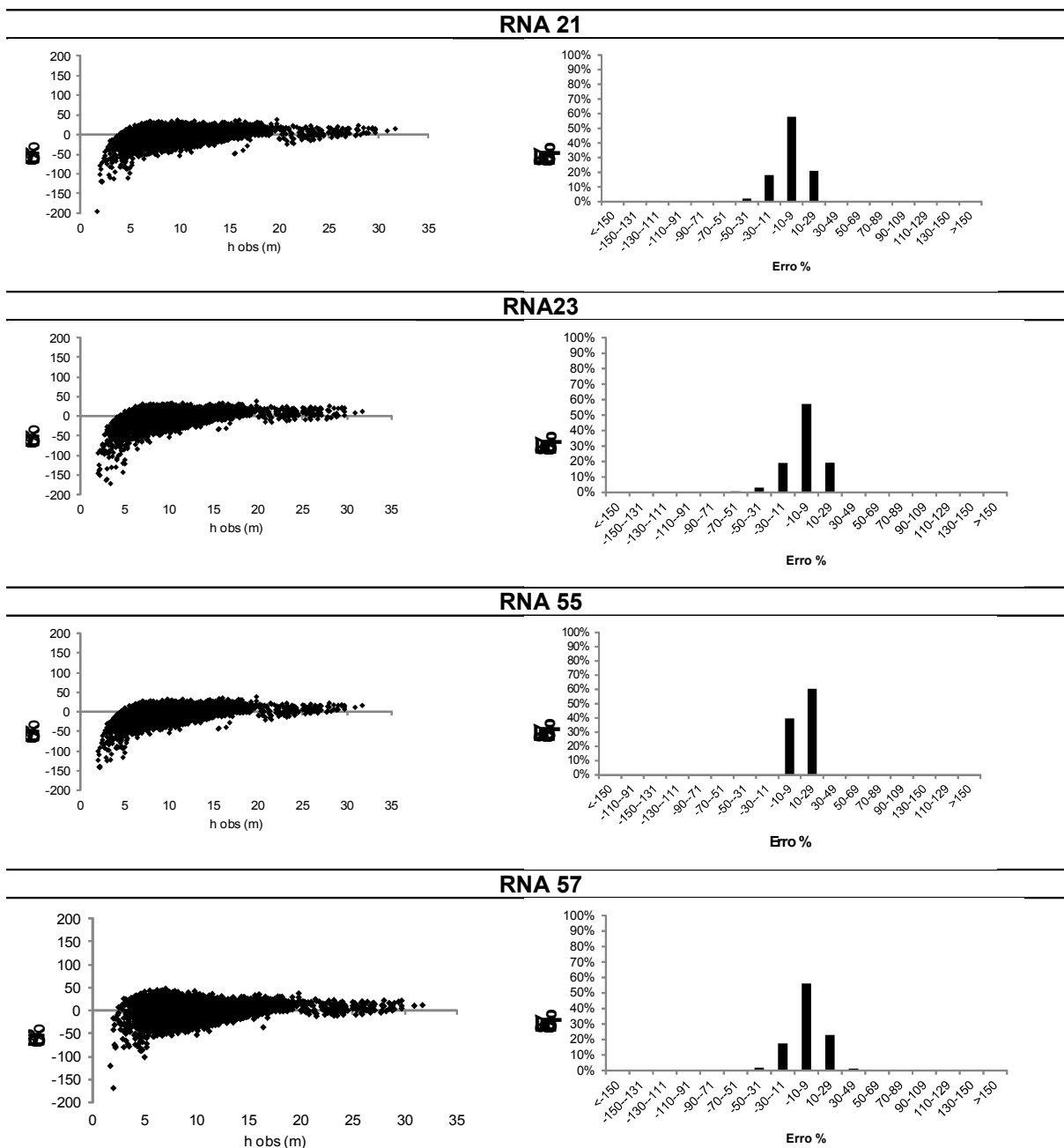


FIGURA 4. Gráficos de dispersão de resíduos e gráficos de frequência percentual dos erros percentuais para as estimativas de altura com redes neurais

Comparações entre o Método de Regressão x Redes Neurais Artificiais nos Dados Estratificados

Ao aplicar o modelo de regressão com as melhores estatísticas e a RNA com melhor desempenho no banco de dados de validação por estrato, ambos apresentaram comportamentos similares, com algumas exceções. Foi observado um comportamento superior dos modelos de regressão nos estratos 2 e 3, o que pode ser explicado pelo reduzido número de observações nesses estratos, pois como o treinamento das redes foi realizado com o banco de dados total, esses dois estratos não foram tão representativos como os demais, proporcionando desempenho inferior nesse método em relação à regressão. Já no método de regressão os modelos foram ajustados com os dados dos estratos separadamente, fato que pode explicar as melhores estimativas (Tabela 8).

