



Estimativas volumétricas em povoamentos de eucalipto utilizando máquinas de vetores de suporte e redes neurais artificiais

Volumetric estimates in eucalyptus stands using support vector machines and artificial neural networks

Márcio Assis Cordeiro^{1*}, Julio Eduardo Arce², Fabiane Aparecida Retslaff Guimarães¹, Izabel Passos Bonete¹, Anthoinny Vittória dos Santos Silva³, Jadson Coelho de Abreu³ e Daniel Henrique Breda Binoti⁴

¹ Universidade Estadual do Centro-Oeste-Unicentro. Irati, Paraná, Brasil.

² Universidade Federal do Paraná. Curitiba, Paraná, Brasil.

³ Universidade do Estado do Amapá. Departamento de Engenharia Florestal. Macapá, Amapá, Brasil.

⁴ Universidade Federal de Viçosa. Viçosa, Minas Gerais, Brasil.

* autor de correspondência.
marcio.engflorestal@gmail.com

RESUMO

Este estudo teve por objetivo avaliar o desempenho de redes neurais artificiais (RNA) e máquinas de vetor de suporte (MVS) na modelagem volumétrica em povoamentos de eucalipto. Utilizou-se dados oriundos de plantios comerciais não desbastados, localizados em quatro municípios na mesorregião sul do estado do Amapá e foram disponibilizados por uma empresa privada. Foram ajustados modelos volumétricos consagrados na literatura e comparados com a técnica de MVS e de RNA. Os dados foram divididos em 80% para treinamento e 20% para validação dos modelos, as mesmas variáveis dendrométricas utilizadas pelos modelos de regressão (dap e altura) foram utilizadas pelas MVS e RNA. Para o treinamento e generalização das MVS, foram utilizadas quatro configurações, formadas a partir de duas funções de erro e duas funções de *Kernel*. Para configuração, treinamento e generalização das RNA, foi utilizado o software *NeuroForest - Volumetric*, no qual foram utilizadas configurações de redes do tipo Adaline (*Adaptive Linear Element*); *Multilayer Perceptron* (MLP) e Funções de Base Radial (RBF). A qualidade dos ajustes dos modelos de regressão, e das metodologias utilizando RNA e MVS, foram avaliadas utilizando-se o coeficiente de correlação entre os volumes individuais observados e estimados ($r_{y\hat{y}}$), a raiz quadrada do erro médio, expresso em porcentagem da média (RMSE%), análise gráfica dos resíduos (Res%). Considerando os resultados, MVS e RNA obtiveram desempenho ligeiramente melhores, comparados à metodologia tradicional, nas estimativas de volume individual, demonstrando serem técnicas que se adequaram bem para aplicações na área de mensuração e manejo florestal.

PALAVRAS CHAVE: análise de regressão, aprendizado de máquina, volumetria.

ABSTRACT

This study aimed to evaluate the performance of artificial neural networks (ANN) and support vector machines (SVM) in volumetric modeling in eucalyptus stands. Data from commercial plantations, located in four municipalities in the southern mesoregion of the state of Amapá, were used and were provided by a private company. Volumetric models established in the literature were adjusted and compared with the SVM and ANN techniques. Data were divided into 80% for training and 20% for model validation. The same dendrometric variables used by the regression models (DBH and height) were used by the SVM and ANN. For training and generalization of the SVM, four configurations were used, formed from two error functions and two *Kernel* functions. For configuration, training, and generalization of the ANN, the *NeuroForest-Volumetric* software was used, in which network configurations such as Adaline (*Adaptive Linear Element*) were used; Multilayer Perceptron (MLP) and Radial Base Functions (RBF). The quality of the adjustments of the regression models, and of the methodologies using ANN and SVM, were evaluated using the correlation coefficient between the observed and estimated individual volumes ($r_{y\hat{y}}$), the root mean square error, expressed as a percentage of the mean (RMSE%), and graphical analysis of residues (Res%). Considering the results, SVM and ANN performed slightly better, compared to the traditional methodology, in individual volume estimates, demonstrating that they are techniques that are well suited for applications in the area of measurement and forest management.

KEYWORDS: regression analysis, machine learning, volumetry.

INTRODUÇÃO

Os inventários florestais são um componente integral do monitoramento e gestão de recursos naturais (Fankhauser et al., 2018). Os inventários florestais são tradicionalmente realizados por meio de amostragem de campo intensiva, com o objetivo de fornecer aos gestores uma compreensão da composição e estrutura de uma floresta (Goodbody et al., 2017).

Estimar os volumes das árvores individuais é, na maioria das vezes, a principal finalidade dos levantamentos florestais, notadamente quando se trata de povoamentos destinados para fins comerciais (Machado e Figueiredo-Filho, 2009).

As estimativas volumétricas podem ser obtidas por meio de fatores de forma, quociente de forma, equações de volume, equações de múltiplos volumes ou de afilamento (Burkhart e Tome, 2012; Campos e Leite, 2017). A equação de volume é a forma mais usual de realizar a estimação do volume. Ela é uma expressão em que o volume da madeira é apresentado como função de outras grandezas ou variáveis da árvore (normalmente o diâmetro a altura do peito e a altura) que podem ser medidas e estimadas por meio não destrutivo (Batista et al., 2014). Porém nos últimos anos, as equações volumétricas têm sido utilizadas em estudos comparativos juntamente com aplicações de ferramentas da inteligência artificial, haja vista os bons resultados obtidos em alguns trabalhos na ciência florestal, dada à sua flexibilidade no treinamento e modelagem das relações entre variáveis, capacidade de aprendizado de informações de um conjunto de dados e a generalização desse aprendizado para dados desconhecidos (Binoti et al. 2016; Bonete et al., 2019; Abreu et al., 2020).

Avaliar novas abordagens de modelagem preditiva é uma importante ação na busca por modelos mais precisos e na superação de problemas comuns às técnicas convencionais. Os avanços nas técnicas computacionais, como algoritmos de aprendizado de máquina, têm sido cada vez mais utilizados para modelar dados biológicos. Esses algoritmos permitem o uso de dados categóricos, com ruídos estatísticos (outliers) e dados incompleto, portanto,

podem atender às necessidades sob diferentes cenários de conjunto de dados (Breiman, 2001). O aprendizado de máquina, de acordo com Bell (2015), é um ramo da inteligência artificial, no qual são efetuados sistemas que são induzidos a aprender por meio de dados a serem treinados, computacionalmente, o modelo vai aprendendo por meio de aperfeiçoamentos com a experiência no decorrer do tempo. Recentemente, uma nova cultura de modelagem estatística - aprendizado de máquina - ganhou impulso e foi aplicada para resolver questões desafiadoras em diversas áreas da ciência e tecnologia (Dalla-Corte, 2020).

