

ESTIMATIVA DE DIÂMETRO A 1,30 M DO SOLO UTILIZANDO REDES NEURAI ARTIFICIAIS

ESTIMATIVE OF DIAMETER AT 1.30 M ABOVE GROUND USING ARTIFICIAL NEURAL NETWORKS

Emanuelly Canabrava Magalhães¹ Sthefany Mendes Zuba Thais Sales Gonçalves
Paulo Ricardo Santos Miranda Adriana Leandra de Assis Carlos Alberto Araújo Júnior

RESUMO

O objetivo deste estudo foi avaliar o desempenho das Redes Neurais Artificiais (RNAs) para estimativa da variável diâmetro a 1,30 m do solo (DAP), utilizando como variáveis independentes Idade, Área útil por planta e Altura das árvores. Foi realizado dois tipos de treinamento, para o primeiro os dados referentes ao último ano do inventário florestal não foram considerados, e para o segundo foi utilizado os dados de todos os anos sendo estes aleatorizados. Utilizou-se o *software* NeuroForest para treinamento das RNAs. Após o processamento foram calculadas as estatísticas BIAS, RQME, correlação e erro percentual médio; gerados gráficos de dispersão, considerando os valores estimados e observados; e histograma de resíduos. Observou-se que, o treinamento considerando os dados aleatorizados gerou as melhores estatísticas, com valores de correlação altos tanto para a fase de treinamento bem como para a validação, isto em decorrência da maior representatividade dos dados durante o treinamento. Os gráficos gerados para esse treinamento obtiveram menor dispersão dos pontos estimados, podendo inferir a maior precisão da estimativa quando comparado ao treinamento no qual os dados do último ano não foram considerados.

Palavras-chave: mensuração florestal; dendrometria; inteligência artificial.

ABSTRACT

The aim of this study was to evaluate the performance of Artificial Neural Networks (ANNs) to estimate the diameter variable at 1.30 m above ground (DAP), using as independent variables Age, Area per plant and Total height of trees. Two types of training were carried out; for the first, the data referring to the last year of the forest inventory were not considered, and for the second, the data of all the years were used and these were randomized. NeuroForest software was used to train ANNs. After the processing, the Bias, RQME, correlation and mean percent error statistics were calculated; dispersion charts were generated, considering the estimated and observed values; and histogram of residues. It was observed that training considering the randomized data generated the best statistics, with high correlation values both for the training phase as well as for the validation, due to the greater representativeness of the data during the training. The graphs generated for this training obtained less dispersion of the estimated points, and can infer the greater precision of the estimate when compared to the training in which the data of the last year were not considered.

Keywords: forest measurement; dendrometry; artificial intelligence.

INTRODUÇÃO

A mensuração de variáveis dendrométricas de uma floresta, em especial diâmetro e altura, demanda bastante atenção, visto que os dados coletados em campo serão utilizados para estimar a produtividade de um determinado povoamento florestal. A princípio, as medições eram realizadas manualmente com o auxílio de instrumentos como fita centimétrica e régua graduada, evoluindo para sutas eletrônicas e clinômetros digitais sendo que, atualmente, os dados são obtidos a partir de sensores remotos terrestres ou aerotransportados (CABOA et al., 2018; KOREN; MOKROŠ; BUCHA, 2017).

No entanto, ainda existem dificuldades em se coletar tais dados. A aferição da altura de árvores, por exemplo, é dependente da visualização das copas que sofre influência da presença de ventos fortes no local, espaçamentos muito adensados e terrenos de relevo acidentado, tornando-se um processo oneroso e impreciso agravado por erros na utilização dos equipamentos de medição (BINOTI et al., 2013).

Em função disso, são coletadas alturas de alguns indivíduos sendo a altura dos demais estimada a partir de modelos de regressão linear e não linear, comumente utilizados. Com o avanço tecnológico, têm-se

¹Estudante de graduação em Engenharia Florestal, Instituto de Ciências Agrárias, Universidade Federal de Minas Gerais, Av. Universitária, 1000 - Universitário, CEP 39404-547, Montes Claros (MG), Brasil. E-mail: emanuellymagalhaes1@gmail.com

desenvolvido ferramentas capazes de coletar tais dados de forma otimizada, dentre essas ferramentas se destaca o uso de laser para coleta de dados.

Os sistemas de varredura a laser (*Ligth Detection and Ranging* – LiDAR) podem ser utilizados para coleta de parâmetros dendrométricos do povoamento, sendo uma ferramenta de auxílio ao inventário florestal. Tal tecnologia apresenta alta precisão em seus sensores, o que possibilita a detecção da altitude do terreno e da copa das árvores para gerar informações de altura, porém tal ferramenta não permite a extração de informações diretas sobre os diâmetros das árvores, pois os pulsos são bloqueados pelas copas antes de atingirem a base do tronco. Logo, é necessário o uso de métodos para estimativa dessa variável a partir das informações de escaneamento (COSENZA, 2016).