TABELA 8. Estatísticas para a comparação entre os modelos de regressão e redes neurais para estimar altura de *Pinus taeda*

Estrato	Redes	R	Syx%	V	MD	DPD	Total	Classificação
1	rede 21	0,73	9,7%	(2)- 0,87	(2) 1,70	(2) 2,22	6	2
	modelo 4	0,84	7,2%	(1)-0,25	(1) 1,29	(1) 1,72	3	1
2	rede 21	0,81	11,0%	(2) 2,30	(2) 2,47	(2) 1,88	6	2
	modelo 3	0,82	6,3%	(1)- 0,29	(1) 1,20	(1) 1,64	3	1
3	rede 21	0,79	14,1%	(2) 2,80	(2) 2,91	(2) 1,61	6	2
	modelo 3	0,88	5,9%	(1)-0,24	(1) 1,07	(1) 1,31	3	1
4	rede 21	0,92	4,4%	(1) 0,29	(1) 0,66	(2) 0,87	4	1
	modelo 4	0,95	4,6%	(2)-0,46	(2) 0,67	(1) 0,70	5	2
5	rede 21	0,90	9,3%	(2) 0,29	(2) 1,00	(2) 1,25	6	2
	modelo 3	0,91	8,5%	(1)- 0,11	(1) 0,89	(1) 1,16	3	1
6	rede 21	0,87	10,6%	(2) 1,08	(2) 1,22	(1) 1,21	5	2
	modelo 4	0,86	8,3%	(1) 0,02	(1) 0,99	(2) 1,26	4	1
7	rede 21	0,90	11,1%	(2)- 0,23	(2) 1,05	(1) 1,29	5	1
	modelo 3	0,89	11,1%	(1) 0,07	(1) 1,04	(2) 1,30	4	2
8	rede 21	0,91	12,3%	(2)-0,31	(2) 1,02	(1) 1,25	5	2
	modelo 4	0,91	12,3%	(1) 0,01	(1) 1,00	(2) 1,28	4	1
9	rede 21	0,92	12,8%	(2) 0,28	(1) 0,96	(1) 1,20	4	1
	modelo 4	0,91	13,2%	(1)-0,04	(2) 1,00	(2) 1,28	5	2
10	rede 21	0,85	12,4%	(2) 0,26	(2) 0,98	(1) 1,23	5	2
	modelo 4	0,85	12,2%	(1)-0,01	(1) 0,96	(2) 1,23	4	1
11	rede 21	0,88	14,7%	(2) 0,41	(1) 1,08	(1) 1,39	4	1
	modelo 3	0,88	13,5%	(1) 0,27	(2) 2,45	(2) 3,08	5	2

Segundo BARROS (2002), a aplicação de equações genéricas que permitam estimativas para um conjunto de dados originados de povoamentos florestais com características silviculturais diferentes deve ser testada, pois permite diminuir o custo de processamento eliminando o trabalho de ajuste e a seleção de equações mais adequadas para cada unidade amostral.

Apesar da RNA gerar estatísticas próximas às da regressão com a aplicação de uma única rede para todos os estratos, o que pode parecer prático e desejável, o treinamento da RNA levou cerca de 17 horas para ser completado, enquanto o ajuste dos modelos regressão para todos os estratos foi realizado em menos de 8 horas de forma prática e rápida pelo software Statistica 7.0

Ao contrário de algumas bibliografias que alegam a simplicidade, rapidez e praticidade para o treinamento das RNA, nesse estudo isso não foi evidenciado. De forma prática, na dinâmica de empresas de grande porte, as informações de inventários florestais contínuos chegam diariamente e os dados devem ser importados e processados rapidamente para gerar informações que subsidiam os processos de colheita, silvicultura e/ou investidores.

De posse de uma base de dados rica, como a do presente estudo, é possível gerar modelos de regressão estratificados que podem ser utilizados e reajustados anualmente com a incorporação de novos estratos, materiais genéticos, sítios etc. Enquanto a RNA aplicada hoje nas empresas, deve ser treinada para cada nova massa de dados, fora do banco de dados e depois das alturas estimadas, as informações são importadas para o banco de dados.