As redes neurais artificiais (RNA) são ferramentas do aprendizado de máquina, que segundo Feltrin (2019) apresentam neurônios que estão ligados por meio de estrutura de entradas, intermediárias e saída. No qual esse neurônio se denomina *Perceptron*, que matematicamente efetua interpretação de dados de entradas, além dos pesos empregados sobre eles e as funções de ativação para o processamento final, basicamente se apresentando como um neurônio biológico, no qual as RNA são baseadas em sistemas biológicos com objetivo de execução de uma determinada tarefa (Haykin, 2007; Bulinaria, 2016).

Para Haykin (2007), as Máquinas de Vetores de Suporte (*Support Vector Machines*) compreendem um algoritmo de aprendizagem de máquina supervisionado, isto é, um algoritmo com capacidade de explorar um conjunto de dados de treinamento a fim de descobrir relações entre os atributos (variáveis independentes) e uma variável alvo (dependente).

OBJETIVOS

Com a necessidade de estimativas mais precisas, haja vista a potencialidade da aplicação das MVS e das RNA, o objetivo do presente estudo é aplicar essas metodologias na modelagem de volume individual com casca em povoamentos de eucalipto, efetuando comparação dos métodos com a modelagem tradicional.

MATERIAIS E MÉTODOS

Os dados desse estudo são de plantios comerciais não desbastados de *Eucalyptus urophylla* St Blake, com idades



variando entre 22 meses e 88 meses, localizados em quatro municípios (Ferreira Gomes, Itaubal, Porto Grande e Tartarugalzinho) da mesorregião sul do estado do Amapá. O clima equatorial úmido ou tropical super úmido, marcado por altas temperaturas, e elevados índices pluviométricos (média anual de 2500 mm), a temperatura média varia entre 36 °C a 20 °C. Em linhas gerais os solos contidos nas áreas operacionais são distróficos, altamente coesos, com baixa fertilidade. O relevo predominante nas áreas operacionais é o plano a suave ondulado, com altitudes variando de 30 m a 120 m.

Foi realizada a cubagem rigorosa de 214 árvores-amostra, onde seus respectivos diâmetros foram mensurados ao longo do fuste em posições fixas, nas seguintes medidas: 0 cm (base); 0,50 m; 1,0 m; 2,0 m e a partir desse ponto de 2,0 m em 2,0 m até o diâmetro mínimo de 4,0 cm, sendo o volume individual determinado pelo método de Smalian.

Foram ajustados os modelos de Schumacher e Hall (1933) na sua forma linearizada (1) e o modelo de Spurr (1952) (2).

$$\ln v = \beta_0 + \beta_1 \ln d + \beta_2 \ln h + \varepsilon_i \quad (1)$$

$$v = \beta_0 + \beta_1 d^2 h + \varepsilon_i \quad (2)$$

Em que:

v = Volume total com casca (m^3)

d = Diâmetro a 1,30 m de altura (cm)

h = Altura total (m)

β_0, β_1 = Coeficientes a serem estimados

ε_i = erro aleatório

As amplitudes mínimas, máximas e médias das variáveis dendrométricas utilizadas para estimar o volume são apresentadas na tabela 1.

Utilizou-se 80% dos dados para ajuste/treinamento ($n = 171$), e 20% para validação/generalização ($n = 43$) das estimativas de volumes individuais obtidos pelas metodologias avaliadas no presente trabalho. Tanto para MVS quanto para RNA, as variáveis de entrada foram o dia

e a altura, sendo o volume total com casca a variável de saída. Para o treinamento e generalização de todas as estimativas de volume individual com a técnica MVS, foram utilizadas quatro configurações, formadas a partir de duas funções de erro e duas funções de *Kernel*, sendo que as funções de erro foram: função do tipo I (3) e do tipo II (7).

Função do tipo I:

$$\text{Minimizar } \frac{1}{2} \cdot W^T W + C \cdot \sum_{i=1}^N \xi_i + C \cdot \sum_{i=1}^N \xi_i * \quad (3)$$

Sujeito as seguintes restrições:

$$W^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i \quad (4)$$

$$y_i - W^T \phi(x_i) - b \leq \varepsilon + \xi_i \quad (5)$$

$$\xi_i, \xi_i * \geq 0, i = 1, \dots, N, \varepsilon \geq 0 \quad (6)$$

Em que:

W = vetor de coeficientes

C = parâmetro de penalidade do erro

$\xi_i, \xi_i *$ variáveis de folga que caracterizam, respectivamente, o erro acima e abaixo do ε tubo

i = casos de treinamento

N = número total de casos de treinamento

$\phi(x_i)$ = kernel utilizado

b = bias

y_i = dados de saída

ε = erro máximo admitido

Função do tipo II:

$$\text{Minimizar } \frac{1}{2} \cdot W^T W - C \left(v \cdot \varepsilon + \frac{1}{N} \cdot \sum_{i=1}^N \xi_i + C \cdot \sum_{i=1}^N (\xi_i + \xi_i *) \right) \quad (7)$$

Sujeito as seguintes restrições:

$$(W^T \phi(x_i) + b) - y_i \leq \varepsilon + \xi_i \quad (8)$$

$$yi - (W^T \cdot \phi(x_i) + b) \leq \varepsilon + \xi_i \quad (9)$$

$$\xi_i, \xi_i * \geq 0, i = 1, \dots, N, \varepsilon \geq 0 \quad (10)$$

em que:

ν = parâmetro que regula o número de vetores de suporte

As quatro configurações de MVS utilizadas foram função de base radial (RBF) e linear, tipo I e tipo II, apresentadas matematicamente conforme tabela 2.

Para as redes neurais, utilizou-se configurações de redes do tipo Adaline (*Adaptive Linear Element*); *multilayer perceptron* (MLP) e funções de base radial (*radial basis function-RBF*), tendo como arquitetura genérica, a camada de entrada são as variáveis independentes utilizadas na regressão convencional, uma camada intermediária com n neurônios e uma camada de saída (volume).

Utilizou-se o algoritmo *Resilient Propagation* na sua variação *RPROP+* por ser um algoritmo eficiente em trabalhos recentes na área florestal (Tavares-Júnior et al., 2019; Freitas et al., 2020; Silva et al., 2020).

