

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
CÂMPUS DE CHAPADÃO DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

JEAN DE JESUS DA SILVA

**ESTIMATIVAS DO VOLUME DE ÁRVORES DE EUCALIPTO
COM DIFERENTES TÉCNICAS DE MODELAGEM E TAMANHO
DO BANCO DE DADOS**

CHAPADÃO DO SUL – MS

2022

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
CÂMPUS DE CHAPADÃO DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

JEAN DE JESUS DA SILVA

**ESTIMATIVAS DO VOLUME DE ÁRVORES DE EUCALIPTO
COM DIFERENTES TÉCNICAS DE MODELAGEM E TAMANHO
DO BANCO DE DADOS**

Orientador: Dr. Gileno Brito de Azevedo

Dissertação apresentada à Universidade Federal de Mato Grosso do Sul, como requisito para obtenção do título de Mestre em Agronomia, área de concentração: Produção Vegetal.

CHAPADÃO DO SUL – MS

2022



Serviço Público Federal
Ministério da Educação
Fundação Universidade Federal de Mato Grosso do Sul



PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

CERTIFICADO DE APROVAÇÃO

DISCENTE: Jean de Jesus da Silva

ORIENTADOR: Dr. Gileno Brito de Azevedo

TÍTULO: Estimativas do volume de árvores de eucalipto com diferentes técnicas de modelagem e tamanho do banco de dados

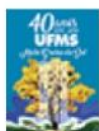
AVALIADORES:

Prof. Dr. Gileno Brito de Azevedo

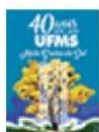
Prof. Dr. Paulo Eduardo Teodoro

Prof. Dr. Renato Vinicius Oliveira Castro

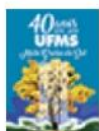
Chapadão do Sul, 24 de fevereiro de 2022.



Documento assinado eletronicamente por **Gileno Brito de Azevedo**, Professor do Magisterio Superior, em 24/02/2022, às 10:31, conforme horário oficial de Mato Grosso do Sul, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Paulo Eduardo Teodoro**, Professor do Magisterio Superior, em 24/02/2022, às 10:31, conforme horário oficial de Mato Grosso do Sul, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Renato Vinicius Oliveira Castro**, Usuário Externo, em 24/02/2022, às 11:04, conforme horário oficial de Mato Grosso do Sul, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

LISTA DE TABELAS

Tabela 01. Síntese das estatísticas de qualidade para as redes com melhor desempenho obtido durante a etapas de treinamento e validação.....	14
Tabela 02. Desempenho de estimativa e coeficientes obtidos pelo ajuste do modelo de Schumacher e Hall nas etapas de ajuste e validação.....	15

LISTA DE FIGURAS

Figura 01. Boxplot de distribuição dos dados dendrométricos volume (V), altura (H), diâmetro à altura do peito (D) obtidos e selecionados para treinamento (T) e validação (V) para cada clone (C1: AEC0144, C2: AEC244, C3: VM01 e C4: H77) e rotação (R1: primeira rotação e R2: segunda rotação)	11
Figura 02. Análise de qualidade das estatísticas dos modelos e RNA's para treinamento e validação; MOD sem = Somente variável quantitativa; RNA sem = Somente variável quantitativa; RNA com = Variáveis quantitativas e qualitativas; Narv = Número de árvore. r = Coeficiente de determinação; RMSE% = Raiz quadrada do erro médio; MAE% = Média absoluta do erro; MBE = Erro de viés médio.....	16
Figura 03. Dispersão dos resíduos do erro relativo (er - %) em função do volume observado para cada banco de dados (V – m ³); RNA's para validação; MOD sem = Somente variável quantitativa; RNA sem = Somente variável quantitativa; RNA com = Variáveis quantitativas e qualitativas; 38 = 10%; 75 = 20%; 112 = 30%; 149 = 40%; 186 = 50%; 223 = 60%; 261 = 70%; 297 = 80%; 335 = 90%; 372 = 100%.....	17
Figura 04. Análise do comportamento do erro relativo (er - %) em função dos clones, rotações e banco de dados.....	18

RESUMO

A quantificação do volume das árvores é atividade fundamental no manejo dos recursos florestais. Contudo, é uma atividade bastante onerosa para obtenção de forma direta. Como alternativa surge a utilização de métodos indiretos como os modelos de regressão e as redes neurais artificiais (RNAs). Portanto, objetivou-se avaliar a precisão das estimativas do volume de árvores de eucaliptos utilizando diferentes técnicas de modelagem e tamanho do banco de dados. Os objetivos específicos foram: (i) avaliar se a redução do número de árvores utilizadas na modelagem do volume influencia na precisão dessas estimativas; (ii) avaliar se o desempenho das RNAs é superior ao dos modelos de regressão; e (iii) avaliar se a inclusão de variáveis categóricas nas RNAs contribui para melhoria das estimativas. Os dados do estudo foram obtidos em oito plantios clonais de eucalipto (quatro clones x duas rotações), implantados no município de Ribas do Rio Pardo, Estado de Mato Grosso do Sul. Foram cubadas rigorosamente 465 árvores para obter o volume. Desse total, 20% foram destinados à validação das estimativas e 80% para o ajuste dos modelos e treinamento das RNAs. Para avaliar o efeito do tamanho do banco de dados sobre a precisão das estimativas, o treinamento foi realizado com diferentes frações dos dados destinados ao treinamento (10% a 100%). Para cada fração do banco de dados foi ajustado o modelo volumétrico de Schumacher e Hall e treinadas 500 RNAs do tipo Multilayer Perceptron. Com a redução do tamanho do banco de dados utilizado na modelagem do volume de árvores de eucalipto, por modelos de regressão e RNAs, foi possível manter a precisão das estimativas. Independentemente do tamanho do banco de dados, a precisão das estimativas do volume geradas por RNAs foram ligeiramente superiores às dos modelos de regressão. A inclusão de variáveis explicativas qualitativas nas RNAs proporcionou estimativas de volume ligeiramente superiores às aquelas que não utilizaram essas variáveis.

Palavras-chave: modelos volumétricos; rede neurais artificiais; inventário florestal.

