

ESTIMAÇÃO DA ALTURA DE PLANTIOS FLORESTAIS DE EUCALIPTO POR REGRESSÃO E REDES NEURAS ARTIFICIAIS

ESTIMATION FROM HEIGHT OF FOREST PLANTATIONS OF EUCALYPTUS BY REGRESSION AND ARTIFICIAL NEURAL NETWORKS

Matheus Teixeira Martins¹, Gabriel Paes Marangon², Emanuel Arnoni Costa³,
Bruna Denardin da Silveira⁴, Rafael Cubas⁵, Jean Pierre Cavalli⁶

^{1, 2, 4} Universidade Federal do Pampa, São Gabriel, RS, Brasil – mteixeiramartins@gmail.com,
gabrielmarangon@unipampa.edu.br & brunadenardin@gmail.com

³ Universidade Federal de Uberlândia, Monte Carmelo, MG, Brasil – emanuelarnonicost@hotmail.com

⁵ Universidade do Oeste de Santa Catarina, Canoinhas, SC, Brasil – florestal.rafael@gmail.com

⁶ Universidade Federal de Santa Maria, Santa Maria, RS, Brasil – jeancavalli@gmail.com

RESUMO

A modelagem estatística na predição de alturas de árvores em florestas plantadas é uma forma de reduzir o tempo e custo do levantamento de dados do inventário florestal. Nesse sentido, o presente estudo teve como objetivo estimar a altura de árvores de *Eucalyptus grandis* W. Hill através de modelos de regressão (MR) e redes neurais artificiais (RNA). Para isso, foram utilizados 713 pares de dados de altura e diâmetro de árvores individuais medidas no inventário florestal, sendo que 70% dos dados foram utilizados para o ajuste dos modelos de regressão e treino das RNA e 30% utilizados para validação das técnicas. Foram ajustados cinco modelos hipsométricos tradicionais, cinco modelos em função da variável dap e da variável idade na forma aritmética, quadrática, logarítmica, inversa e raiz quadrada, totalizando vinte e cinco novos modelos e, por fim, foram treinadas cinco redes neurais do tipo *Multilayer Perceptron*. As técnicas foram avaliadas estatisticamente através da correlação (r_{YY}), raiz quadrada do quadrado médio do erro (RQME) e análise gráfica de resíduos. Tanto no treino como na validação as RNA obtiveram melhores resultados estatísticos, na validação a melhor RNA obteve r_{YY} de 0,941 e RQME de 1,238 m. O modelo de regressão de relação h/d obteve r_{YY} de 0,928 e RQME de 1,373 m. O modelo de regressão com inserção da variável idade apresentou r_{YY} de 0,936 e RQME de 1,289 m. Ambas as técnicas poderiam ser utilizadas para estimar a altura das árvores de *E. grandis*, porém as RNA são mais acuradas.

PALAVRAS-CHAVE: Aprendizado de máquinas, Inteligência artificial, Mensuração florestal.

ABSTRACT

Statistical modeling in predicting tree heights in planted forests is a way to reduce the time and cost of collecting forest inventory data. In this sense, the present study aimed to estimate the height of *Eucalyptus grandis* W. Hill trees through regression models (MR) and artificial neural networks (RNA). For this, 713 pairs of height and diameter data of individual trees measured in the forest inventory were used, being that 70% of the data were used to adjust the regression and training models of RNA and 30% used to validate the techniques. Five traditional hypsometric models were adjusted, five models as a function of the dap variable and the age variable in arithmetic, quadratic, logarithmic, inverse and square root form, totaling twenty-five new models and, finally, five Multilayer Perceptron neural networks were trained. The techniques were statistically evaluated by correlation (r_{YY}), square root mean square error (RQME) and graphical analysis of residuals. In both training and validation, RNA obtained better statistical results, in validation the best RNA obtained r_{YY} of 0.941 and RQME of 1.238 m. The h/d ratio regression model obtained r_{YY} of 0.928 and RQME of 1.373 m. The regression model with insertion of the age variable presented r_{YY} of 0.936 and RQME of 1.289 m. Both techniques could be used to estimate the height of *E. grandis* trees, but RNAs are more accurate.

KEYWORDS: Machine learning, Artificial intelligence, Forest measurement.

INTRODUÇÃO

A quantificação do estoque volumétrico é realizada por meio de inventários florestais contínuos ou temporários. Ambos consistem basicamente na medição de amostras representativas da população, denominadas de unidades amostrais, também conhecidas como parcelas. As estimativas de estoques de crescimento e de colheita são importantes elementos do manejo florestal, uma vez que fornecem informações quantitativas dos povoamentos, auxiliando na definição de planos de manejo (BINOTI et al., 2013).

Existem muitos modelos para estimação da altura total de árvores, muitos deles podendo ser utilizados em várias espécies. A mensuração da altura de árvores é de grande importância nos inventários florestais, apresentando um custo significativo, sendo assim, torna-se necessário a realização de estudos voltados à modelagem, aos procedimentos e equipamentos utilizados para mensuração (VENDRUSCOLO et al., 2017).

Além das técnicas de modelagem por regressão o aprendizado de máquinas está se destacando no meio florestal, devido a sua capacidade de gerar estimativas mais acuradas que as equações obtidas com os modelos. Na mensuração florestal, as Redes Neurais Artificiais (RNA) é uma das técnicas de aprendizado de máquinas que estão sendo estudadas e ganhando espaço nesse meio, tais como: Binoti et al. (2013) que utilizaram RNA para estimar a altura de povoamentos equiâneos de *Eucalyptus* sp.; Vendruscolo et al. (2017) estimaram altura de árvores de *Tectona grandis* L.f.; Mendonça et al. (2018) estimaram altura de árvores de ipê felpudo utilizando RNA; Martins et al. (2019a) estimaram altura de árvores de *Parapiptadenia rigida* (Benth.) Brenan com uso de RNA; entre outros.

Uma RNA consiste em múltiplas unidades de processamento simples, denominados neurônios artificiais, que estão conectados entre si e organizados em camadas, formando um sistema computacional paralelo para executar uma determinada tarefa (BULLINARIA, 2014). Estes fornecem estimações precisas, tornando o alcance dos resultados relacionados à dendrometria florestal mais célere, menos onerosa e dispendiosa economicamente, além de ser menos laboriosa (FERREIRA et al., 2014).

Objetivou-se com este trabalho comparar modelos de regressão e redes neurais artificiais para estimar a altura de árvores de *E. grandis* de povoamentos florestais presentes em duas mesorregiões do estado do Rio Grande do Sul.

MATERIAL E MÉTODOS

Caracterização e localização da área do estudo

Os dados foram coletados em povoamentos florestais de *E. grandis*, localizados no estado do Rio Grande do Sul, nos municípios de Santa Cruz do Sul, Lajeado, Estrela e Cachoeira do Sul, pertencentes a mesorregião Centro Oriental Rio-grandense, além dos municípios de Montenegro, Gramado, Canela, Porto Alegre, Osório e Camaquã, pertencentes a mesorregião Metropolitana.