A definição do número ideal de neurônios da camada intermediária foi realizada conforme Bonete (2020), foi avaliada a função de ativação sigmoidal na camada oculta, combinando com a função sigmoidal e logarítmica na camada de saída, observando-se a ineficiência ou a não capacidade de ajuste da rede quando na presença de poucos neurônios na camada intermediária, bem como valores baixos da raiz quadrada do erro quadrático médio percentual (RMSE%) no treinamento e valor crescente da RMSE% na validação, quando na presença de muitos neurônios na camada intermediária. O número de neurônios na camada intermediária é responsável, principalmente, pela extração de características de não linearidade dos dados (Leite et al., 2016).

TABELA 1. Informações das variáveis dendrométricas para estimativa de volume individual

Variável	Ajuste/Treinamento			Validação/Generalização		
	Mín.	Máx.	Média	Mín.	Máx.	Média
Dap(cm)	6,3	23,7	13,4	5,8	21,5	13,8
Ht(m)	12	26,7	19,7	10	25,7	19,3
Volume (m³)	0,02	0,447	0,145	0,02	0,327	0,149

DAP = diâmetro à altura de 1,30 m; Ht = altura total; V = volume.

TABELA 2. Funções de *Kernel* testadas nas máquinas de vetor de suporte.

Tipo de Kernel	Funções	Parâmetros	Nº
Linear	$K(X_i - X_j)$	-	(11)
RBF	$K(X_i - X_j) = \exp(-\gamma \ X_i - X_j\)$	γ	(12)

Em que: $K(X_i - X_j) = \langle \phi(X_i) | \phi(X_j) \rangle$ e representa a função de Kernel aplicada aos dados de entrada; γ = parâmetro de forma; d = grau do polinômio; C = parâmetro de penalidade do erro.



Como critério de parada do treinamento das redes foi adotada a formatação padrão do *software* Neuro 4.0, em que a RNA interrompe o ajuste dos pesos ao atingir o erro médio de 0,0001 ou 3000 ciclos de treinamento.

A qualidade dos ajustes dos modelos de regressão e das estimativas foram avaliadas utilizando-se o coeficiente de correlação entre os volumes individuais observados e estimados ($r_{Y\hat{Y}}$), a raiz quadrada do erro médio, expresso em porcentagem da média ($RMSE\%$) e distribuição gráfica dos resíduos, demonstrados na tabela 3. E para verificar a significância dos parâmetros estimados dos modelos de

regressão, utilizou-se o teste t a 5% de probabilidade de erro.

O ajuste das regressões e o treinamento da máquina de vetor de suporte foram efetuados utilizando o *software* R (R Core Team, 2019), por meio do pacote e1071 (Meyer et al., 2019). E as redes neurais artificiais foram treinadas utilizando o *software* NeuroForest – *Volumetric*.

RESULTADOS E DISCUSSÃO

Na tabela 4 estão os resultados dos ajustes dos modelos de volume comercial com casca. Todos os modelos tiveram parâmetros significativos ($p < 0,05$) pelo teste t .

TABELA 3. Estatísticas utilizadas para avaliar o desempenho dos modelos de regressão, das MVS e das RNA

Estatísticas	Fórmulas	Nº
Coeficiente de correlação	$r_{Y\hat{Y}} = \frac{\sum_{i=1}^n y_i \hat{y}_i - \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n \hat{y}_i)}{n}}{\sqrt{\sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n}} \cdot \sqrt{\sum_{i=1}^n \hat{y}_i^2 - \frac{(\sum_{i=1}^n \hat{y}_i)^2}{n}}} \quad (13)$	
Raiz quadrada do erro quadrático médio (%)	$RMSE\% = \frac{100}{\bar{y}_i} \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (14)$	
Resíduo	$Res \% = \frac{(y_i - \hat{y}_i)}{y_i} \cdot 100 \quad (15)$	

Em que: y_i = valores observados; \hat{y}_i = valores estimados; \bar{y}_i = média dos valores da variável dependente e n é o número total de observações.

TABELA 1. Coeficientes estimados e estatísticas utilizadas para avaliar o desempenho do ajuste dos modelos volumétricos na estimação do volume individual.

Autor/Fonte	Modelo	Coeficientes	$r_{Y\hat{Y}}$	RMSE%	Nº
Schumacher e Hall (1933)	$lnv = \beta_0 + \beta_1 lnDAP + \beta_2 lnh$	$\beta_0 = 9,9085574$ $\beta_1 = 1,68596775$ $\beta_2 = 1,168174$	0,9917	7,92	4
Spurr (1952)	$v = \beta_0 + \beta_1 d^2 h$	$\beta_0 = 0,0170752$ $\beta_1 = 0,00003107$	0,9899	8,73	5

Em que: β s = coeficientes estimados; $r_{Y\hat{Y}}$ = coeficiente de correlação entre volumes observados e estimados; RMSE% = raiz do erro quadrático médio percentual.

Ao analisar o coeficiente de correlação dos dois modelos, pode-se dizer que os mesmos apresentaram resultados semelhantes. Quanto a estatística RMSE%, o modelo de Schumacher e Hall demonstrou maior precisão. Para Campos e Leite (2017), esse modelo tem sido o mais difundido por resultar quase sempre em estimativas não tendenciosas. Outras pesquisas na literatura florestal, também demonstram a superioridade do modelo de Schumacher e Hall, como Müller et al. (2014) que selecionaram modelos alométricos de *Eucalyptus* para estimativas volumétricas sendo que o modelo de Schumacher & Hall se destacou com estatísticas ligeiramente superiores aos demais; no trabalho de Hernandes Ramos et al. (2017) na quantificação de volume comercial de clones e sementes de *Eucalyptus urophylla* no sudeste do México, onde os melhores modelos foram o de Schumacher-Hall seguido do modelo de Spurr. No entanto, na modelagem da relação altura-diâmetro e o volume de mognos jovens africanos estabelecidos em sistemas agroflorestais sucessórios no nordeste do Brasil estudo de Santos et al. (2019), o modelo Spurr de dupla entrada foi mais preciso do que os outros 6 modelos avaliados.

Na figura 1, estão as distribuições dos erros percentuais para as estimativas de volumes individuais, para cada um dos modelos ajustados.

Nota-se uma certa semelhança entre os volumes estimados e observados na distribuição dos resíduos, no modelo de Spurr, uma tendência de superestimação para as árvores com diâmetros menores que 10 cm foi observada.

Na tabela 5 são demonstrados os resultados das estatísticas dos modelos volumétricos para os dados de validação.