ABSTRACT

The quantification of the volume of trees is a fundamental activity in the management of forest resources. However, it is a very expensive activity to obtain directly. As an alternative, there is the use of indirect methods such as regression models and artificial neural networks (ANNs). Therefore, the objective was to evaluate the precision of the estimates of the volume of eucalyptus trees using different modeling techniques and database size. The specific objectives were: (i) to assess whether the reduction in the number of trees used in volume modeling influences the accuracy of these estimates; (ii) assess whether the performance of ANNs is superior to that of regression models; and (iii) to assess whether the inclusion of categorical variables in the ANNs contributes to improving the estimates. The study data were obtained in eight eucalyptus clonal plantations (four clones x two rotations), implanted in the municipality of Ribas do Rio Pardo, State of Mato Grosso do Sul. 465 trees were rigorously cubed to obtain the volume. Of this total, 20% were used to validate the estimates and 80% to adjust the models and train the ANNs. To assess the effect of the size of the database on the accuracy of the estimates, training was performed with different fractions of the training data (10% to 100%). For each fraction of the database, the volumetric model of Schumacher and Hall was adjusted and 500 ANNs of the Multilayer Perceptron type were trained. By reducing the size of the database used to model the volume of eucalyptus trees, using regression models and ANNs, it was possible to maintain the precision of the estimates. Regardless of the size of the database, the accuracy of volume estimates generated by ANNs was slightly higher than that of regression models. The inclusion of qualitative explanatory variables in the ANNs provided slightly higher volume estimates than those that did not use these variables.

Keywords: volumetric models; Artificial neural networks; Forest Inventory.

SUMÁRIO

INTRODUÇÃO.....	8
MATERIAIS E MÉTODOS.....	9
ÁREA DE ESTUDO	9
OBTENÇÃO DOS DADOS	10
FRACIONAMENTO DO BANCO DE DADOS	10
AJUSTE DE MODELO DE REGRESSÃO	12
REDES NEURAIS ARTIFICIAIS	12
AVALIAÇÃO DA QUALIDADE DAS ESTIMATIVAS.....	13
RESULTADOS	13
DISCUSSÃO.....	18
CONCLUSÕES	20
REFERÊNCIAS	21

INTRODUÇÃO

O Brasil é líder mundial na produção de aço a partir de carvão vegetal e o segundo maior produtor de celulose do mundo, com a produção de 19,7 milhões de toneladas fabricadas (IBÁ, 2020). A atividade florestal também representa 1,2% do PIB Nacional e receita bruta total de R\$ 97,4 bilhões. Conta com uma área de 6,97 milhões de hectares em plantios de eucalipto, os quais apresentam uma produtividade média de 35,3 m³ ha⁻¹ ano⁻¹. É considerada uma das maiores do mundo, subsidiada por investimentos constantes em pesquisa para desenvolvimento de produtos que estejam alinhados à bioeconomia, retratando o perfil de indústria de olho no futuro (IBÁ, 2020).

A representação global exercida pelo Brasil é explicada pelos avanços nos métodos de produção e nas tecnologias aplicadas. Desse modo, para o gerenciamento de um plantio comercial é fundamental a quantificação do crescimento e da produção (LEITE; ANDRADE, 2002; CASTRO et al., 2015; SANTOS et al., 2017). O prognóstico do desenvolvimento é realizado por meio da modelagem da produção e do crescimento, em que o volume é estimado em função de diversas variáveis, tais como: o diâmetro a altura do peito (DAP) obtido a 1,30 m do solo, a altura, a idade do povoamento e índices de local (AZEVEDO et al., 2015; CASTRO et al., 2016).

Entre os métodos de determinação de volume de árvores em pé, destaca-se em precisão o método da equação de volume, contrapondo-se aos métodos de volume cilíndrico e da área basal (BINOTI et al., 2013). A vantagem das equações de volume é o cálculo de volume sólido, árvore a árvore, por meio de modelos estatísticos, especialmente testados para apresentar erros factíveis (THOMAS et al., 2006; CAMPOS; LEITE 2013). Entretanto, as equações de volume, cujos modelos incluem como variável independente, a altura e o DAP da árvore, são mais gerais, podendo abranger sítios diferentes (MARTINS et al., 2015). Diversos modelos de regressão lineares e não lineares foram investigados para a estimação do volume de povoamentos florestais (LEITE, 2011), contudo, alguns modelos e situações podem apresentar tendenciosidades e níveis de erro inaceitáveis para os objetivos do manejo florestal (BINOTI et al., 2014).

Novas ferramentas têm sido introduzidas na área florestal com o intuito de aumentar a precisão das estimativas e diminuir os custos, entre elas destacam-se as Redes Neurais Artificiais (RNAs) (BINOTI et al., 2009; GORGENS et al., 2009; SILVA et al., 2009; BINOTI et al., 2014). As RNAs são uma tentativa de representação de um modelo

computacional do sistema nervoso, a fim de simular a capacidade de aprendizado do cérebro humano. Os primeiros estudos surgiram com o advento dos primeiros computadores eletrônicos, em meados de 1943 (WARREN; WALTER, 1943).

Assim, as RNAs são modelos estatísticos baseados em modelos biológicos neurais, formadas por vários neurônios conectados se tornando um modelo robusto e distribuído que pode ser usada como alternativa à regressão (HAYKIN, 2001; BRAGA et al., 2007; GORGENS et al., 2009; SILVA et al., 2009). Estas unidades, ou neurônios, computam funções matemáticas e seus resultados são processados em conjunto nas demais camadas da rede (HAYKIN, 2001; GORGENS et al., 2009). As conexões simulam sinapses biológicas e possuem pesos associados às suas entradas. Esses pesos são ajustados à medida que o conjunto todo é treinado (FERNANDES et al., 2018).

Em razão do elevado custo de cubagem de árvores nas empresas florestais, quanto menor for o tamanho do banco de dados a ser utilizado na modelagem volumétrica, desde que suficientes para manter a precisão das estimativas, mais economicamente viável será para as empresas florestais. Diante do exposto, este estudo objetivou avaliar a precisão das estimativas do volume de árvores de eucaliptos utilizando diferentes técnicas de modelagem e tamanho do banco de dados. Os objetivos específicos foram:

- i. Avaliar se a redução do número de árvores utilizadas na modelagem do volume influencia na precisão dessas estimativas.
- ii. Avaliar se o desempenho das redes são superiores aos modelos de regressão.
- iii. Avaliar se nas redes neurais a inclusão de variáveis qualitativas contribui para melhoria das estimativas.

MATERIAIS E MÉTODOS

ÁREA DE ESTUDO

Os dados do estudo foram obtidos em plantios clonais de quatro clones de eucalipto (C1: AEC0144, C2: AEC244, C3: VM01 e C4: H77), cultivados em primeira (R1) e segunda rotação (R2) cada, no espaçamento 3 x 3 m, com idades que variam de 5,9 a 12,9 anos. Os plantios foram conduzidos em áreas do Grupo Mutum no município

de Ribas do Rio Pardo, Estado de Mato Grosso do Sul, com altitude de 370 m e nas coordenadas 20° 19' 14'' S e 53° 17' 28'' W.

A pluviosidade média anual para região é de aproximadamente 1.400 mm ao ano, com maior acúmulo pluviométricos entre os meses de dezembro a março e menor entre julho a setembro (WEATHER SPARK, 2020). A temperatura média está entre 27 °C e 21 °C para as estações quente e frias do ano respectivamente. Destaca-se que as maiores temperaturas estão entre os meses de dezembro a março e as menores temperaturas entre julho a setembro (ALVAREZ et al., 2013).