A longo prazo, a aplicação de RNA fora do banco de dados não é sustentável, podendo ter o mesmo risco de erro humano no cadastro de parâmetros nos modelos de regressão, além da falta de controle das RNA utilizadas. Por isso, para a estimativa de altura utilizando os mesmos dados de entrada para a RNA e modelos de regressão, recomenda-se a utilização dos métodos tradicionais de regressão de forma estratificada.

Avaliação da performance dos métodos

A média das alturas estimadas pela RNA diferiu da média das alturas observadas nos estratos 2, 3, 6, 7, 8 e 10. Por outro lado, as médias das alturas estimadas pela regressão foi igual às médias das alturas observadas em todos os estratos (Tabela 9).

TABELA 9. Comparação de médias entre os três tratamentos com seus respectivos valores de f e p para estimativa de altura de *Pinus taeda*

Métodos	Estrato	Estrato 2	Estrato 3	Estrato 4	Estrato 5	Estrato 6	Estrato 7	Estrato 8	Estrato 9	Estrato 10	Estrato 11
	1										
	Médias										
H obs	24,28a	26,44a	22,74a	19,16a	13,44 ^a	15,31a	11,77a	10,45a	9,70a	10,09a	9,82a
H Regressão	24,70a	26,92a	23,15a	19,46a	13,73 ^a	15,36a	11,73a	10,46a	9,73a	10,10a	9,74a
H RNA	25,15a	23,34b	19,94b	20,08a	13,91 ^a	14,23b	12,00b	10,77b	9,42a	9,83b	9,41a
F	2,100	14,960	29,090	0,82	0,87	7,73	3,75	13,08	15,8	9,84	1,34
P	0,125	0,000	0,000	0,4478	0,4209	0,0005	0,0237	0,000	0,000	0,0001	0,2641

CONCLUSÃO

Os modelos de regressão apresentaram melhores estimativas de altura que as RNA. Ou seja, quando são utilizadas as mesmas variáveis de entrada para a regressão e RNA, empregar modelos hipsométricos estratificados gera estimativas mais precisas.

REFERÊNCIAS

BARTOSZECK, A. C. P. S.; MACHADO, S. A.; FILHO, A. F.; OLIVEIRA, E. B. Dinâmica da relação hipsométrica em função da idade, do sítio e da densidade inicial de povoamentos de bracatinga da Região Metropolitana de Curitiba, PR. **Revista Árvore**, Viçosa, v. 28, n. 4, jul/ago. 2004. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622004000400006>. doi: 10.1590/S0100-67622004000400006

BARROS, D. A.; MACHADO, S. A.; ACÉRBI JUNIOR, F. W.; SCOLFORO, J. R. Comportamento de modelos hipsométricos tradicionais e genéricos para plantações de *Pinus oocarpa* em diferentes tratamentos. **Boletim de Pesquisa Florestal**, Colombo, n. 45, p. 3-28, jul./dez. 2002. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/CNPF-2009-09/33608/1/pag3-28.pdf>

BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**. Viçosa, v. 37, n. 4, jul/ago. 2013. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000400007. doi:10.1590/S0100-67622013000400007

BINOTI, D. H. B. B.; BINOTI, M. L. M. S. Redes Neurais Artificiais Aplicadas à Ciência Florestal. Instituto de Inovação em Engenharia e Meio Ambiente-AMBINOVA: **Material de Apoio**. Viçosa, 83 p, 2014.

BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G.; GARCIA, S. L. R.; FERREIRA, M. Z.; RODE, R.; SILVA, A. A. L. Redes neurais artificiais para estimação do volume de árvores. **Revista Árvore**, Viçosa v. 38, n. 2, mar/abr. 2014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622014000200008>. doi:10.1590/S0100-67622014000200008

CASTRO, R. V. O.; SOARES, C. P. B.; MARTINS, F. B.; LEITE, H. G. Crescimento e produção de plantios comerciais de eucalipto estimados por duas categorias de modelos **Pesquisa Agropecuária Brasileira**, Brasília, v. 48, n. 3, p.287-295, mar. 2013. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-204X2013000300007>. doi:10.1590/S0100-204X2013000300007

CASTRO, R. V. O.; CUNHA, A. B.; SILVA, L. V.; LEITE, H. G.; SILVA, A. A. L. Modelagem do crescimento e produção para um povoamento de *Eucalyptus* utilizando dois métodos para quantificação do índice de local. **Scientia Forestalis**, Piracicaba, v. 43, n. 105, p.83-90, mar. 2015. Disponível em: <<http://www.ipef.br/publicacoes/scientia/leitura.asp?Article=08&Number=105&p=s>>.