Na etapa de validação, o modelo de Schumacher e Hall, obteve um maior coeficiente de correlação (r_{yy}) e menor raiz do erro quadrático médio (RMSE%). Pode-se observar na figura 1 que o modelo de Spurr manteve a tendência em superestimar o volume nos diâmetros

menores, sendo que alguns resíduos ultrapassaram 40%. Já o modelo de Schumacher e Hall, apresentou distribuição residual livre de tendenciosidades.

As estatísticas de avaliação da MVS são exibidas na tabela 6. As configurações que fizeram uso das funções *Kernel* linear, não se ajustaram aos dados, gerando estimativas volumétricas com valores negativos, provavelmente isso ocorreu em decorrência das funções de *Kernel* do tipo Linear não conseguirem extrair relações entre as variáveis de entrada (diâmetro e altura) e a variável de saída (volume). Uma característica importante que uma função de *Kernel* deve possuir é satisfazer condições do teorema de Mercer. De forma geral, um *kernel* satisfaz as condições do teorema de Mercer se der origem à matriz positiva semidefinida (Lorena e Carvalho, 2007). Mais detalhes sobre o teorema de Mercer podem ser encontrados em Burges (1998).

A MVS 4 obteve ligeiramente maior correlação e menor RMSE% que a MVS 2 e também que o modelo de regressão de Schumacher e Hall, que apresentou correlação de 0,9917 e RMSE% 7,92.

Resultados de maior exatidão utilizando a MVS para a estimativa do volume, superando o modelo de Schumacher e Hall, também foram encontrados por Binoti et al., (2016) e Cordeiro et al. (2015). Esses autores concluíram que a MVS pode ser utilizada para a predição volumétrica de espécies do gênero *Eucalyptus* e *Acacia mangium*, respectivamente. Em ambos os trabalhos, a MVS conseguiu maior exatidão nas estimativas, Vale ressaltar que no trabalho desses autores foi utilizada apenas uma configuração da MVS, composta pela função do tipo II e pela função de *Kernel* RBF.

No gráfico de distribuição dos resíduos, houve superestimação de volumes individuais da MVS 2, nas árvores com diâmetros menores que 10 cm. Tendo em vista esses resultados, a MVS 4 foi a que obteve os melhores resultados, conforme apresenta a figura 2.

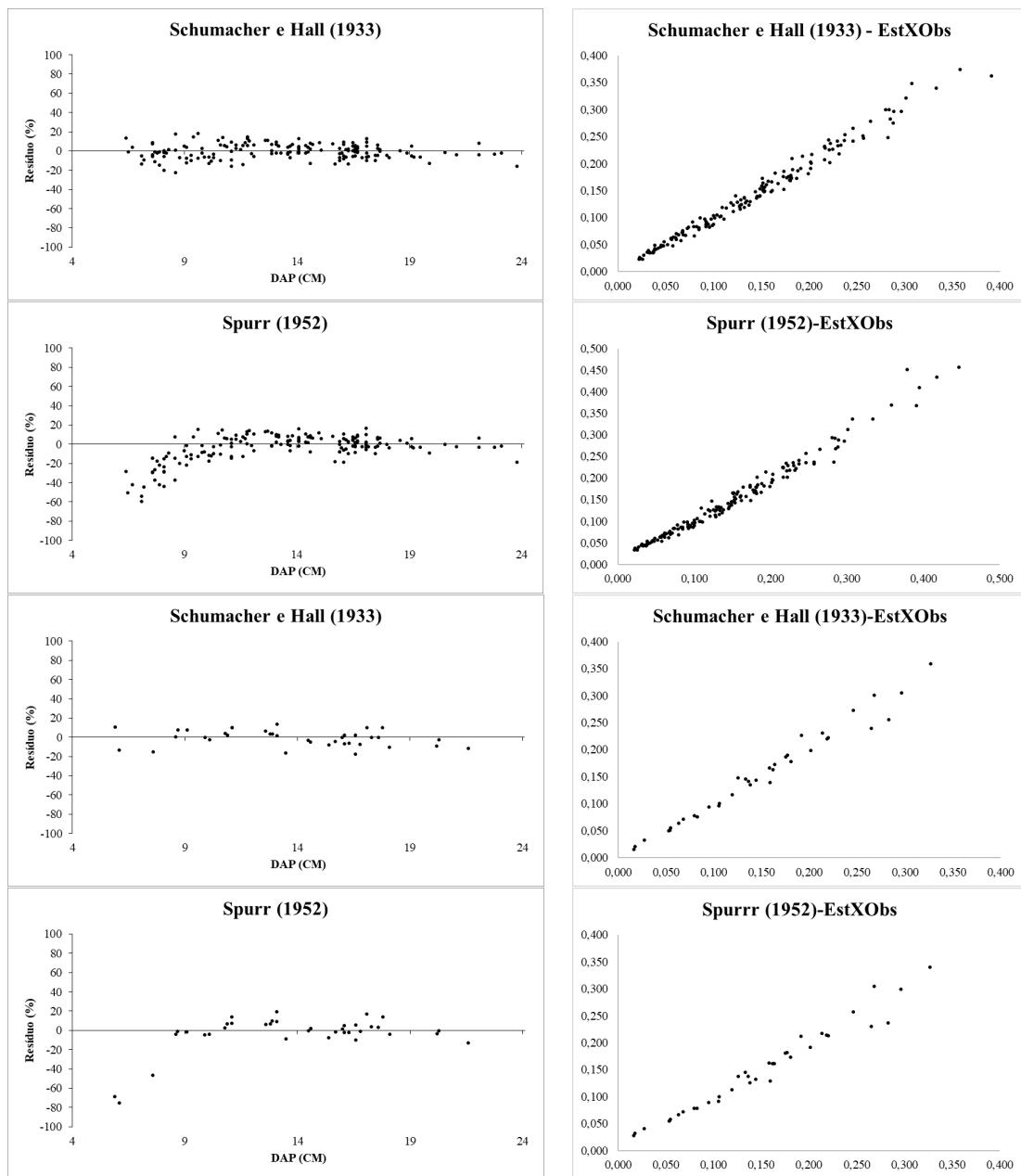


FIGURA 1. Distribuição de resíduos e volumes estimados e valores observados, na metodologia convencional, para os dados de ajuste e validação.