OBTENÇÃO DOS DADOS

Foram selecionadas aleatoriamente de 40 a 50 árvores em cada plantio (clone/rotação), totalizando 465 fustes para obtenção do volume. As árvores escolhidas foram cubadas pelo método de Smalian (SOARES et al., 2011), com auxílio de uma trena e suta. Foram mensurados os diâmetros com casca nas posições de 0,15m; 0,40m; 0,70m; 1,0m; 1,30m; 2,0m; e, a partir desta posição, as medições foram realizadas a cada dois metros até o diâmetro comercial de 5 centímetros. Também foi obtida a altura total (H) de cada fuste.

FRACIONAMENTO DO BANCO DE DADOS

Inicialmente os dados foram divididos em treinamento e validação, com 372 fustes destinados ao treinamento (80%) e 93 à validação (20%). Os dados foram divididos de forma aleatória dentro de cada clone e rotação, de modo a garantir a mesma representatividade de cada um no banco de dados. Em cada clone e rotação os dados de treinamento e validação apresentam distribuições semelhantes (Figura 01).

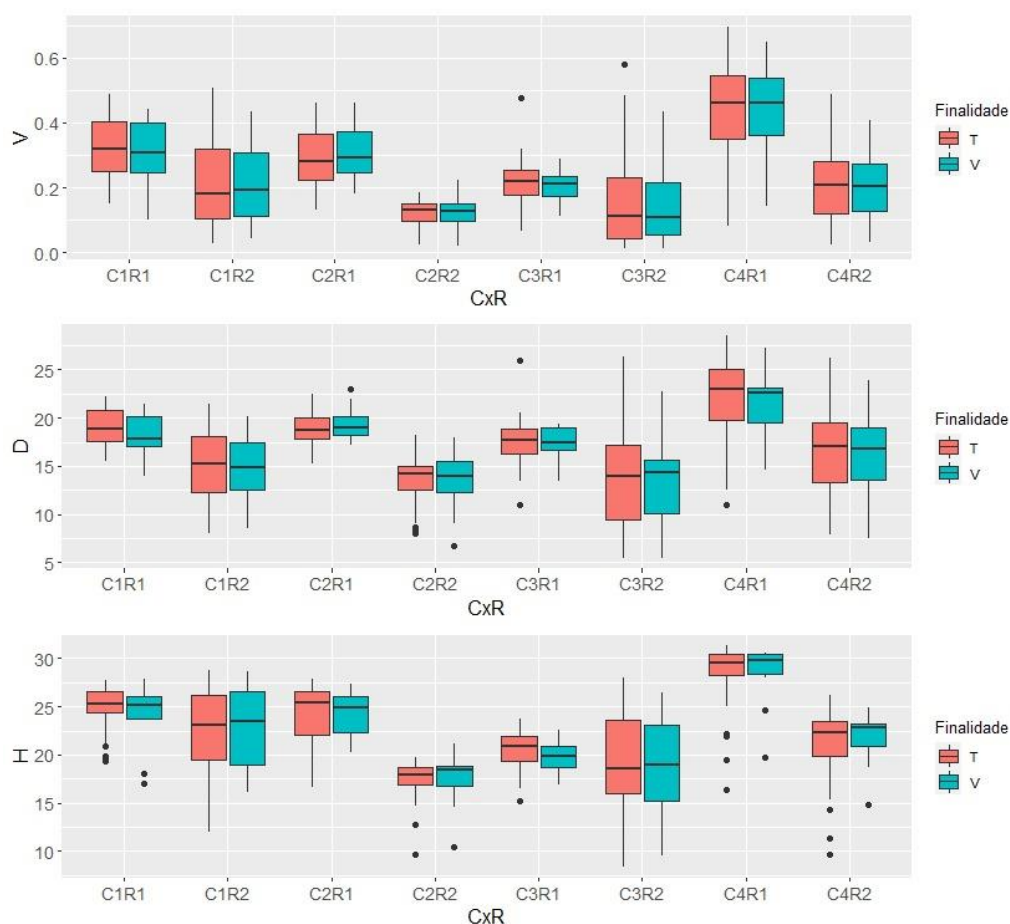


Figura 01. Boxplot de distribuição dos dados dendrométricos volume (V), altura (H), diâmetro à altura do peito (D) obtidos e selecionados para treinamento (T) e validação (V) para cada clone (C1: AEC0144, C2: AEC244, C3: VM01 e C4: H77) e rotação (R1: primeira rotação e R2: segunda rotação).

Para avaliar o efeito do tamanho do banco de dados sobre a precisão das estimativas do volume do fuste, o treinamento foi realizado com diferentes frações dos dados destinados ao treinamento: 10% (38 fustes), 20% (75 fustes), 30% (112 fustes), 40% (149 fustes), 50% (186 fustes), 60% (223 fustes), 70% (261 fustes), 80% (297 fustes), 90% (335 fustes) e 100% (372 fustes). O treinamento foi realizado separadamente para cada fração, enquanto para a validação foi mantido 20% do total do banco de dados para essa finalidade, sendo este independente dos dados do treinamento.

AJUSTE DE MODELO DE REGRESSÃO

Para cada fração do banco de dados foi ajustado o modelo volumétrico proposto por Schumacher e Hall (1933):

$$V = \beta_0 * DAP^{\beta_1} * H^{\beta_2}$$

onde: V= volume com casca (m³), DAP = Diâmetro à altura do peito (cm), H = Altura total (m).

O ajuste dos modelos foi realizado com auxílio do software R, por meio do pacote *nlme* e função *nls* (R CORE TEAM, 2021).

REDES NEURAIS ARTIFICIAIS

Para cada fração do banco de dados também foram treinadas RNAs com auxílio da ferramenta *Intelligent Problem Solver* do *software* STATISTICA 7.0 (STATSOFT, 2007). A topologia de rede adotada foi a *Multilayer Perceptron* (MLP), sendo utilizada a função de ativação logística. As camadas de entrada foram compostas pelas variáveis quantitativas (DAP e H) e/ou qualitativas (rotação, bifurcação e clone). A camada intermediária foi composta por no máximo 10 neurônios e a camada de saída foi composta por Volume com casca. Foram testadas RNAs, com ou sem as variáveis qualitativas como variáveis preditoras, sendo:

$$V = f(DAP, H)$$

$$V = f(DAP, H, R, B, C)$$

onde: V = Volume com casca (m³), DAP = Diâmetro a altura do peito (cm), H = Altura total (m), R = Rotação (Primeira ou segunda rotação), B = Bifurcação (Com ou sem bifurcação, na segunda rotação), C = Clone.