Essas mesorregiões, as quais os povoamentos estão inseridos, possuem clima, segundo a classificação de Köppen, do tipo Cfa, com temperatura anual média de 17,9 °C e a precipitação média anual de 1826,0 mm (ALVARES et al., 2013). Os solos das mesorregiões avaliadas são argilosos, profundos, de coloração avermelhada em toda sua extensão do perfil, bem drenados e derivados de siltitos finos (LEMONS et al., 1973).

Amostragem dos dados

O conjunto de dados utilizado para realização do estudo foram obtidos através do inventário florestal utilizando parcelas permanentes de 30 x 20 m, obtendo a altura da árvore e o diâmetro a altura do peito (DAP), com uma abrangência de idade entre 4 e 27 anos.

Foram utilizados 713 pares de dados de altura e DAP de árvores individuais medidas no inventário florestal. Esse conjunto foi dividido em 70% (500 indivíduos) para realização do ajuste dos modelos e treino de RNA e 30% (213 indivíduos) para validação do melhor modelo sem e com inserção da variável idade e da melhor RNA.

Análise dos dados

Foram ajustados 5 modelos hipsométricos utilizando a planilha eletrônica Excel de tal modo que estes foram ajustados na forma original e com a inserção da variável idade (t - anos). Essa variável foi inserida de diferentes formas nos modelos, são elas: $[t; t^2; \ln(t); 1/t$ e \sqrt{t}], assim, foram ajustados 25 modelos, que obtiveram o diâmetro (d) e a idade (t) como variáveis independentes (Tabela 1).

Os modelos de regressão de natureza logarítmica $[\ln(h)]$ foram recalculados para a unidade original da variável dependente $[h]$ e corrigidos pelo fator de correção $[F_c = e(QME/2)]$, referente a discrepância logarítmica. Os modelos ajustados foram avaliados quanto a

colinearidades nos coeficientes de regressão, sendo considerada a presença de multicolinearidade quando obtidos valores do fator de inflação de variâncias (VIF) superiores a 10. (SCHNEIDER et al., 2009).

Tabela 1. Modelos testados para estimar altura de árvores em povoamentos de *E. grandis* em duas mesorregiões no estado do Rio Grande do Sul.

Nomenclatura	Modelos	Inclusão da idade (t)
MR (h/d)-1	$h = \beta_0 + \beta_1 \cdot d + \beta_2 \cdot d^2$	-
MR (h/d)-2	$\ln h = \beta_0 + \beta_1 \cdot \ln d$	-
MR (h/d)-3	$h = \beta_0 + \beta_1 \cdot d + \beta_2 \cdot (1/d)$	-
MR (h/d)-4	$h = \beta_0 + \beta_1 \cdot (1/d) + \beta_2 \cdot d^2$	-
MR (h/d)-5	$\ln h = \beta_0 + \beta_1 \cdot \sqrt{d}$	-
MR-6-10	$h = \beta_0 + \beta_1 \cdot d + \beta_2 \cdot d^2 + \beta_3 \cdot x$	
MR-11-15	$\ln h = \beta_0 + \beta_1 \cdot \ln d + \beta_2 \cdot x$	
MR-16-20	$h = \beta_0 + \beta_1 \cdot d + \beta_2 \cdot (1/d) + \beta_3 \cdot x$	t; t ² ; ln(t); 1/t; \sqrt{t}
MR-21-25	$h = \beta_0 + \beta_1 \cdot (1/d) + \beta_2 \cdot d^2 + \beta_3 \cdot x$	
MR-26-30	$\ln h = \beta_0 + \beta_1 \cdot \sqrt{d} + \beta_2 \cdot x$	

h = altura (m); d = diâmetro à altura do peito (cm); t = idade (anos); β_i = coeficientes de regressão estimados; x = transformações para a variável idade (t): t; t²; ln(t); 1/t e \sqrt{t} .

Os modelos de regressão de natureza logarítmica [ln(h)] foram recalculados para a unidade original da variável dependente [h] e corrigidos pelo fator de correção [Fc = e(QME/2)], referente a discrepância logarítmica. Os modelos ajustados foram avaliados quanto a colinearidades nos coeficientes de regressão, sendo considerada a presença de multicolinearidade quando obtidos valores do fator de inflação de variâncias (VIF) superiores a 10. (SCHNEIDER et al., 2009).

A colinearidade ou multicolinearidade ocorre quando há correlação significativa entre variáveis preditoras ou variáveis independentes de um modelo proposto. Como regra geral, o problema ocorre quando as variáveis independentes apresentam maior grau de correlação significativa entre si do que com a variável dependente (SCHNEIDER et al., 2009).

Com a ocorrência de multicolinearidade as variâncias associadas aos estimadores dos coeficientes da regressão podem atingir valores demasiadamente elevados, tornando-os pouco confiáveis, e ainda, os parâmetros gerados podem apresentar valores absurdos ou sem nenhuma coerência com o fenômeno biológico estudado (CARVALHO, 1994).

Uma elevada multicolinearidade pode resultar em

estimativas inconsistentes dos coeficientes de regressão e na superestimativa dos efeitos diretos das variáveis explicativas sobre a variável resposta e, por consequência, levar à interpretação errada dos resultados (CRUZ & CARNEIRO, 2003). Valores de VIF maiores que 10 indicam graves problemas de multicolinearidade (FREUND & WILSON, 1998).

Após o ajuste dos modelos de regressão, foi selecionado o melhor modelo sem a inserção da variável idade e o melhor modelo com inserção da variável idade, e estes foram validados, para assim se ter um resultado mais realista das estimativas.

Foram treinadas 100 RNA do tipo *Perceptron* de múltiplas camadas, comumente conhecidas como MLP (*Multilayer Perceptron*) para estimar a altura das árvores. O treinamento foi realizado com uso do *software* Neuro 4.0, que permite a determinação do número de neurônios na camada oculta. Foram utilizados quatro neurônios na camada oculta, com as variáveis de entrada idade (quantitativa) e DAP (quantitativa), e a variável de saída altura (quantitativa).

Martins et al. (2019b) realizaram um estudo com objetivo de testar diferentes configurações de RNA para estimar altura de *E. grandis*, foram obtidas melhores estimativas utilizando apenas três neurônios na camada oculta da RNA. Segundo Vendruscolo et al. (2015) o uso de três a quatro neurônios na camada oculta da RNA é o suficiente para aplicações em inventário e biometria florestal, pois o excesso de neurônios acarreta em erros maiores, podendo ocorrer superajuste da RNA.

O algoritmo utilizado foi o *Resilient Propagation* (RPROP+), com função de ativação sigmoideal na camada oculta e de saída. Este algoritmo tem como vantagem a facilidade de calcular e adquirir aprendizagem sobre determinado problema, pelo fato de seu ajuste de pesos dependerem mais do sinal dos gradientes de erro, além de ser mais eficiente e recomendado para RNA do tipo *Multilayer Perceptron* (RIEDMILLER & BRAUN, 1993).