Nieto *et al.* (2016), visando estimar o volume sem casca de árvores de *Eucalyptus globulus*, confrontaram duas técnicas de inteligência artificial como alternativa aos modelos estatísticos convencionais de dupla entrada. Esses autores concluíram que a MVS foi a melhor técnica para estimar o volume sem casca para os dados em questão em relação às RNA e aos modelos de regressão. No trabalho desses

autores, foram testadas três funções de *Kernel* (Polinomial, sigmoidal e RBF), com melhores estimativas geradas pela função de *Kernel* RBF. É importante ressaltar que nesse estudo foram testadas quatro funções de *Kernel* (polinomial, sigmoidal, RBF e linear), constatou-se que a função de *Kernel* RBF, obteve as melhores estimativas.

TABELA 5. Estatísticas de validação para os modelos testados para estimar volume.

Autor/Fonte	Coeficientes	$ry\hat{y}$	RMSE%	Nº
Schumacher e Hall (1933)	$\beta_0 = -9,9085574$			
	$\beta_1 = 1,68596775$	0,9842	9,38	4
	$\beta_2 = 1,168174$			
Spurr (1952)	$\beta_0 = 0,0170752$			
	$\beta_1 = 0,00003107$	0,9827	9,81	5

Em que: Coeficientes de correlação ($ry\hat{y}$); raiz do erro quadrático médio percentual (RMSE%); ln = logaritmo natural; h= altura estimada; d= DAP = diâmetro à altura do peito (cm).

TABELA 6. Estatísticas utilizadas para avaliar o desempenho das MVS na estimação do volume, para os dados de treinamento.

MVS	Função	Kernel	$ry\hat{y}$	RMSE%
2	Tipo I	RBF	0,9940	6,72
4	Tipo II	RBF	0,9943	6,60

Em que: $ry\hat{y}$ = coeficiente de correlação entre os volumes observados e estimados; RMSE% = raiz do erro quadrático médio percentual.

Na etapa da validação, a MVS 4 também demonstrou estatísticas ligeiramente melhores que a MVS 2, as estatísticas dessa etapa da generalização são mostradas na tabela 7.

A MVS 4 do tipo II, obteve estimativas mais precisas na generalização indicando ser a melhor configuração de MVS para os dados desse estudo. A análise gráfica demonstra, menor dispersão de erros residuais, conforme figura 2.

Os resultados obtidos pela MVS 2 são similares aos encontrados por Abreu et al. (2020) ao avaliar alternativas para estimar volumes individuais em diferentes formações florestais no estado de Minas Gerais, relatou que mesmo apresentando bons resultados, a MVS, para alguns volumes pequenos, não conseguiu apresentar boas estimativas, comparando-se a outras metodologias avaliadas. Montaño (2016), ao modelar variáveis dendrométricas para plantios comerciais, também utilizou MVS para biomassa seca em florestas tropicais e nesse estudo, as MVS também foram inferiores a outras metodologias utilizadas no trabalho desse autor.

Os resultados estatísticos de avaliação do treinamento das três configurações das RNA, sendo elas Adaline (*Adaptive linear Element*), Perceptron de múltiplas camadas (*Multilayer Perceptron - MLP*) e Funções de Base Radial (*Radial Basis Function-RBF*), são demonstrados na tabela 8.

As redes do tipo Adaline e MLP (1 e 2 respectivamente) obtiveram estatísticas semelhantes, sendo que a rede 3 obteve maior coeficiente de correlação ($ry\hat{y}$) e menor RMSE% (raiz do erro quadrático médio percentual).

Analizando os gráficos de dispersão de resíduos, observou-se uma leve tendência em superestimar árvores com diâmetros menores, comportamento comum em uma série de dados biológicos, por se tratar de erro relativo. Isso demonstra a importância de se testar diferentes configurações de RNA e MVS, combinando as variantes existentes (função do tipo I e II) com diferentes funções de *Kernel*, buscando aquela que melhor se adeque a determinado problema e que obtenha maior exatidão na generalização.