Na etapa de treinamento, para cada fração do banco de dados, com e sem variáveis qualitativas, foram treinadas 500 RNAs, totalizando 10.000 RNAs treinadas no estudo. Foram utilizadas para a validação apenas as 5 de melhor desempenho em cada condição, totalizando a validação de 100 redes.

AVALIAÇÃO DA QUALIDADE DAS ESTIMATIVAS

A precisão das estimativas nas etapas de treinamento e validação foi avaliada a partir dos seguintes critérios estatísticos: Correlação de Pearson – r ; Raiz do erro médio quadrático – RMSE (%); Erro absoluto médio – MAE (%); Erro de viés médio – MBE (%); Distribuição de resíduos (AZEVEDO et al., 2020).

$$r = \frac{COV(y_i, \hat{y}_i)}{\sigma_{y_i} * \sigma_{\hat{y}_i}}$$

$$RMSE(\%) = \frac{100}{\bar{Y}} \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}$$

$$MAE(\%) = \frac{100}{\bar{Y}} \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

$$MBE(\%) = \frac{100}{\bar{Y}} \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}$$

$$Resíduo(\%) = \frac{y_i - \hat{y}_i}{y_i} * 100$$

onde: y_i = valores observados, $\hat{y}_i$ = valores estimados, $\bar{y}$ = média dos valores observados, Σ = Somatória, n = número de observações.

Para facilitar a visualização do comportamento das estatísticas de precisão das estimativas em função do tamanho do banco de dados no ajuste, as mesmas foram plotadas graficamente. Também foi realizada a análise gráfica dos resíduos para avaliação de possíveis tendências nas estimativas.

RESULTADOS

O modelo de Schumacher e Hall proporcionou desempenho semelhante nas estimativas de volume geradas a partir de diferentes tamanhos de bancos de dados. Na validação, as correlações entre os valores observados e estimados variaram de 0,981 a 0,989, RMSE variou de 9,075% a 11,769%, MAE variou de 6,414% a 8,666% e MBE variou de -0,868% a 1,956% (Tabela 1).

Tabela 01. Desempenho de estimativa e coeficientes obtidos pelo ajuste do modelo de Schumacher e Hall nas etapas de ajuste e validação.

B.D (%)	β_0	β_1	β_2	R	Treinamento			R	Validação		
					RMSE (%)	MAE (%)	MBE (%)		RMSE (%)	MAE (%)	MBE (%)
10	8,16E-05	1,061	1,576	0,983	11,105	8,313	-0,837	0,981	11,769	8,666	1,956
20	5,80E-05	1,640	1,163	0,986	10,243	6,485	-0,924	0,988	9,351	6,632	-0,868
30	6,78E-05	1,283	1,436	0,982	11,539	8,210	-0,683	0,987	9,601	6,986	0,476
40	5,65E-05	1,471	1,320	0,983	11,257	7,553	-0,348	0,989	9,132	6,441	0,866
50	7,40E-05	1,358	1,342	0,982	11,406	7,927	-0,771	0,988	9,306	6,822	-0,203
60	6,03E-05	1,339	1,421	0,982	11,557	7,860	-0,609	0,988	9,342	6,689	0,670
70	6,30E-05	1,312	1,432	0,982	11,620	7,993	-0,690	0,988	9,460	6,823	0,650
80	5,75E-05	1,419	1,364	0,983	11,126	7,429	-0,555	0,989	9,075	6,419	0,264
90	6,72E-05	1,375	1,355	0,981	11,679	8,033	-0,638	0,988	9,282	6,661	0,414
100	6,35E-05	1,383	1,365	0,983	11,318	7,722	-0,534	0,988	9,243	6,592	0,553

B.D. = Banco de dados; r = Coeficiente de correlação; RMSE% = Raiz quadrada do erro média; MAE% = Média absoluta do erro; MBE = Erro de viés médio; Modelo de Schumacher & Hall ajustada; β_0 , β_1 e β_3 = coeficientes de ajuste do modelo.

As estatísticas obtidas para as RNAs em diferentes bancos de dados foram ligeiramente superiores às obtidas pelos modelos de regressão. Na tabela 2 são apresentadas as RNAs de melhor desempenho na etapa de validação, obtidas no treinamento para cada banco de dados com e sem a inclusão de variáveis qualitativas. Para as RNAs sem a inclusão de variáveis qualitativas a correlação variou de 0,988 a 0,992, RMSE variou de 7,785% a 9,396%, MAE variou de 5,756% a 6,316% e MBE variou de 0,333% a 1,740%. Já quando incluídas as variáveis qualitativas, esses valores foram de 0,989 a 0,993 para correlação, 7,160% a 9,017% para RMSE, 4,983% a 6,137% para MAE e 0,293% a 2,158% para MBE. De forma geral houve melhoria da precisão das estimativas com a inclusão das variáveis qualitativas nas RNAs.

Tabela 02. Síntese das estatísticas de precisão para as redes com melhor desempenho obtido durante a etapas de treinamento e validação.

B.D	CAT	Treinamento				Validação			
		R	RMSE %	MAE %	MBE %	R	RMSE %	MAE %	MBE %
10	S	0,992	7,784	6,040	-0,070	0,988	9,396	6,316	1,740
10	C	0,995	6,041	4,427	0,062	0,989	9,017	5,613	2,158
20	S	0,987	9,778	6,413	-0,026	0,992	7,785	5,855	0,348
20	C	0,992	7,713	5,308	0,125	0,993	7,305	5,242	0,293
30	S	0,986	10,134	6,609	-0,014	0,991	7,970	5,809	0,800
30	C	0,992	7,862	5,026	0,176	0,990	8,551	6,137	1,494
40	S	0,985	10,372	6,984	-0,077	0,991	8,047	6,033	1,447
40	C	0,991	8,184	5,707	0,049	0,992	7,586	5,441	1,354
50	S	0,987	9,739	6,264	-0,010	0,991	7,902	5,756	0,333
50	C	0,989	8,925	5,829	-0,027	0,993	7,160	5,116	0,517
60	S	0,984	10,926	7,041	0,015	0,991	8,128	5,854	1,239
60	C	0,987	9,797	6,717	-0,107	0,992	7,550	5,488	1,405
70	S	0,985	10,444	6,724	-0,038	0,991	8,005	5,835	1,206
70	C	0,989	9,088	5,992	0,045	0,993	7,317	4,983	1,556
80	S	0,985	10,527	6,799	0,037	0,991	7,940	5,830	0,913
80	C	0,988	9,361	6,378	0,032	0,993	7,301	5,297	1,219
90	S	0,985	10,431	6,755	0,006	0,991	8,005	5,881	1,083
90	C	0,989	8,892	5,942	0,069	0,992	7,493	5,258	1,467
100	S	0,986	10,231	6,595	0,040	0,991	7,944	5,831	1,165
100	C	0,989	8,939	5,996	-0,004	0,993	7,358	5,174	1,152

B.D. = Banco de dados (%); r = Coeficiente de correlação; RMSE% = Raiz quadrada do erro média; MA% = Média absoluta do erro; Rot = Rotação; MBE = Erro de viés médio.