Os critérios de parada do treinamento das RNA foram, o número total de ciclos igual a 3.000 ou o erro quadrático médio inferior a 1%, portanto o treinamento foi encerrado quando um dos critérios foi atingido.

A comparação das estimativas obtidas por regressão e RNA, foram realizadas por critérios estatísticos, que foram o coeficiente de correlação ($r_{Y\hat{Y}}$), raiz quadrada do quadrado médio do erro (RQME) e análise gráfica dos resíduos em percentagem (E%), representados pelas seguintes (Equações 1, 2 e 3). Foram atribuídos pesos aos resultados estatísticos calculados de [$r_{Y\hat{Y}}$ e RQME], com valores de 1 a n, onde o menor valor atribuído advindo da

soma dos pesos indicou a equação selecionada.

$$r_{Y\hat{Y}} = \frac{\text{Cov}(Y, \hat{Y})}{\sqrt{S^2(Y)S^2(\hat{Y})}} \quad (1)$$

$$\text{RQME} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (2)$$

$$E\% = \frac{(Y - \hat{Y})}{Y} \cdot 100 \quad (3)$$

Em que: Y = altura observada; \hat{Y} = altura estimada; S^2 = variância; Cov. = covariância; n = número de observações.

RESULTADOS E DISCUSSÃO

Foi verificada a altura mínima das árvores de 11,0 m e máxima de 34,1 m no conjunto de dados de treino, para o conjunto de dados de validação foi verificada altura mínima de 10,5 m e máxima de 33,0 m. Para o diâmetro no conjunto de dados de treino o mínimo foi de 8,0 cm e o máximo de 23,5 cm, para o conjunto de dados de validação foi de 7,6 cm e máximo de 23,0 cm. Para a altura em ambos os conjuntos teve abrangência de idades dos quatro aos vinte e sete anos de idade (Tabela 2).

De acordo com a Tabela 3 os modelos ajustados apresentaram desempenho estatístico satisfatórios, com valores semelhantes $r_{Y\hat{Y}}$ e RQME, com destaque, em relação aos modelos de relação h/d (MR h/d), o modelo MR (h/d)-2, o qual apresentou $r_{Y\hat{Y}}$ de 0,933 e RQME de 1,315 m e VIF de apenas 1,0.

O modelo MR (h/d)-4 também se destacou positivamente, apresentando resultados semelhantes ao MR (h/d)-2, exceto pelo VIF de 6,8. Tendo em vista que valores de VIF inferiores a 10 não prejudicariam o uso da equação, esse modelo também poderia ser utilizado para estimar a altura das árvores de *E. grandis*.

Os modelos MR (h/d)-1 e MR (h/d)-3 apresentaram resultados estatísticos melhores em relação aos modelos MR (h/d)-2 e MR (h/d)-4, porém os mesmos obtiveram VIF maior que 10, ou seja, a multicolinearidade inviabilizou o uso dessas equações.

Os modelos de menor desempenho foram MR (h/d)-5, com $r_{Y\hat{Y}}$ de 0,929 e RQME de 1,359 m. A inserção da variável idade destacou o modelo MR-29, com $r_{Y\hat{Y}}$ de 0,945 e RQME de 1,192 m, o qual também apresentou probabilidade não significativa para seus coeficientes e VIF menor que 10. Além disso, a variável idade proporcionou resultados estatísticos melhores. O resultado estatístico de menor relevância para os modelos com inserção da variável idade foi obtido com o modelo MR-27, com $r_{Y\hat{Y}}$ de 0,932 e RQME de 1,325 m.

Segundo Mendonça et al. (2011), a utilização de mais variáveis além do diâmetro para estimar a altura de árvores possui grande potencial na modelagem florestal, pois aumenta a precisão das estimativas.

Com relação ao treino das RNA, destacaram-se as RNA-31 e RNA-32 com $r_{Y\hat{Y}}$ de 0,952 e RQME de 1,111 m. Os resultados estatísticos obtidos apontam uma maior aplicabilidade das RNA quando comparadas as equações ajustadas tão somente para relação h/d quanto para as equações com inserção da variável idade. Dentre as RNA, a RNA-35 apresentou os piores resultados estatísticos com $r_{Y\hat{Y}}$ de 0,951 e RQME de 1,119 m, e ainda assim, foram resultados superiores aos modelos de regressão ajustados.

Na Figura 1 A são apresentados os resultados dos pesos estatísticos para as equações ajustadas e RNA treinada. Com base no critério estatístico $r_{Y\hat{Y}}$, os piores resultados observados referem-se aos modelos ajustados apenas com relação h/d, em que a equação MR (h/d)-5 alcançou soma de 35 pesos.

Os modelos ajustados com inserção da variável idade tiveram melhores resultados estatísticos quando comparados aos modelos h/d, com menores pesos referentes ao critério $r_{Y\hat{Y}}$ para os modelos MR-9, 14, 19, 24 e 29, os quais apresentaram em comum a variável 1/t. As RNA apresentaram os menores pesos gerais para $r_{Y\hat{Y}}$, sendo os destaques as RNA-31 e 32 com valores de pesos de 1 e 2 respectivamente.

Tabela 2. Estatísticas das variáveis mensuradas nos povoamentos de *E. grandis* localizados em duas mesorregiões no estado do Rio Grande do Sul.

Variável	Tipo	Mínimo	Média	Mediana	Máximo	Desvio Padrão
d (cm)	Treino (n=500)	8,0	16,5	16,6	23,5	2,4
h (m)		11,0	23,5	24,1	34,1	3,7
t (anos)		4,0	16,0	15,0	27,0	5,5
d (cm)	Validação (n=213)	7,6	16,5	16,6	23,0	2,4
h (m)		10,5	23,3	23,6	33,0	3,8
t (anos)		4,0	16,0	15,0	27,0	5,4

Tabela 3. Critérios estatísticos das equações ajustadas (MR) com seus respectivos coeficientes e redes neurais artificiais (RNA) treinadas para estimativa de altura da espécie *E. grandis* em duas mesorregiões no estado do Rio Grande do Sul.