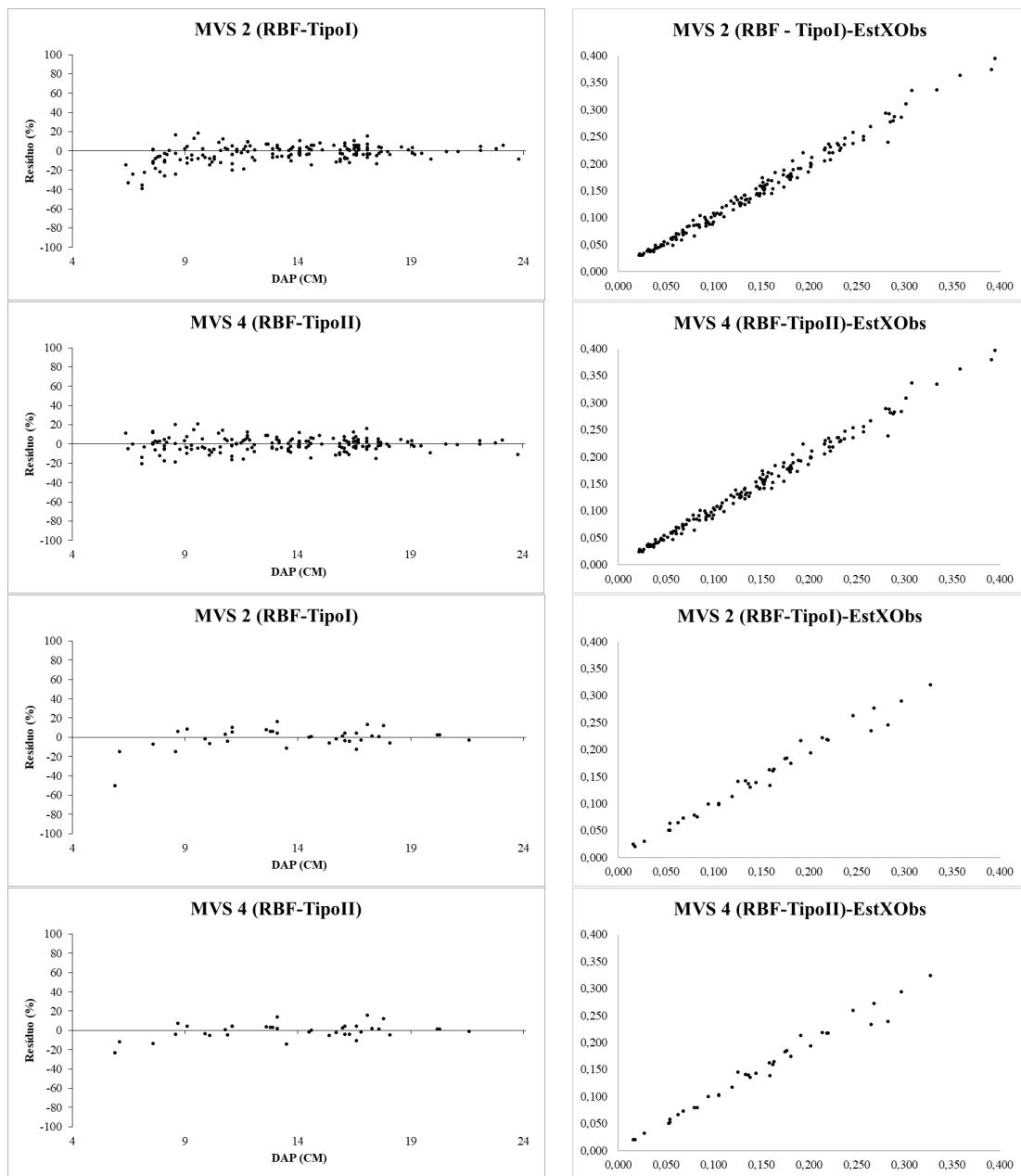


FIGURA 2. Distribuição de resíduos e valores estimados e valores observados, na metodologia de MVS, para os dados de treinamento e generalização.

TABELA 7. Estatísticas utilizadas para avaliar o desempenho da Máquina de Vetor de Suporte (MVS) na estimativa do volume individual, para os dados de generalização.

MVS	Função	Kernel	$ry\hat{y}$	RMSE%
2	Tipo I	RBF	0,9891	7,81
4	Tipo II	RBF	0,9895	7,64

Em que: $ry\hat{y}$ = coeficiente de correlação entre volumes observados e estimados; RMSE% = raiz do erro quadrático médio percentual.

TABELA 8. Estatísticas utilizadas para avaliar o desempenho das RNA na estimativa do volume de árvores, para os dados de treinamento.

RNA	Tipo*	FA			Arquitetura**	Treinamento	
		CO	CS			$r_{y\hat{y}}$	RMSE%
1	Adaline	S	S		2-1	0,9939	6,79
2	MLP	S	S		2-7-1	0,9940	6,74
3	RBF	S	LOG		2-2-1	0,9967	4,98

Em que: FA = Função de ativação; *Tipo de RNA: Adaline = Perceptron, MLP = Multilayer Perceptron, RBF = Radial Basis Function; CO = camada oculta; CS = camada de saída; S= Sigmoidal; LOG= Logarítmica; ** Número de neurônios em cada camada; $r_{y\hat{y}}$ = coeficiente de correlação entre volumes observados e estimados; RMSE% = raiz do erro quadrático médio percentual.

TABELA 9. Estatísticas utilizadas para avaliar o desempenho das RNA na estimativa de volumes individuais, para os dados de generalização.

RNA	Tipo*	FA			Arquitetura**	Generalização	
		CO	CS			$r_{y\hat{y}}$	RMSE%
1	Adaline	S	S		2-1	0,9884	8,05
2	MLP	S	S		2-7-1	0,9884	8,06
3	RBF	S	LOG		2-2-1	0,9960	4,71

Em que: FA = Função de ativação; *Tipo de RNA: Adaline = Perceptron, MLP = Multilayer Perceptron, RBF = Radial Basis Function; CO = camada oculta; CS = camada de saída; S= Sigmoidal; LOG= Logarítmica; ** Número de neurônios em cada camada; $r_{y\hat{y}}$ = coeficiente de correlação entre volumes observados e estimados; RMSE% = raiz do erro quadrático médio percentual.

Todas as configurações de rede obtiveram valores estatísticos superiores aos modelos de regressão conforme trabalhos de Azevedo et al. (2020), que utilizaram diferentes estratégias para modelar volumes de *Eucalyptus urophylla*, em Ribas do Rio Pardo, Mato Grosso do Sul, Brasil, onde foi treinada 1000 redes utilizando o layout *Multilayer Perceptron* (MLP) com algoritmo *Backpropagation*, em comparação também com o modelo de Schumacher e Hall, as RNA demonstraram ser mais adequadas do que os modelos de regressão na estimativa de vários volumes de eucaliptos, revelando maior precisão e praticidade; Abreu et al. (2017), que avaliou alternativas para estimar volumes individuais de uma floresta nativa semidecidual, na cidade de Viçosa, nos estados de Minas Gerais, Brasil, no referido trabalho, utilizou-se modelos lineares e lineares híbridos, 8 configurações de MVS com 4 funções *kernel* (Linear, Polinomial, Sigmoidal e RBF), RNA do tipo Adaline e RBF,

utilizando os algoritmos *Backpropagation* e *Simulated Annealing* respectivamente. MVS e RNA demonstraram resultados estatisticamente melhores comparados como o uso dos modelos de regressão, para aquela tipologia florestal.

Na tabela 9 estão indicados os resultados das estatísticas obtidas na fase de generalização para as três configurações de RNA que foram utilizadas nesse estudo, onde a rede do tipo RBF apresentou maior correlação e menor RMSE%. Bons resultados na modelagem de volumes individuais de árvores de *Eucalyptus spp.*, com a rede neural RBF, são encontrados nos trabalhos de Silva-Júnior et al. (2018) e Dantas et al. (2020).

Na figura 3 é demonstrado que graficamente na RNA 3, a distribuição dos resíduos está mais próxima ao eixo-x, que a redes Adaline e MLP nessa fase de generalização das redes.