De forma geral, nos modelos de regressão e nas RNAs sem a inclusão de variáveis qualitativas, as estimativas obtidas a partir de 20% do banco de dados apresentaram desempenho semelhante entre si e superiores aos obtidos com 10% do banco de dados, além de apresentar melhor desempenho na validação (Figura 2). Quando incluídas as variáveis qualitativas nas RNAs, observou-se que foi necessário um banco de dados a partir de 40% para garantir melhor desempenho nas estimativas.

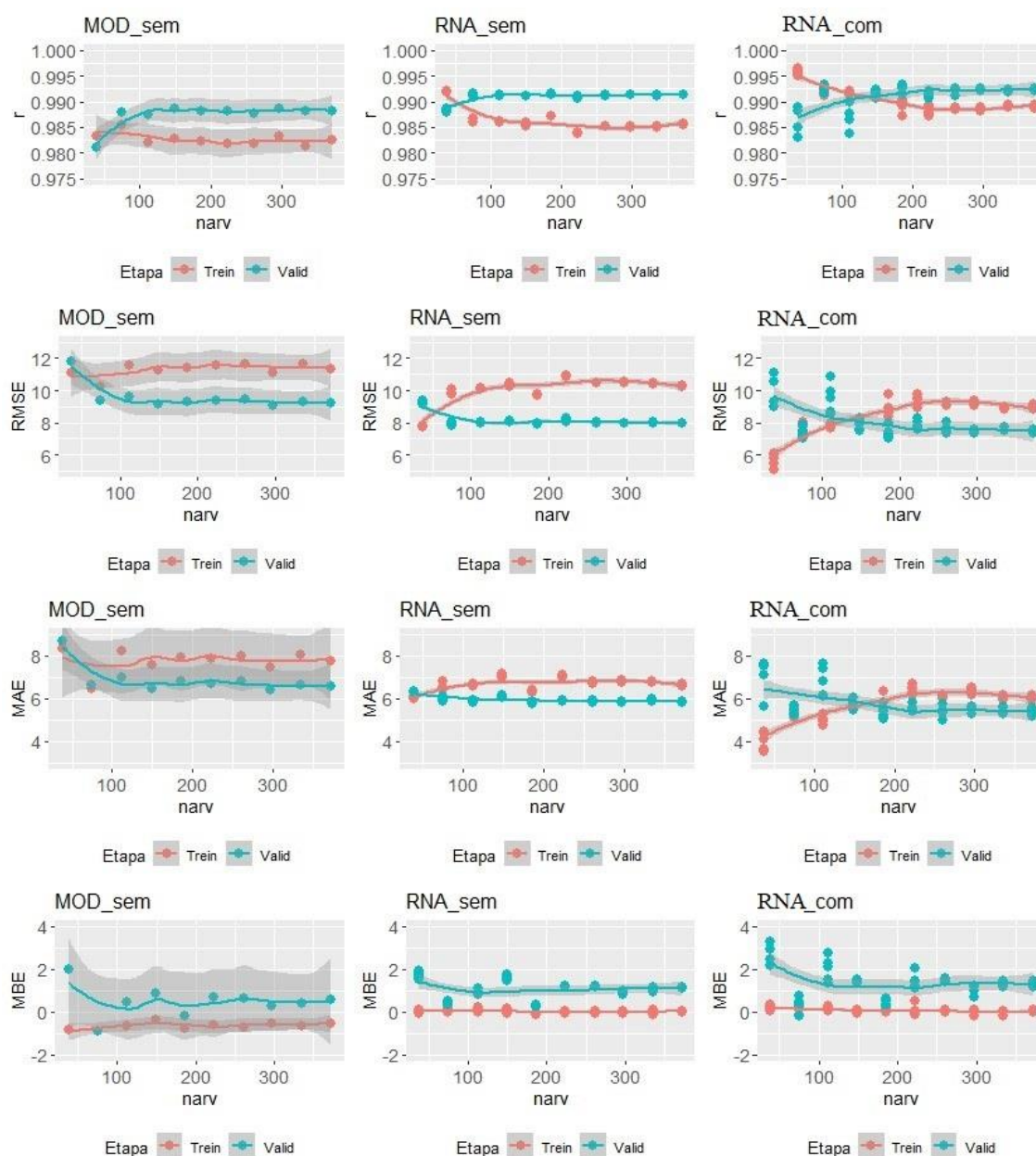


Figura 02. Análise de qualidade das estatísticas dos modelos e RNA's para treinamento e validação; MOD sem = Somente variável quantitativa; RNA sem = Somente variável quantitativa; RNA com = Variáveis quantitativas e qualitativas; Narv = Número de árvore. r = Coeficiente de determinação; RMSE% = Raiz quadrada do erro médio; MAE% = Média absoluta do erro; MBE = Erro de viés médio.

A análise gráfica de distribuição dos resíduos para as estimativas realizadas no banco de dados da validação demonstrou a maior dispersão dos valores para a estimativa dos menores volumes, independentemente do método e tamanho do banco de dados

(Figura 03). Os modelos de regressão apresentaram maiores tendenciosidades nas estimativas, superestimando os menores volumes e subestimando os maiores volumes.