Equação	β_0	β_1	β_2	β_3	Fc	r_{YY}	RQME (m)
MR (h/d)-1	-3,8366 (0,0212) [0,0]	1,8839 (<0,0001) [72,1]	-0,0135 (0,0421) [72,1]			0,934	1,306
MR (h/d)-2	0,2305 (<0,0001) [0,0]	1,0434 (<0,0001) [1,0]			1,0017	0,933	1,315
MR (h/d)-3	3,8176 (0,1532) [0,0]	1,3157 (<0,0001) [13,1]	-32,4518 (0,0974) [13,1]			0,933	1,308
MR (h/d)-4	22,8372 (<0,0001) [0,0]	-120,2953 (<0,0001) [6,8]	0,0293 (<0,0001) [6,8]			0,933	1,315
MR (h/d)-5	0,9700 (<0,0001) [0,0]	0,5366 (<0,0001) [1,0]			1,0018	0,929	1,359
MR-6	-0,7870 (0,6351) [0,0]	1,4406 (<0,0001) [80,0]	-0,0031 (0,6347) [76,5]	0,0864 (<0,0001) [1,6]		0,939	1,252
MR-7	-2,2478 (0,1739) [0,0]	1,6880 (<0,0001) [74,7]	-0,0095 (0,1433) [73,1]	0,0019 (<0,0001) [1,4]		0,937	1,274
MR-8	0,8087 (0,6229) [0,0]	0,9098 (<0,0001) [96,4]	0,0113 (0,0967) [88,0]	1,6780 (<0,0001) [2,2]		0,942	1,219
MR-9	14,3581 (<0,0001) [0,0]	0,0736 (0,7823) [138,1]	0,0350 (<0,0001) [120,0]	-24,2953 (<0,0001) [3,2]		0,944	1,195
MR-10	-0,4855 (0,7664) [0,0]	1,2197 (<0,0001) [86,1]	0,0028 (0,6680) [80,6]	0,7822 (<0,0001) [1,8]		0,941	1,236
MR-11	0,4075 (<0,0001) [0,0]	0,9561 (<0,0001) [1,6]	0,0042 (<0,0001) [1,6]		1,0016	0,938	1,265
MR-12	0,3483 (<0,0001) [0,0]	0,9920 (<0,0001) [1,4]	0,000089 (<0,0001) [1,4]		1,0016	0,936	1,286
MR-13	0,4144 (<0,0001) [0,0]	0,9021 (<0,0001) [1,9]	0,0778 (<0,0001) [1,9]		1,0015	0,941	1,230
MR-14	0,7977 (<0,0001) [0,0]	0,8653 (<0,0001) [2,3]	-0,9494 (<0,0001) [2,3]		1,0014	0,944	1,204
MR-15	0,3985 (<0,0001) [0,0]	0,9304 (<0,0001) [1,7]	0,0374 (<0,0001) [1,7]		1,0015	0,939	1,248
MR-16	1,0923 (0,6732) [0,0]	1,3054 (<0,0001) [13,1]	-8,3134 (0,6630) [13,6]	0,0869 (<0,0001) [1,6]		0,939	1,252
MR-17	3,4891 (0,1808) [0,0]	1,2747 (<0,0001) [13,2]	-25,2719 (0,1867) [13,2]	0,0019 (<0,0001) [1,4]		0,937	1,274
MR-18	-6,2101 (0,0243) [0,0]	1,4107 (<0,0001) [13,4]	31,6832 (0,1084) [15,3]	1,6553 (<0,0001) [2,1]		0,942	1,220
MR-19	-8,1068 (0,0030) [0,0]	1,6370 (<0,0001) [15,2]	102,5778 (<0,0001) [20,7]	-23,9274 (<0,0001) [3,1]		0,945	1,194
MR-20	-2,2436 (0,3968) [0,0]	1,3449 (<0,0001) [13,2]	7,9253 (0,6804) [14,2]	0,7793 (<0,0001) [1,8]		0,941	1,236
MR-21	19,6767 (<0,0001) [0,0]	-92,8573 (<0,0001) [7,3]	0,0293 (<0,0001) [6,8]	0,0905 (<0,0001) [1,6]		0,939	1,254
MR-22	21,7681 (<0,0001) [0,0]	-108,9269 (<0,0001) [6,9]	0,0285 (<0,0001) [6,8]	0,0020 (<0,0001) [1,3]		0,936	1,278
MR-23	13,5889 (<0,0001) [0,0]	-57,8354 (<0,0001) [8,7]	0,0319 (<0,0001) [6,9]	1,7206 (<0,0001) [2,1]		0,942	1,221
MR-24	14,9976 (<0,0001) [0,0]	0,1789 (0,9918) [12,5]	0,0372 (<0,0001) [7,9]	-24,7844 (<0,0001) [3,1]		0,944	1,195
MR-25	16,7624 (<0,0001) [0,0]	-78,3458 (<0,0001) [7,8]	0,0303 (<0,0001) [6,8]	0,8106 (<0,0001) [1,8]		0,940	1,238
MR-26	1,0944 (<0,0001) [0,0]	0,4879 (<0,0001) [1,5]	0,0045 (<0,0001) [1,5]		1,0016	0,935	1,295
MR-27	1,0554 (<0,0001) [0,0]	0,5089 (<0,0001) [1,4]	0,000094 (<0,0001) [1,4]		1,0017	0,932	1,325
MR-28	1,0589 (<0,0001) [0,0]	0,4557 (<0,0001) [1,8]	0,0883 (<0,0001) [1,8]		1,0015	0,940	1,242

Tabela 3. Continuação...

MR-29	1,4792 (<0,0001) [0,0]	0,4313 (<0,0001) [2,1]	-1,1245 (<0,0001) [2,1]	1,0014	0,945	1,192
MR-30	1,0646 (<0,0001) [0,0]	0,4727 (<0,0001) [1,7]	0,0417 (<0,0001) [1,7]	1,0016	0,938	1,271
RNA-31					0,952	1,111
RNA-32					0,952	1,111
RNA-33					0,952	1,115
RNA-34					0,951	1,117
RNA-35					0,951	1,119

h = altura (m); d = diâmetro à altura do peito (cm); β_i = coeficientes de regressão estimados; Fc - fator de correção; $r_{Y\hat{Y}}$ - coeficiente de correlação; RQME - raiz quadrada do quadrado médio do erro; ($Pr > |t|$ dos coeficientes) – probabilidade dos coeficientes; [] - fator de inflação da variância (VIF).

Na Figura 1B são apresentados os resultados dos pesos estatísticos para o critério RQME (m), que foram semelhantes aos obtidos para o critério $r_{Y\hat{Y}}$, com os piores resultados obtidos pelos modelos h/d, bons resultados com os modelos com inserção da idade e os melhores resultados com as RNA.

A Figura 1C apresenta o somatório geral dos pesos de $r_{Y\hat{Y}}$ e RQME para os modelos relação h/d, modelos com inserção da variável idade e RNA. As maiores somas de pesos foram obtidas nos modelos de relação h/d, na qual o MR (h/d)-1 apresentou soma de 60 pesos, o MR (h/d)-5 teve soma de 70 pesos e o modelo de destaque, MR (h/d)-2, apresentou soma de 64 pesos.

Os modelos com inserção da variável idade apresentaram menor somatório de pesos quando comparados aos modelos de relação h/d, no qual se teve destaque para o modelo MR-29 com 12 pesos e maior soma de peso para o modelo MR-27 com pesos.

As RNA apresentaram os menores somatórios de pesos quando comparadas aos modelos de relação h/d e modelos com inserção da variável idade, com destaque para a RNA-31 que apresentou apenas 2 pesos.