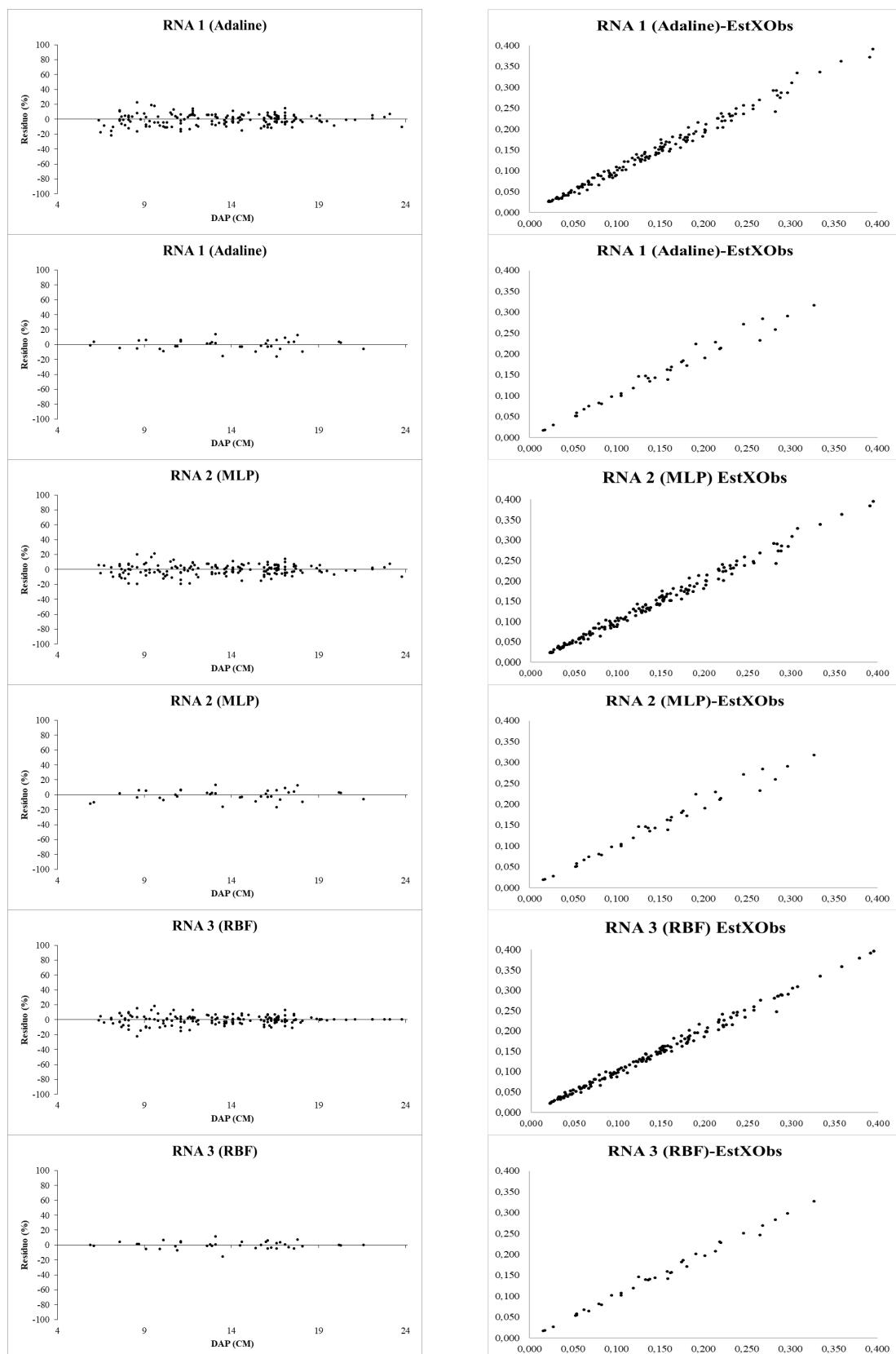


FIGURA 3. Distribuição de resíduos e valores estimados e valores observados, na metodologia de RNA, para os dados de treinamento e generalização.

CONCLUSÕES

Com o desenvolvimento desse estudo, verificou-se que as metodologias que utilizaram aprendizado de máquina, MVS e RNA demonstraram ser alternativas eficazes e estatisticamente satisfatórias à análise de regressão convencional, nas estimativas de volumes individuais para *Eucalyptus urophylla* St Blake, sendo que a metodologia que utilizava RNA, obtiveram melhores resultados.

Diante dos resultados, a eficiência e flexibilidade na aplicação das metodologias de MVS e RNA, apresentaram-se como boas alternativas em estudos de mensuração e manejo florestal.

REFERÊNCIAS

- Abreu, J. C. de, Soares, C. P. B., Leite, H. G., Binoti, D. H. B., & Silva, G. F. (2020). Alternatives to estimate the volume of individual trees in forest formations in the state of Minas Gerais-Brazil. *Cerne*, 26(3), 393-402. <https://doi.org/10.1590/01047760202026032728>
- Abreu, J. C. de, Soares, C. P. B., & Leite, H. G. (2018). Assessing alternatives to estimate the stem volume of a seasonal semi-deciduous forest. *Floresta*, 47(4), 375-382. <https://doi.org/10.5380/rf.v47i4.54259>
- Azevedo, G. B., Tomiazzi, H. V., Azevedo, G. T. D. O. S., Teodoro, L. P. R., Teodoro, P. E., de Souza, M. T. P., Batista, T. S., Eufrade-Junior, H. J., & Guerra, S. P. S. (2020). Multi-volume modeling of *Eucalyptus* trees using regression and artificial neural networks. *Plos One*, 15(9), e0238703. <https://doi.org/10.1371/journal.pone.0238703>
- Batista, J. L. F.: Couto, H. T. Z.; & Silva Filho, D. F. (2014). *Quantificação dos recursos florestais: árvores, arvoredos e florestas*. São Paulo, Oficina de Textos.
- Bell, J. (2015). *Machine learning: Hands-on for developers and technical professionals*. John Wiley & Sons.
- Binoti, D. H. B., Binoti, M. L. M. S., & Leite, H. G. (2014). Configuração de Redes Neurais Artificiais para estimação do volume de árvores. *Ciência da Madeira*, 5(1), 58-67. <https://doi.org/10.12953/2177-6830.v05n01a06>
- Binoti, D. H. B., Binoti, M. L. M. D. S., Leite, H. G., Andrade, A. V., Nogueira, G. S., Romarco, M. L., & Pitangui, C. G. (2016). Support vector machine to estimate volume of eucalypt trees. *Revista Árvore*, 40(4), 689-693. <https://doi.org/10.1590/0100-67622016000400012>
- Bonete, I. P., Arce, J. E., Filho, A. F., Retslaff, F. A. S., & Lansanova, L.R. (2019). Artificial neural networks and mixed-effects modeling to describe the stem profile of *Pinus taeda* L. *Floresta*, 50(1), 1123-1132. <https://doi.org/10.5380/rf.v50i1.61764>
- Bullinaria, J. A. (2016). *Introduction to neural computation: Apuntes de clase*. <http://www.cs.bham.ac.uk/~jxb/inc.html>. Acceso en: 20 enero 2020.
- Burkhart, H. E., & Tomé, M. (2012). *Modeling forest trees and stands*. Dordrecht: Springer.
- Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2), 121-167.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.
- Campos, J. C. C., & Leite, H. G. (2017). *Mensuração florestal: perguntas e respostas* (5^a ed.). Editora UFV.
- Cordeiro, M. A., de Jesus Pereira, N. N., Binoti, D. H. B., da Silva Binoti, M. L. M., & Leite, H. G. (2015). Estimativa do volume de *Acacia mangium* utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. *Pesquisa Florestal Brasileira*, 35(83), 255-261. <https://doi.org/10.4336/2015.pfb.35.83.596>
- Dalla-Corte, A. P., Souza, D. V., Rex, F. E., Sanquetta, C. R., Mohan, M., Silva, C. A., Zambrano, A. M. A., Prata, G., de Almeida, D. R. A., Trautenmüller, J. W., Klauberg, C., de Moraes, A., Sanquetta, M. N., Wilkinson, B., & Broadbent, E. N. (2020). Forest inventory with high-density UAV-Lidar: Machine learning approaches for predicting individual tree attributes. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2020.105815>
- Dantas, D., Calegario, N., Acerbi-Júnior, F. V., Carvalho, S. P. C., Isaac-Júnior, M. A., & Melo, E.A. (2020). Multilevel nonlinear mixed-effects model and machine learning for predicting the volume of *Eucalyptus spp.* trees. *Cerne*, 26(1), 48-57. <https://doi.org/10.1590/01047760202026012668>
- Freitas, E. C. S de., de Paiva, H. N., Neves, J. C. L., Marcatti, G. E., & Leite, H. G. (2020). Modeling of eucalyptus productivity with artificial neural networks. *Industrial Crops and Products*. <https://doi.org/10.1016/j.indcrop.2020.112149>
- Fankhauser, K., Strigul, N., & Gatziolis, D. (2018). Augmentation of traditional forest inventory and airborne laser scanning with unmanned aerial systems and photogrammetry for forest monitoring. *Remote Sensing*. <https://doi.org/10.3390/rs10101562>
- Feltrin, F. B. (2019). *Ciência de dados e aprendizado de máquina: Uma abordagem prática as redes neurais artificiais*. Kindle Edition.
- Nieto, P. G., García-Gonzalo, E., Arbat, G., Duran-Ros, M., de Cartagena, F. R., & Puig-Bargués, J. (2016). A new predictive model for the filtered volume and outlet parameters in micro-