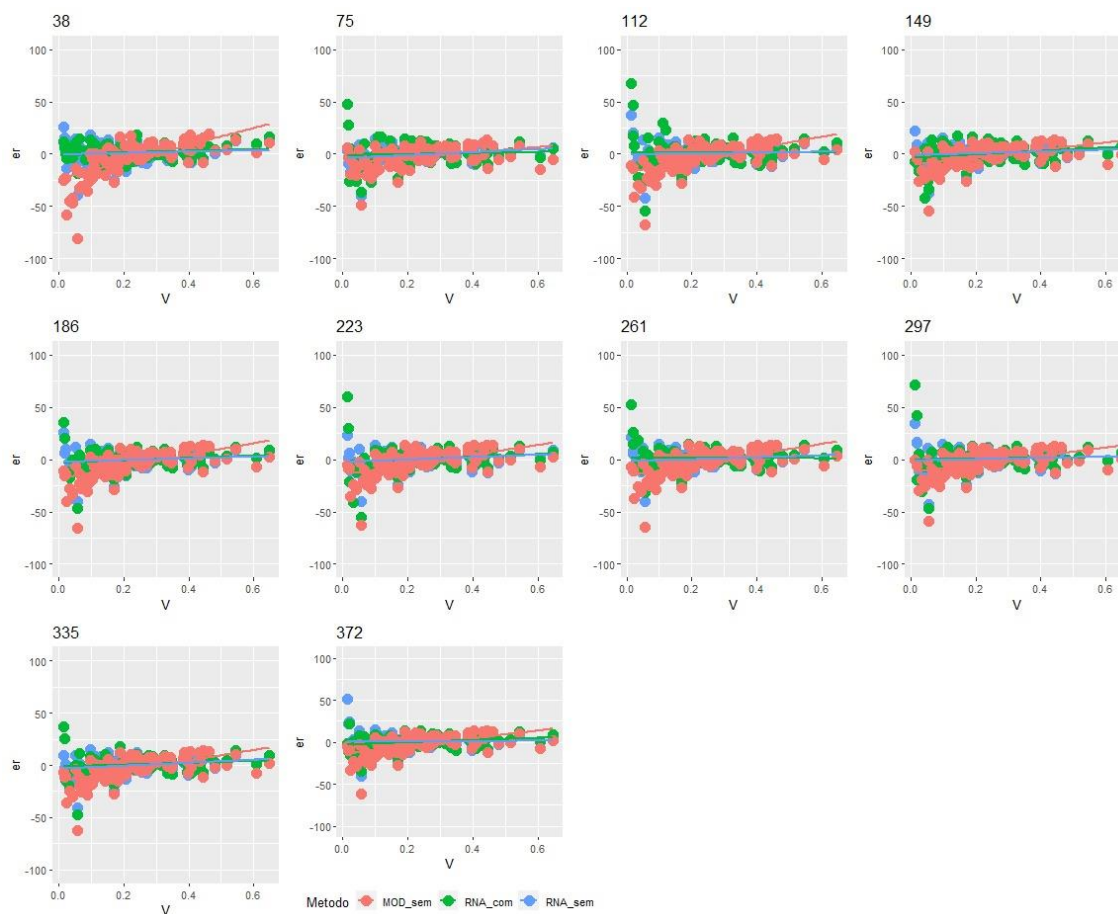


Figura 03. Dispersão dos resíduos do erro relativo (er - %) em função do volume observado para cada banco de dados (V – m³); RNA's para validação; MOD sem = Somente variável quantitativa; RNA sem = Somente variável quantitativa; RNA com = Variáveis quantitativas e qualitativas; 38 = 10%; 75 = 20%; 112 = 30%; 149 = 40%; 186 = 50%; 223 = 60%; 261 = 70%; 297 = 80%; 335 = 90%; 372 = 100%.

Na Figura 04 é apresentada análise gráfica dos resíduos para condições específicas da área de estudo em função da modelagem volumétrica realizada no presente estudo. Normalmente na segunda rotação houve maior dispersão dos resíduos em relação à primeira rotação, independente da estratégia utilizada na modelagem, do tamanho do banco de dados e do clone (Figuras 04a, 04b e 04d). Em relação às estratégias utilizadas, a maior dispersão dos resíduos ocorreu para a RNA com a inclusão de variáveis categóricas (Figuras 04a, 04c, 04d e 04f). Em relação ao tamanho do banco de dados, a

dispersão dos resíduos se mostrou semelhante entre os clones e estratégias de modelagem (Figura 04b e 04f), apesar da modelagem por regressão apresentar ligeira tendência em superestimar os volumes.

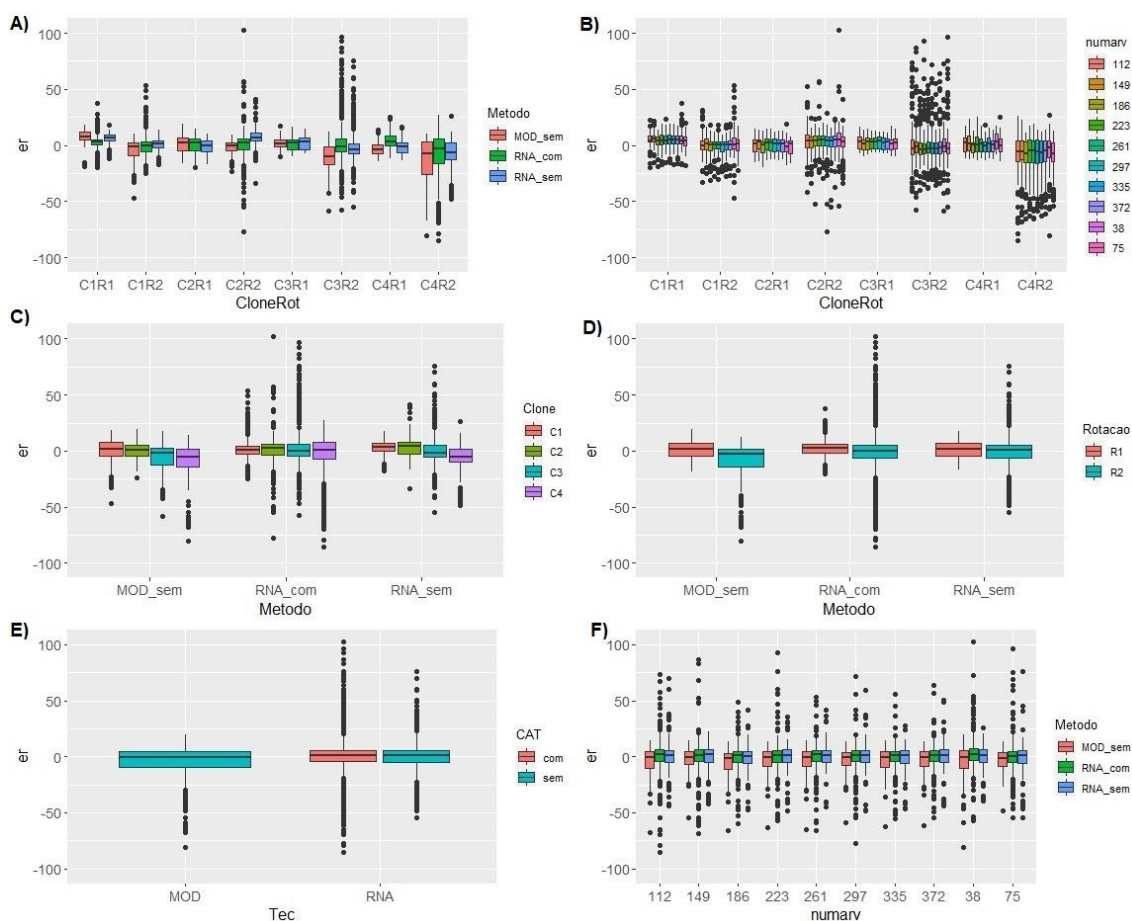


Figura 04. Análise do comportamento do erro relativo (er - %) em função dos clones, rotações e banco de dados.

DISCUSSÃO

A redução do tamanho do banco de dados utilizado na modelagem do volume de árvores de eucalipto proporcionou estimativas precisas tanto para a modelagem por regressão quanto por RNAs. Quando consideradas apenas a modelagem utilizando variáveis explicativas quantitativas, a utilização do banco de dados com número de observações a partir de 20% (73 fustes) dos dados completos (372 fustes) já foram suficientes para obter estimativas com precisão semelhantes às obtidas com o banco de

dados completos. Para as RNAs com inclusão de variáveis explicativas qualitativas, esse comportamento foi obtido com bancos de dados a partir de 40% (149 fustes).