Na Figura 2A é apresentada a representação gráfica da altura observada em função da altura estimada e na Figura 2B a distribuição de valores residuais em porcentagem para o ajuste do modelo MR (h/d)-2, que para essa técnica apresentou os melhores resultados estatísticos. É notável que as estimativas foram inferiores as demais, devido aos pontos de altura estimada não recobrirem boa parte dos dados observados. Na verificação do percentual de erro, foi obtido intervalo de -24,6 % a 12,2% de erro.

As Figuras 2C e 2D apresentam as representações gráficas do modelo com inserção da variável idade, no qual o modelo MR-29 apresentou o melhor ajuste. Foi

verificado melhor recobrimento dos pontos de altura estimada, portanto os resultados foram mais precisos comparado ao modelo MR (h/d)-2. Com relação a distribuição residual em porcentagem foi obtido intervalo de -18,3% a 12,0%.

As Figuras 2 E e F tratam das representações gráficas da melhor RNA treinada, que apresentou resultados bastante próximos ao modelo MR-29, recobrando boa parte dos pontos de altura observada, portanto foi uma técnica eficiente. Com relação à distribuição residual, foi obtido intervalo de -18,3% a 11,2%.

Diante dos resultados obtidos, onde as RNA apresentam melhor precisão das estimativas, Alcântara (2015) cita que as ferramentas com uso de inteligência artificial, vêm sendo utilizadas no setor florestal, em substituição aos modelos de regressão.

Foi realizada a validação do melhor modelo de relação h/d (MR (h/d)-2) e melhor modelo com a variável idade (MR-29), para as RNA foram realizadas as validações para as cinco redes treinadas (Tabela 4).

Tabela 4. Resultados estatísticos obtidos na validação do melhor modelo de relação h/d (MR (h/d)-2), melhor modelo com inserção da variável idade (MR-29) e das RNA.

Tipo	$r_{Y\hat{Y}}$	RQME (m)
MR (h/d)-2	0,928	1,373
MR-29	0,936	1,289
RNA-31	0,934	1,313
RNA-32	0,937	1,285
RNA-33	0,941	1,249
RNA-34	0,941	1,238
RNA-35	0,938	1,278

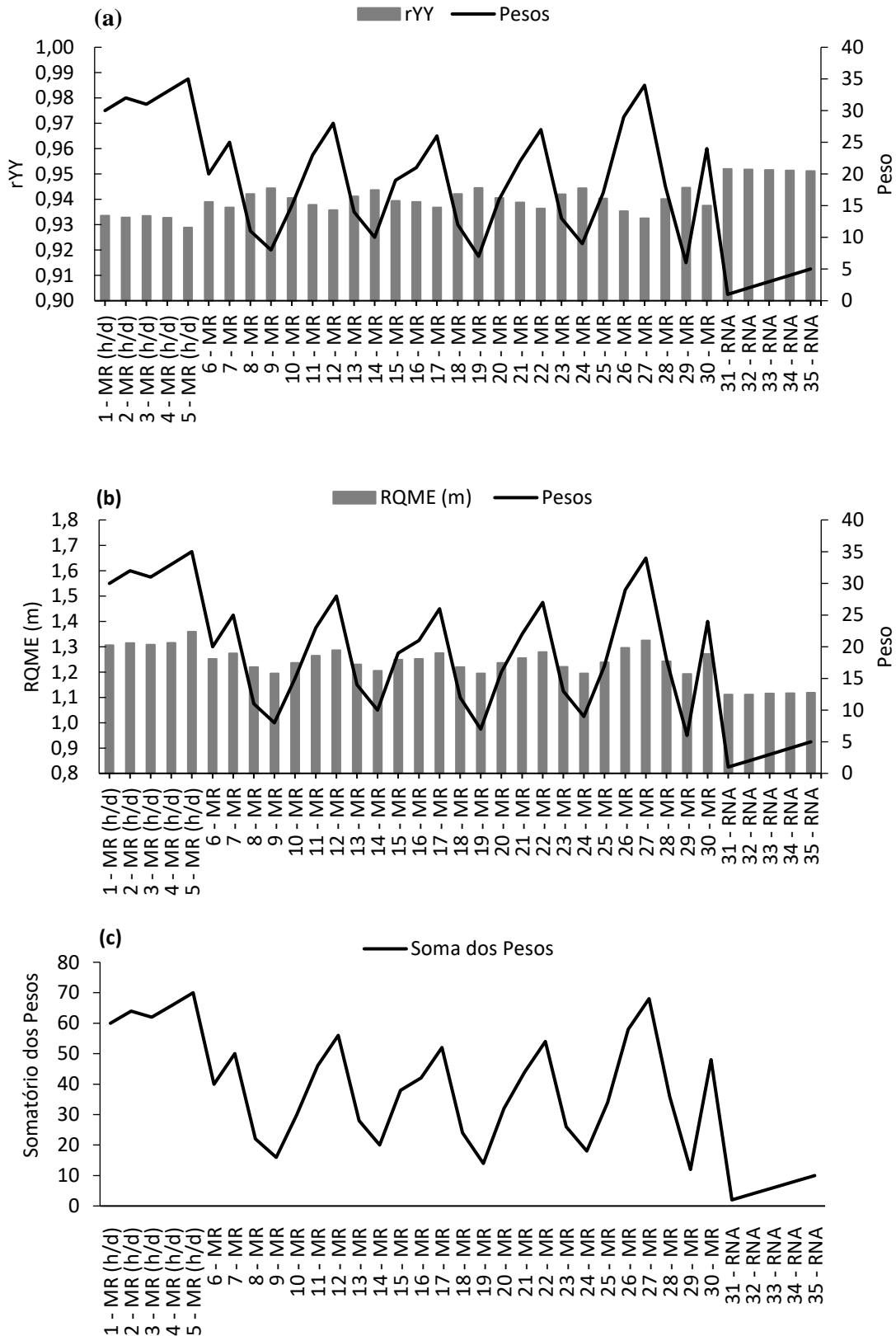


Figura 1. Representações gráficas das equações ajustadas (MR) e redes neurais artificiais (RNA) treinadas para estimativa de altura da espécie *E. grandis*, em duas mesorregiões no estado do Rio Grande do Sul, em relação aos critérios estatísticos de coeficiente de correlação (r_{YY}) (A), raiz quadrada do quadrado médio do erro (RQME) (B) e seus respectivos somatórios de pesos (C).