- irrigation sand filters fed with effluents using the hybrid PSO-SVM-based approach. *Computers and Electronics in Agriculture*, 125, 74-80.
- Goodbody, T. R., Coops, N. C., Marshall, P. L., Tompalski, P., & Crawford, P. (2017). Unmanned aerial systems for precision forest inventory purposes: A review and case study. *The Forest Chronicle*, 93, 71-81. <https://doi.org/10.5558/tfc2017-012>
- Haykin, S. (2007). *Redes neurais: princípios e prática* (3a. ed.). Bookman.
- Hernández-Ramos, J., Santos-Posadas, H. M., Valdez-Lazalde, J. R., Tamarit-Urias, J. C., Ángeles-Pérez, G., Hernández-Ramos, A., Méndez-Lopez, B., & Peduzzi, A. (2017). *Eucalyptus urophylla* merchantable volume estimation with total volume and ratio models. *Agrociencia*, 51(5), 561-580.
- Leite, M. V. S., Silva, J. P. M., Cabacinha, C. D., & Assis, A. L. (2016). *Redes neurais artificiais para estimativa de altura dominante em povoamentos de Eucalipto*. IV SEEFLOR-BA, Vitória da Conquista-Bahia.
- Lorena, A. C., & Carvalho, C. P. L. F. (2007). Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, 14(2), 43-67.
- Machado, S. D., & Figueiredo Filho, A. (2009). *Dendrometria* (2a ed.). Editora Unicentro.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leish, F. (2019). Package ‘e1071’. *The R Journal*.
- Miguel, E. P., Rezende, A. V., Leal, F. A., Matricardi, E. A. T., Vale, A. T. D., & Pereira, R. S. (2015). Artificial neural networks for modeling wood volume and aboveground biomass of tall Cerrado using satellite data. *Pesquisa Agropecuária Brasileira*, 50(9), 829-839.
- Montaño, R. A. N. R. (2016). *Aplicação de técnicas de aprendizado de máquina na mensuração florestal* [Tese de doutorado, Universidade Federal do Paraná].
- Müller, M. D., Salles, T. T., Paciullo, D. S. C., Brighenti, A. M., & de Castro, C. R. T. (2014). Equações de altura, volume e afilamento para eucalipto e acácia estabelecidos em sistema silvipastoril. *Floresta*, 44(3), 473-484.
- Neuroforest (2017). <https://www.neuroforest.ucoz.com/>
- Santos, F. M., Terra, G., Chaer, G. M., & Monte, M. A. (2019). Modeling the height-diameter relationship and volume of young African mahoganies established in successional agroforestry systems in northeastern Brazil. *New Forests*, 50(3), 389-407.
- Schumacher, F. X., & Hall, F. S. (1933). Logarithmic expression of timber-tree volume. *Journal of Agricultural Research*, 47(9), 719-734, 1933.
- Silva, S., Oliveira Neto, S. N., Leite, H. G., Alcântara, A. E. M., Oliveira Neto, R. R., & Souza, G. S. A. (2020). Productivity estimate using regression and artificial neural networks in small familiar areas with agrosilvopastoral systems. *Agroforestry Systems*, 94(6), 2081-2097.
- Silva-Júnior, E. M., Maia, R. D., & Cabacinha, C. D. (2018). Bee-inspired RBF network for volume estimation of individual trees. *Computers and Electronics in Agriculture*, 152, 401-408.
- Spurr, S. H. (1952). *Forestry inventory*. Ronald Press.
- Tavares Júnior, I. S., Rocha, J. E. C., Chaves, A. S., Zanuncio, J. C., Farias, A. A., & Leite, H. G. (2019). Artificial neural networks and linear regression reduce sample intensity to predict the commercial volume of *Eucalyptus* clones. *Forests*, 10(3), 268.
- Team, R. C. (2019). *R: A language and environment for statistical computing*. <http://www.R-project.org/>

Manuscrito recebido em 25 de agosto de 2020

Aceito em 14 de outubro de 2021

Publicado em 3 de março de 2022

Este documento deve ser citado como:

Cordeiro, M. A., Arce, J. E., Guimarães, F. A., R., Bonete, I. P., Silva, A. V. S., Abreu, J. C. de, & Binoti, D., H. B. (2022). Estimativas volumétricas em povoamentos de eucalipto utilizando máquinas de vetores de suporte e redes neurais artificiais. *Madera y Bosques*, 28(1), e2812252. <https://doi.org/10.21829/myb.2022.2812252>



Madera y Bosques pelo Instituto de Ecología, A.C. é distribuída sob uma Licença Internacional Creative Commons Atribuição-NãoComercial ShareAlike 4.0.