Em trabalho desenvolvido por Freitas e Andrade (2017) com 35 árvores-amostra cubadas em plantios de *Corymbia citriodora* com 49 meses de idade, ao avaliar a variação de amplitude de classes diamétricas e do número de árvores-amostra cubadas, verificou-se que houve uma redução de 28,57% no número árvores cubadas. Essa redução é bem inferior a obtida no presente estudo. Isso pode ser explicado principalmente por diferenças metodológicas entre os estudos, uma vez que no presente estudo trabalhou com um banco de dados com maiores números de árvores cubadas, além da modelagem ser realizada de forma generalista para diferentes clones e rotações. Para Souza et al. (2018), os indicadores de precisão não dependem apenas do tamanho da amostra, mas também da variância, e consequentemente, as amostras pequenas tendem a apresentar um erro amostral maior, com resultados menos confiáveis.

Dessa forma, independentemente da técnica utilizada para a modelagem do volume, os resultados do presente estudo mostram a importância do planejamento adequado do número de árvores a serem utilizadas na cubagem rigorosa, o que pode contribuir para a redução de custos nos inventários florestais. A cubagem rigorosa é uma atividade trabalhosa e onerosa (MIGUEL et al., 2018; OLIVEIRA et al., 2020) e, consequentemente, estratégias que permitam reduzir a quantidade de árvores a serem cubadas são de grande relevância no processo de quantificação do volume de madeira nos povoamentos florestais.

As estimativas do volume obtida por regressão e RNAs se mostraram satisfatórias, com boa precisão das estimativas e ausência de tendências. Porém, de forma geral, as estimativas geradas pelas RNAs se mostraram ligeiramente superiores às do modelo de Schumacher e Hall. Esse comportamento também foi observado por diferentes autores ao gerar estimativas de volume para diferentes condições, como clones de eucalipto (SOUZA et al., 2018; TAVARES JÚNIOR et al., 2019; LEAL et al., 2020), *Khaya ivorensis* (LOPES et al., 2020) e de espécies nativas (MIGUEL et al., 2015; SILVA et al., 2020), além de múltiplos volumes de árvores de clones de eucalipto em diferentes rotações (AZEVEDO et al., 2020).

Tavares Júnior et al. (2019) citam que as RNAs têm vantagens sobre modelos de regressão lineares devido a fatores como a capacidade de aprender e generalizar,

tolerância ao ruído e modelagem de relações não lineares entre variáveis. Lopes et al. (2020) ainda destacaram vantagens das RNAs em relação aos modelos de regressão, pois uma única rede pode estimar o volume para diferentes locais, dispensando necessidade de estratificação.

Para Souza et al. (2018), a RNA com a adição de uma variável categórica (tipo de clone) na modelagem não apresentou diferenças perceptíveis, não demonstrando ser necessária para a estimação do volume de árvores de eucalipto. Estes autores ainda mencionaram a adição de variáveis categóricas aumentaram o número de neurônios na camada oculta da RNA, o que pode comprometer a habilidade de generalização da RNA, e assim, é importante a avaliar a real necessidade dessa variável para a estimar o volume das árvores. No presente estudo, a inclusão de variáveis explicativas qualitativas (variáveis categóricas) nas RNAs gerou estimativas precisas do volume, e de forma geral, ligeiramente melhores do que aquelas que não utilizaram as variáveis categóricas. Porém, é importante destacar que a inclusão de variáveis qualitativas nas RNAs reduz a sua aplicabilidade a situações mais específicas, o que pode ser indesejável em algumas situações.

CONCLUSÕES

A redução do tamanho do banco de dados utilizado na modelagem do volume de árvores de eucalipto, por modelos de regressão e RNAs, manteve a precisão das estimativas.

Independentemente do tamanho do banco de dados, a precisão das estimativas do volume geradas por RNAs foram ligeiramente superiores às dos modelos de regressão.

A inclusão de variáveis explicativas qualitativas nas RNAs proporcionou estimativas de volume ligeiramente superiores às aquelas que não utilizaram essas variáveis, porém, deixa a RNA menos generalista.

REFERÊNCIAS

- ALVAREZ, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, 2013.
- AZEVEDO, L. P. A.; COSTA, R. B.; MARTINEZ, D. T.; FILHO, A. A. T.; BRONDANI, G. E.; BARETTA, M. C.; AJALA, W. V. Seleção em progênies de *Eucalyptus camaldulensis* em área de cerrado mato-grossense. **Ciência Florestal**, v. 45, n. 11, p. 2001-2006, 2015.
- AZEVEDO, G. B.; TOMIAZZI, H. E.; AZEVEDO, G. T. O. S.; TEODORO, L. P. R.; TEODORO, P. E.; SOUZA, M. T. P.; BATISTA, D. S.; EUFRATE-JUNIOR, H. J. Multi-volume modeling of *Eucalyptus* trees using regression and artificial neural networks. **PLOS ONE**, v. 15, 2020.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDEMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. 2 ed. Rio de Janeiro: LTC, p. 260, 2007.
- BINOTI, M. L. M. S.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Ajuste do modelo de Schumacher e Hall e aplicação de redes neurais artificiais para estimar volume de árvores de eucalipto. **Revista Árvore**, v. 33, n. 6, p. 1133 - 1139, 2009.
- BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G.; SILVA, A. Redução dos custos em inventário de povoamentos equiâneos. **Revista Brasileira de Ciências Agrárias**, v. 8, n. 1, p. 125 - 129, 2013.
- BINOTI, M. L. M. D. S.; BINOTI, D. H. B.; LEITE, H. G.; GARCIA, S. L. R.; FERREIRA, M. Z.; RODE, R.; SILVA, A. A. L. D. Redes neurais artificiais para estimação do volume de árvores. **Revista Árvore**, p. 283 - 288, 2014.
- CAMPOS, J. C. C.; LEITE, H. G. Mensuração florestal: perguntas e respostas. 4.ed., Viçosa: **Editora UFV**, p. 605, 2013.
- 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**, v. 43, n. 105, p. 83-90, 2015.

CASTRO, R. V. O.; ARAÚJO, R. A. A.; LEITE, H. G.; CASTRO, A. F. N. M.; SILVA, A.; PEREIRA, R. S.; LEAL, F. A. Modelagem do Crescimento e da Produção de Povoamentos de Eucalyptus em Nível de Distribuição Diamétrica Utilizando Índice de Local. **Revista Árvore**, v. 40, n. 1, p. 107-116, 2016.