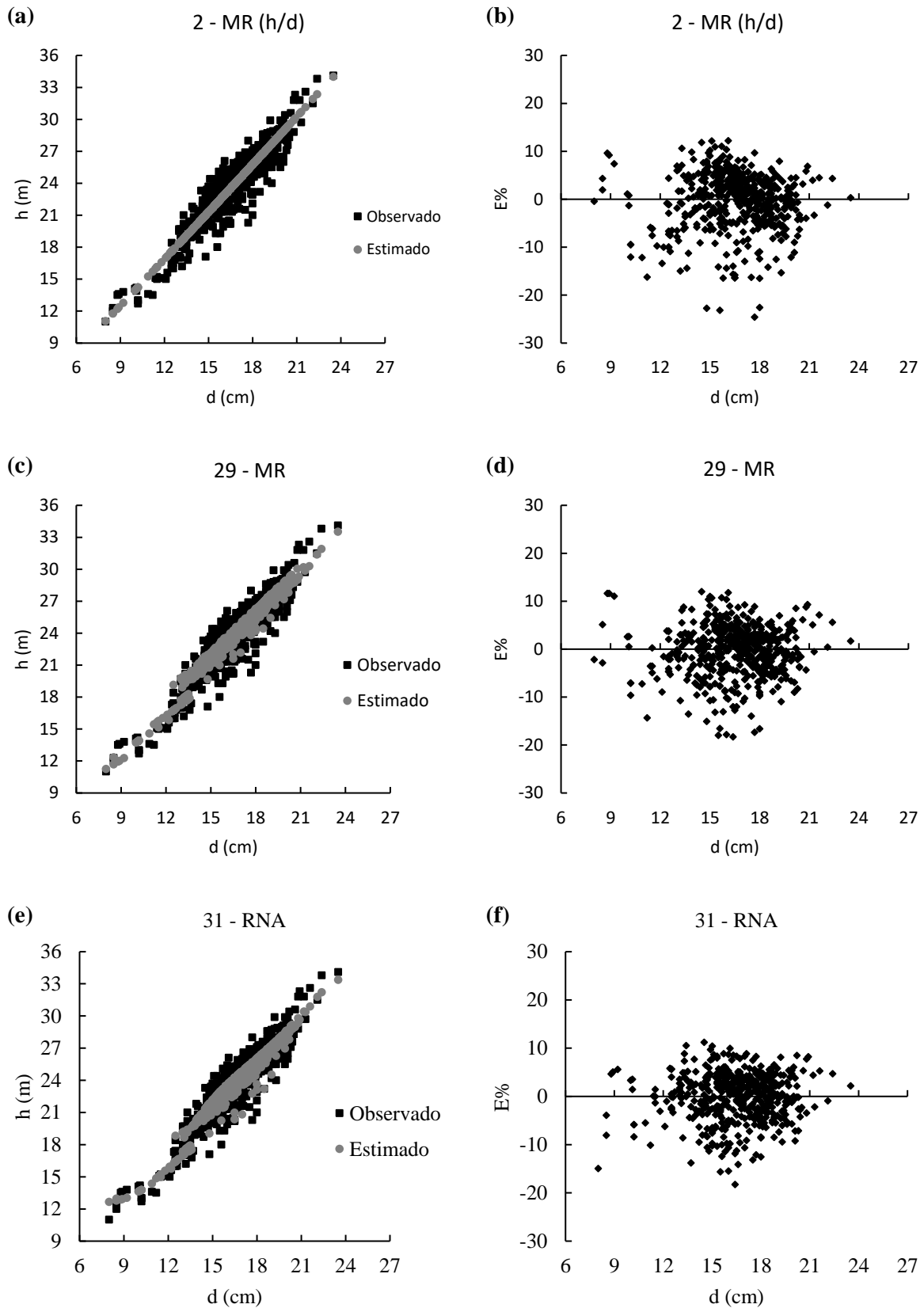


Figura 2. Altura observada versus altura estimada pelas melhores equações ajustadas (MR-h/d)-(A), (MR)-(C) e redes neurais artificiais treinadas (RNA)-(E) , e suas respectivas distribuições de valores residuais (B), (D) e (F) em porcentagem (E%) pelo diâmetro médio (d) observado para a espécie *E. grandis* em duas mesorregiões no estado do Rio Grande Sul.

Com a validação do modelo MR (h/d)-2 foi obtida redução de 0,005 na r_{VY} , para RQME em metros foi obtido aumento de 0,058 m. Para o modelo MR-29 houve redução de 0,009 na r_{VY} e aumento de 0,097 m no RQME.

A RNA-31, que no treinamento apresentou os melhores resultados estatísticos na validação obteve resultado inferior as demais e até mesmo ao melhor modelo com a inserção da variável idade (MR-29). Foi observada redução de 0,018 na r_{VY} e aumento de 0,202 m no RQME. Obtiveram destaque na validação as RNA-33 e 34, ambas com r_{VY} de 0,941 e RQME de 1,249 m e 1,238 m, respectivamente. Esses resultados representam uma redução de 0,011 na r_{VY} de ambas e aumento de RQME de 0,134 m para RNA-33 e 0,121 m para RNA-34.

Para os modelos de regressão não foram observadas diferenças nos valores estatísticos, porém para RNA é de grande importância à realização de validação, pois como foi observado nos resultados obtidos, nem sempre a melhor RNA do treinamento é a mais precisa, pois pode ocorrer o processo de *overfitting* (superajuste). No *software* Neuro, o processo de determinação do número de neurônios na camada oculta é definido pelo usuário, sendo assim, um número excessivo de neurônios pode acarretar a memorização dos dados de treinamento, que leva a acontecer o processo de *overfitting* (VENDRUSCOLO et al., 2015).

Além disso, o procedimento de validação deve ser aplicado para verificar a capacidade de uma rede neural produzir saídas adequadas para entradas que não estavam presentes durante o treinamento (BINOTI et al., 2015). Portanto se a configuração da RNA provoca o processo de superajuste durante o seu treinamento, dificilmente apresentará bons resultados na validação, sendo assim é de grande importância uma configuração adequada da mesma.

Mesmo que não sejam obtidos resultados estatísticos diferentes entre os modelos de regressão e as RNA é notável o potencial e eficiência que essa técnica apresenta.

Nas Figuras 3A e 3B, são apresentados os resultados de altura observada em função da altura estimada e a dispersão residual em percentagem, respectivamente. Em relação ao recobrimento dos pontos que representam as alturas estimadas se manteve o que foi visto no ajuste do modelo. A dispersão residual apresentou intervalo de -21% a 10,7%.

As Figuras 3C e 3D, apresentam as representações gráficas para a validação do modelo MR-29, que também apresentou comportamento semelhante ao observado no ajuste do modelo. A dispersão residual apresentou

intervalo de -19,1% a 9,2%.

Nas Figuras 3 E e F são vistos os resultados obtidos para a RNA-31, que não foi eficiente como se mostrou no treino. Mesmo não sendo a mais eficiente apresentou bom recobrimento das alturas observadas e a dispersão residual teve intervalo de -17,6 a 9,6. As Figuras 3 G e H apresentam as representações gráficas para a segunda melhor RNA, a RNA-32. Houve recobrimento satisfatório das alturas observadas e a dispersão residual teve intervalo de -21,9 a 9,4. As representações gráficas da melhor RNA (RNA-34) são apresentadas nas Figuras 3 I e J, que teve bom recobrimento das alturas observadas e a dispersão residual apresentou intervalo de -17,5 a 9,7.

Todas as RNA validadas apresentaram recobrimento das alturas observadas semelhantes, portanto pode-se dizer que as estimativas de altura para as árvores de *E. grandis* foram eficientes.