FIGUEIREDO FILHO, A.; BORDERS, B. E.; HITCH K. L. Taper equations for *Pinus taeda* plantations in Southern Brazil. **Forest Ecology and Management**. Philadelphia, v. 83, p. 39-46, 1996.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 900 p.

GUIMARÃES, M. A. M.; CALEGÁRIO, N.; CARVALHO, L. M. T.; TRUGILHO, P. F. Height-diameter models in forestry with inclusion of covariates. **Revista Cerne**, Lavras, v.15, n.3, p.313-321, 2009. Disponível em: <<http://www.redalyc.org/html/744/74413021008/>>.

IBÁ – INDÚSTRIA BRASILEIRA DE ÁRVORES. **Anuário Estatístico da Indústria Brasileira de Árvores** - Brasília, 2014. p. 99. Disponível em: <<http://iba.org/pt/>> Acesso em: 19 set. 2015.

PENG, C.; WEN, X. Recent applications of artificial neural networks in forest resource management: an overview. In: AMERICAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE WORKSHOP, 1999, Orlando. Disponível em: <<http://www.aaai.org/Library/Workshops/1999/ws99-07-003.php>>. Acesso em: 19 set. 2015.

MACHADO, S. A. NASCIMENTO, R. G. M.; AUGUSTYNCZIK, A. L. D.; SILVA, L. C. R.; FIGURA, M. A.; PEREIRA, E. M.; TEO, S. J. Comportamento da relação hipsométrica de *Araucaria angustifolia* no capão da engenharia florestal da UFPR. **Boletim de Pesquisa Florestal**, Colombo, n.56, p.5-16. 2008. Disponível em: <<https://pfb.cnpf.embrapa.br/pfb/index.php/pfb/article/view/58>>.

SANQUETTA, C. R.; CORTE, A. P. D.; RODRIGUES, A. L.; WATZLAWICK, L. F. **Inventários Florestais: Planejamento e Execução**. Curitiba: Multi-Graphic, 2014. 406 p.

SILVA, G. F.; CURTO, R. A.; SOARES, C. P. B.; PIASSI, L. C. Avaliação de métodos de medição de altura em florestas naturais. **Revista Árvore**, Viçosa, v. 36, n. 2, p. 341-348, 2012. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622012000200015>. doi:10.1590/S0100-67622012000200015

SCHIKOWSKI, A. B.; CORTE, A. P. D.; SANQUETTA, C. R. Estudo da forma do fuste utilizando redes neurais artificiais e funções de afilamento. **Boletim de Pesquisa Florestal**, Colombo, v. 35, n. 82, p. 119-127, abr./jun. 2015. Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/129445/1/pfb.pdf>>. doi: 10.4336/2015.pfb.35.82.867

SOUZA, C. A. M.; FINGER, C. A. G.; SCHNEIDER, P. R.; SILVA, G. F.; THOMAS, C. Eficiência de um modelo de afilamento ajustado sem e com estratificação por classe de quociente de forma para formação dos sortimentos de *Pinus taeda* L. **Ciência Florestal**, Santa Maria, v. 22, n. 1, p. 125-135, jan./mar. 2012. Disponível em: <<http://www.redalyc.org/articulo.oa?id=53423372012>>.

THIERSCH, C. R.; ANDRADE, M. G.; MOREIRA, M. F. B.; LOIBEL, S. Estimativa da relação hipsométrica em clones de *Eucalyptus* sp. com o modelo de curtis ajustado por métodos bayesianos empíricos. **Revista Árvore**, Viçosa, v. 37, n.1, p.01-08, **ENCICLOPÉDIA BIOSFERA**, Centro Científico Conhecer - Goiânia, v.13 n.24; p.261 2016

2013. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000100001>. doi: 10.1590/S0100-67622013000100001

VENDRUSCOLO, D. G. S.; DRESCHER, R.; SOUZA, H. S. SILVA, R. S. Estratificação hipsométrica em plantios de eucaliptos na região sudeste de Mato Grosso. **Agrarian Academy**, Centro Científico Conhecer - Goiânia, v.2, n.03; p. 52-61, 2015. Disponível em <http://www.conhecer.org.br/Agrarian%20Academy/2015a/estratificacao.pdf>.

ZHANG, S.; BURKHART, H. E.; AMATEIS, R. L. The influence of thinning on tree height and diameter relationships in loblolly pine plantations. **Southern Journal of Applied Forestry**, Ottawa-Canadá, v. 21, p. 199-205, 1997. Disponível em: <https://www.researchgate.net/publication/233663054_The_Influence_of_Thinning_on_Tree_Height_and_Diameter_Relationships_in_Loblolly_Pine_Plantations>.