FERNANDES, M. M.; SOUSA, F. L.; SILVA, J. P. M.; ARAÚJO, E. F.; FERNANDES, M. R. M.; NÓBREGA, R. S. A. Redes Neurais Artificiais na estimação de variáveis biométricas de mudas de espécies florestais produzidas em diferentes substratos. **Revista de Ciência Agroveterinárias**, 2018.

FREITAS, B. C.; ANDRADE, V. C. L. Bancos de dados de cubagem rigorosa de árvores-amostra de *Corymbia citriodora* empregados no ajuste de modelos volumétricos regionais. **Advances in Forestry Science**, v. 4, n. 2, p. 93-97, 2017.

GORGENS, E. B.; LEITE, H. G.; SANTOS, H. N.; GLERIANI, J. M. Estimação do Volume de Árvores Utilizando Redes Neurais Artificiais. **R. Árvore**, v. 33, n. 6, p. 1141-1147, 2009.

HAYKIN, S. Redes neurais: princípios e práticas. Porto Alegre: **Bookman**; 2001.

IBÁ, **Industria Brasileira de Árvores**. Relatório IBA 2020. 2020.

LEAL, F. A.; LEAL, G. S. A.; SILVA, T. C. Redes neurais artificiais e modelos alométricos aplicados para estimativa de volume e altura em *Eucalyptus urophylla* S. T. Blake. **Advances in Forestry Science**, v. 7, n. 3, p. 1181-1188, 2020.

LEITE, H. G.; ANDRADE, V. C. L.; Um método para condução de inventários florestais sem o uso de equações volumétricas. **Revista Árvore**, v. 26, n. 3, p. 321-328, 2002.

LEITE, D. R.; MIGUEL, E. P.; SANTOS, G. A.; ENCINAS, J. M. I.; REZENDE, A. V. Análise comparativa entre dois procedimentos para estimativa da variável volume em um povoamento de *Eucalyptus urophylla*, no município de Mineiros, Goiás. **Enciclopédia biosfera**, v. 7, n. 13, p. 1684-1700, 2011.

LOPES, L. S. S.; RODE, R.; PAULETTO, D.; BALONEQUE, D. D.; SANTOS, F. G.; SILVA, A. R.; BINOTI, D. H. B.; LEITE, H. G. Uso de regressão e redes neurais artificiais na estimativa do volume de *Khaya ivorensis*. **Ciência da Madeira (Brazilian Journal of Wood Science)**, 2020.

MARTINS, A. P. M.; CARVALHO, D. E.; SCHMIDT, L. N.; GORENSTEIN, M. R.; CYSNEIROS, V. C. Eficiência de métodos para estimativa do volume total de *Corymbia citriodora*. **Scientia Agraria Paranaensis**, v. 15, n. 2, p. 178-184, 2015.

MIGUEL, E. P.; NETTO, S. P.; AZEVEDO, G. B.; AZEVEDO, G. T. O. S.; REZENDE, A. V.; PEREIRA, R. S. Alternative methods of scaling *Eucalyptus urophylla* trees in forest stands: compatibility and accuracy of volume equations. **IForest**, v. 11, p. 275-283. 2018.

MÜLLER, M. D.; SALLES, T. T.; PACIULLO, D. S. C.; BRIGHENTI, A. M.; CASTRO, C. R. T. Equações de altura, volume e afilamento para eucalipto e acácia estabelecidos em sistema silvipastoril. **Revista Floresta**, v. 44, n. 3, p. 473-484, 2014.

MIGUEL, E. P.; REZENDE, A. V.; LEAL, F. A.; MATRICARDI, E. A. T.; VALE, A. T.; PEREIRA, R. S. Redes neurais artificiais para a modelagem do volume de madeira e biomassa do cerradão com dados de satélite. **Pesquisa Agropecuária Brasileira**, v. 50, n. 9, p. 829-839, 2015.

OLIVEIRA, J. M. D.; CARDOSO, L. S.; SILVA, L. C.; OLIVEIRA, L. S.; ARAÚJO, E. J.; MONTE, M. A. Quantificação de volume de árvores utilizando cubagem rigorosa, modelo de afilamento e modelo volumétrico. **Agropecuária Científica no Semiárido**. v. 16, n. 2, p. 134-137, 2020.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available in: <<https://www.R-project.org>> (Acessado em agosto, 2021).

SCHUMACHER, F. X.; HALL, F. S. Logarithmic expression of timber-tree volume. **Journal of Agricultural Research**, v. 47, n. 9, p. 719-734, 1933.

SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Adjustment of the Schumacher and Hall model and application of artificial neural networks to estimate volume of eucalypt trees. **Revista Árvore**. 2009.

SILVA, M. V. C.; RAMOS, Y. A.; LIMEIRA, M. M. C.; COELHO, M. C. B.; SANTOS, A. F.; GIONGO, M.; ERPEN, M. L. Uso de redes neurais artificiais e modelos de regressão para estimar volume de espécies nativas em Portel, Pará – Brasil. **Revista Forestal Mesoamericana Kurú**. 2020.

SOARES, C. P. B.; PAULA NETO. F.; SOUZA, A. L. **Dendrometria e inventário florestal**. 2. Ed. Viçosa: Editora UFV. p. 272, 2011.

SOUZA, S. R. R.; SILVA, J. A. A.; FERREIRA, T. A. E.; GUERA, O. G. M. Redes Neurais para Estimativa Volumétrica de Clones de Eucalyptus Spp.No Pólo Gesseiro do Araripe. **Revista Brasileira de Biometria**, v. 36, n. 3, p. 715-729, 2018.

SPARK, W. Clima Brasil. Disponível em <[https://es.441weatherspark.com/y/19346/clima-promedio-en-guayaquilecuador-durante-todo-el-año](https://es.weatherspark.com/y/19346/clima-promedio-en-guayaquilecuador-durante-todo-el-año)>, Acesso em 10 de outubro de 2021.

STATSOFT, INC. 2007. 'STATISTICA (data analysis software system)', Version, 7: 1984-2004.

TAVARES JÚNIOR, I. S.; ROCHA, J. E. C.; EBLING, Â. A.; CHAVES, A. S.; ZANUNCIO, J. C.; FARIAS, A. A.; LEITE, H. G. Artificial Neural Networks and Linear Regression Reduce Sample Intensity to Predict the Commercial Volume of Eucalyptus Clones. **Forests**, v. 10, p. 268, 2019.

THOMAS, C.; ANDRADE, C. M.; SCHNEIDER, P. R.; FINGER, C. A. G. Comparação de equações volumétricas ajustadas com dados de cubagem e análise de tronco. **Ciência Florestal**, v. 16, n. 3, p. 319-327, 2006.

WARREN, S. M.; WALTER, H. P. A Logical Calculus Of The Ideas Immanent In Nervous Activity. **Bulletin of Mathematical Biophysics**. v. 5, p. 115-133, 1943.