Na literatura são encontrados diversos trabalhos comprovando a eficiência das RNA, como é visto em, Campos et al. (2016) que estimaram altura de árvores de pinus e eucalipto com diferentes idades, diferentes ciclos e espaçamentos, os autores verificaram que uma única RNA pode estimar altura de árvores de diferentes espécies em distintas condições, sendo mais eficientes que os modelos convencionais de regressão.

Outros autores como Vendruscolo et al. (2015), Vendruscolo et al. (2017), também comprovaram a eficiência das RNA para estimar a altura de árvores de eucaliptos e teca, obtendo resultados precisos, com erro inferior a 10%, os autores afirmam que tanto as RNA como os modelos convencionais são eficientes, porém as RNA apresentam a vantagem de trabalhar com variáveis qualitativas, melhorando as estimativas.

As RNA também demonstram notável eficiência em estimativas de volumes de árvores, como no estudo realizado por Cordeiro et al. (2015) para estimar volume de árvores da espécie *Acacia mangium* Willd., no qual as redes se mostraram mais acuradas que os modelos volumétricos convencionais.

As RNA são eficientes, apresentando melhor precisão comparadas aos modelos de regressão, podendo ser utilizadas para diversas espécies florestais, como é visto na literatura, além de diversas configurações que podem ser feitas e assim obtendo resultados ainda mais satisfatórios. Portanto as RNA estão ganhando espaço no meio florestal, podendo ser no futuro a técnica mais utilizada para realização de estimativas dendrométricas.

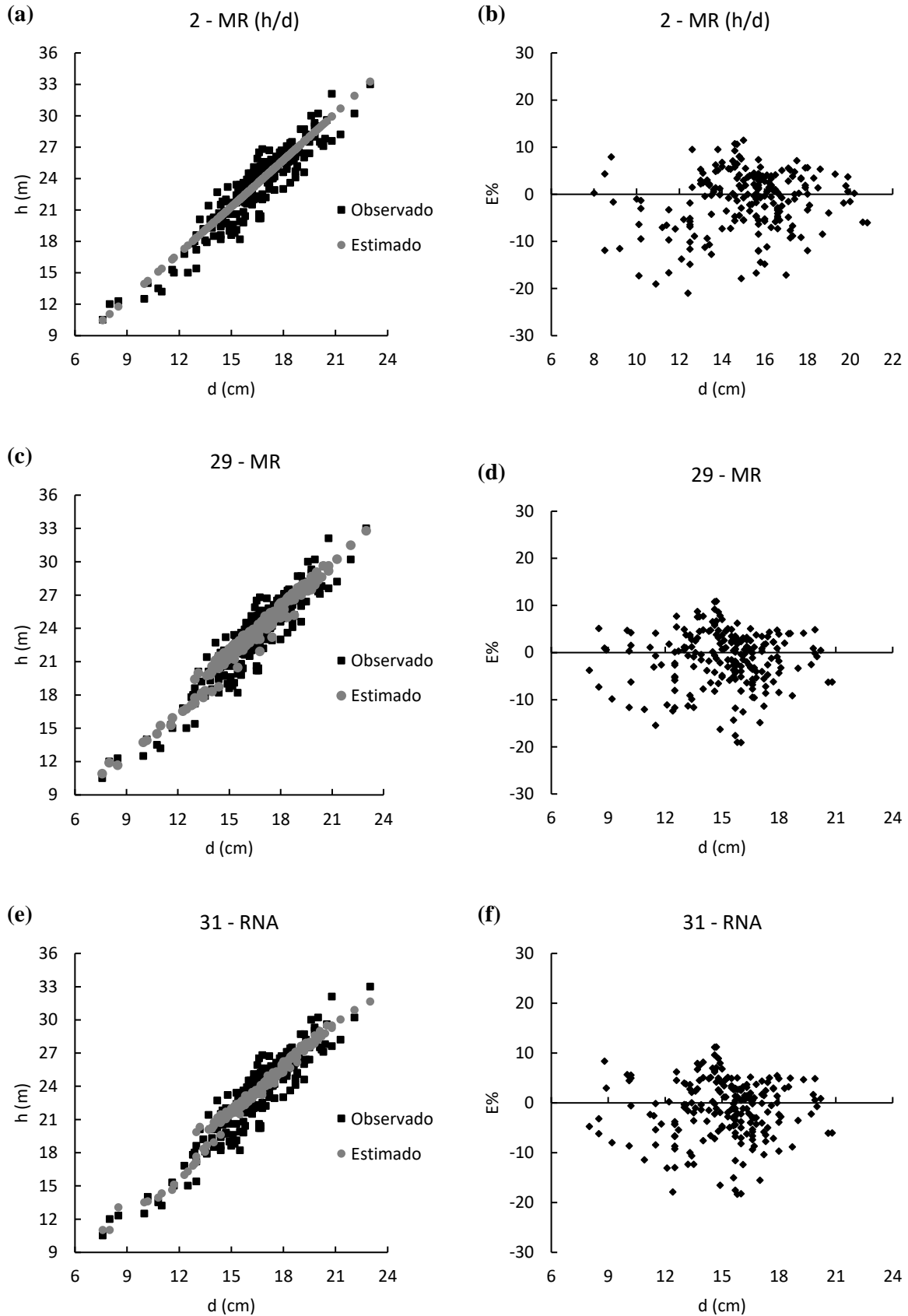


Figura 3. Altura observada versus estimada na validação das melhores equações ajustadas (MR-h/d)-(A), (MR)-(C) e redes neurais artificiais (RNA)-(E,G e I) e suas respectivas distribuições de valores residuais (B, D, F e J) em percentagem (E%) pelo diâmetro médio (d) observado para *E. grandis* em duas mesorregiões no estado do Rio Grande do Sul.

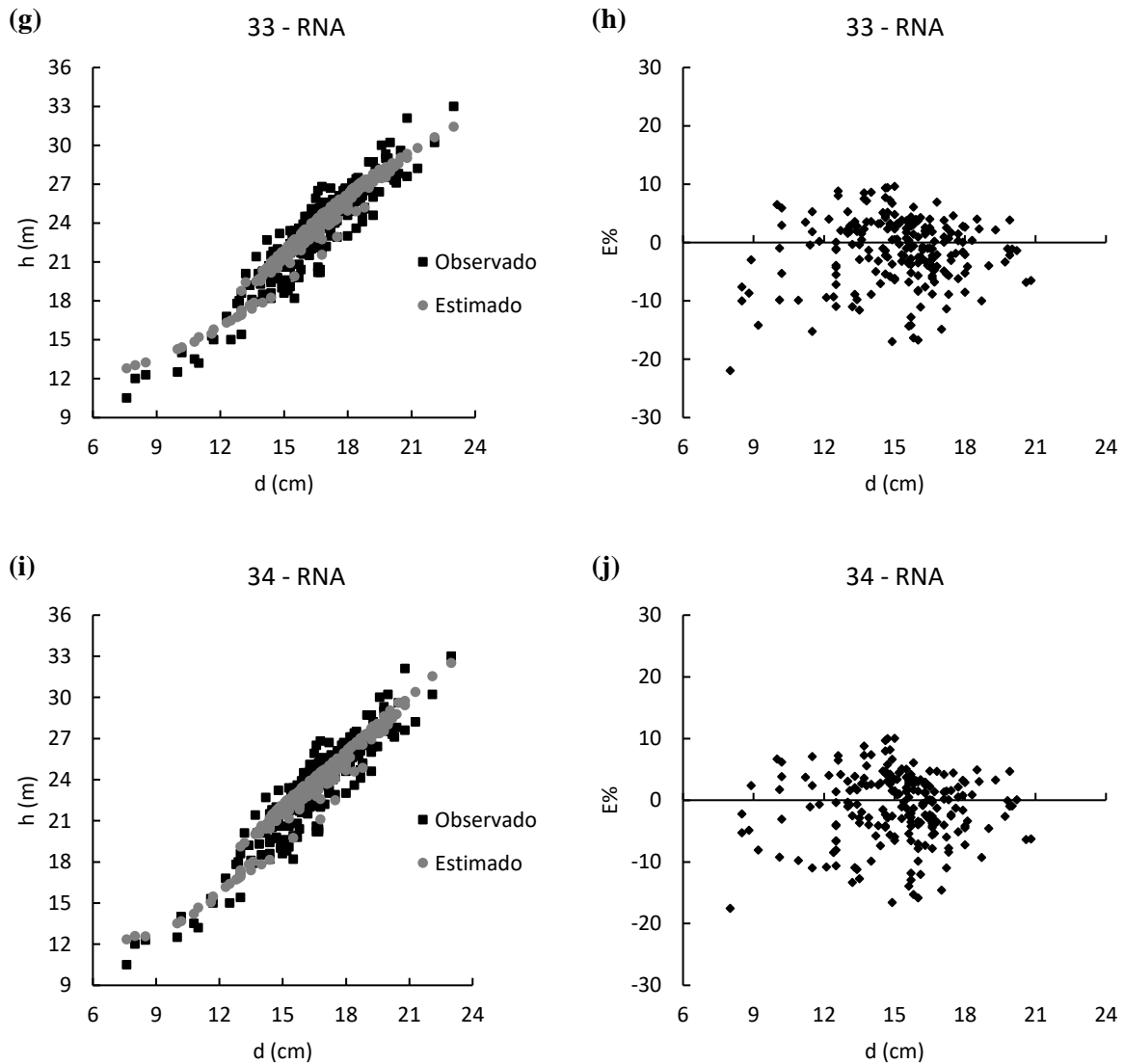


Figura 3. Continuação...

CONCLUSÕES

Ambas as técnicas podem ser utilizadas para estimar a altura das árvores de *E. grandis* nas condições do presente estudo, porém as RNA apresentaram resultados superiores

A multicolinearidade foi um fator importante para determinar o modelo de regressão que poderia ser utilizado para estimar a altura das árvores.

Com a validação foi verificado que nem sempre a melhor RNA obtida no treino apresentará bons resultados quando aplicada em outro conjunto de dados. Com a validação das técnicas foram obtidos resultados mais próximos da realidade.

Observa-se a necessidade da continuidade de estudos relacionados a estimativas de alturas por diferentes técnicas, a saber se o número de observações das

variáveis altura e diâmetro influencia ou não na recomendação de uma técnica específica.

REFERÊNCIAS

- ALCÂNTARA, A.E.M. **Redes neurais artificiais para prognose do crescimento e da produção de povoamentos de eucalipto em Minas Gerais**. 2015. 43p. (Tese de doutorado).
- ALVARES, C.A. et al. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, v.22, n.6, p.711-728, 2013.
- BINOTI, M.L.M.S. et al. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. *Revista árvore*, v.37, n.4, p.639-645, 2013.
- BINOTI, M.L.M.S. et al. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. *Cerne*, v.21, n.1, p.97-105, 2015.
- BULLINARIA, J.A. **Introduction to neural computation**. 2014.

Disponível em: <http://www.cs.bham.ac.uk/~jxb/inc.html>.

Recebido em 30-08-2019

Aceito em 03-12-2019

CAMPOS, B.P.F. et al. Predição da altura total de árvores em plantios de diferentes espécies por meio de redes neurais artificiais. **Pesquisa Florestal Brasileira**, v.36, n.88, p.375-385, 2016.

CARVALHO, S.P. **Métodos alternativos de estimação de coeficientes de trilha e índices de seleção, sob multicolinearidade**. 1994. 163p. (Tese de doutorado).

CORDEIRO, M.A. et al. Estimativa do volume de *Acacia mangium* utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. **Pesquisa Florestal Brasileira**, v.35, n.83, p.255-261, 2015.

CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2003.

FERREIRA, J.C.B et al. Altura de mudas da *Tibouchina granulosa* COGN. (Melastomaceae) estimada por redes neurais artificiais. **Revista da Sociedade Brasileira de Arborização Urbana**, v.9, n.1, p.151-160, 2014.

FREUND, R.J.; WILSON, W.J. **Regression analysis: statistical modeling of a response variable**. San Diego: Academic, 1998.

LEMOS, R.C. et al. **Levantamento e reconhecimento dos solos do estado do Rio Grande do Sul**. Recife: Ministério da Agricultura, 1973.

MARTINS, M.T. et al. Estimativa de altura para povoamentos de eucalipto no Rio Grande do Sul por meio de diferentes configurações de Redes Neurais Artificiais. **Agrarian Academy**, v.6, n.11, p.162-174, 2019b.

MARTINS, M.T. et al. Uso de diferentes técnicas para descrição da altura de um plantio monoespecífico de angico-vermelho. **Agrarian Academy**, v.6, n.11, p.340-352, 2019a.

MENDONÇA, A.R. et al. Estimação da altura total de árvores de ipê felpudo utilizando modelos de regressão e redes neurais artificiais. **Revista Brasileira de Biometria**, v.36, n.1, p.128-139, 2018.

MENDONÇA, A.R. et al. Modelos hiposométricos e de crescimento em altura das árvores dominantes e codominantes para *Pinus caribaea* var. hondurensis. **Scientia Forestalis**, v.39, n.90, p.151-160, 2011.

RIEDMILLER, M.; BRAUN, H. A direct adaptive method for faster back propagation learning: The RPROP algorithm. **IEEE International Conference**, p.586-591, 1993.

SCHNEIDER, P.R. et al. **Análise de regressão aplicada à engenharia florestal**. 2 ed. Santa Maria: FACOS-UFSM, 2009.

VENDRUSCOLO, D.G.S. et al. Estimativa da altura de árvores de *Tectona grandis* L.f. utilizando regressão e redes neurais artificiais. **Nativa**, v.5, n.1, p.52-58, 2017.

VENDRUSCOLO, D.G.S. et al. Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. **Revista Brasileira de Biometria**, v.33, n.4, p.556-569, 2015.