Para as estimativas, busca-se modelos capazes de expressar o comportamento da variável de interesse, de forma que o valor estimado seja o mais próximo possível do valor observado em campo. Assim, a inteligência artificial é um recurso que atende a essa demanda através de treinamentos para determinado problema que possibilitam a estimativa do DAP, diminuindo a necessidade de medições excessivas em campo.

Diversos trabalhos vêm aplicando técnicas de inteligência artificial, como o uso de redes neurais artificiais para estimativa de altura (BINOTI et al., 2013) e diâmetro (LEITE et al., 2011). As Redes Neurais Artificiais (RNAs) são um sistema computacional constituído por várias unidades de processamento denominadas neurônios artificiais, os quais estão conectados entre si para executar uma tarefa específica (BINOTI, 2010). O funcionamento da rede consiste em um processo de aprendizagem e generalização, no qual são fornecidos dados de entrada que serão processados para gerar a saída desejada. Determinar a configuração da RNA é um processo importante, o qual depende do tipo de problema a ser resolvido e dos dados que serão apresentados, além de exigir testes para escolha daquela que permita um melhor desempenho de treinamento e velocidade de processamento, sendo a fase de treinamento das RNAs realizada através de algoritmos de aprendizagem, responsáveis por ajustar os pesos das conexões entre as unidades de processamento (BINOTI, 2010).

Para este estudo foi aplicado o algoritmo de treinamento Resilient Propagation (RPROP), caracterizado por realizar uma adaptação local das atualizações de peso de acordo com o comportamento da função de erro, tal processo não é influenciado pelo tamanho da derivada e sim pelo seu sinal, o que lhe diferencia das demais técnicas adaptativas (CHEN; LIN, 2011; PRASAD; SINGH; LAL, 2013). Assim, o objetivo deste trabalho foi estimar o diâmetro a 1,30 m de árvores do gênero *Eucalyptus sp.* utilizando redes neurais artificiais (RNAs).

MATERIAIS E MÉTODOS

Os dados utilizados neste estudo são oriundos de inventários florestais contínuos realizados em plantios florestais na região do norte do estado de Minas Gerais, durante o período de 2013 a 2016, totalizando 12.131 pares de dados de altura (h) e DAP das árvores medidas em diferentes idades e áreas úteis.

O banco de dados foi dividido em dois conjuntos, sendo que para cada um foram selecionados dados diferentes para treinamento das RNAs e validação do modelo. No conjunto 1 os dados para treinamento foram referentes aos anos de 2013, 2014 e 2015, e para validação foram utilizados os dados do ano de 2016. Já para o conjunto 2, a divisão dos dados foi feita de forma aleatória.

Para estimativa dos pesos das redes neurais artificiais, utilizou-se o software *NeuroForest versão 4.0.2* considerando o algoritmo Resilient Backpropagation, 8 neurônios na camada intermediária, funções de ativação do tipo sigmoideal para neurônios das camadas intermediária e de saída e interrupção do treinamento após 500 ciclos, sendo treinadas 30 RNAs para cada caso e selecionadas as melhores para análise. Tal parametrização foi definida a partir de testes realizados previamente.

Foi avaliado um modelo para estimativa do DAP utilizando como variáveis independentes Idade (Id), Área útil por planta (Ap) e Altura (h) para cada um dos conjuntos. Gráficos de dispersão considerando os valores estimados e observados foram construídos para avaliação dos resultados, bem como histogramas de resíduos. Os valores de erro para cada observação (e_i) foram obtidos utilizando-se a equação 1. Também foram calculadas as estatísticas Bias (equação 2), raiz quadrada do erro médio quadrático (equação 3), correlação entre valores estimados e observados (equação 4) e erro médio percentual (equação 5).

$$e_i = \hat{y}_i - y_i \quad (\text{equação 1})$$

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) \quad (\text{equação 2})$$

$$RQME (\%) = \frac{100}{\bar{y}} \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (\text{equação 3})$$

$$r = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (\text{equação 4})$$

$$EMP = \frac{1}{n} \sum_{i=1}^n \left[100 \cdot \frac{(y_i - \hat{y}_i)}{y_i} \right] \quad (\text{equação 5})$$

Em que: \hat{y}_i = valor estimado; y_i = valor observado para o i-ésimo dado; n = número total de observações; \bar{y} = média dos valores observados; $\bar{\hat{y}}$ = média dos valores estimados.

RESULTADOS E DISCUSSÃO

Analisando as estatísticas calculadas para cada conjunto (Tabela 1), é perceptível que o Conjunto 1 apresentou valor de correlação para validação, principalmente, inferior aos valores referentes ao Conjunto 2, o que pode ser explicado pelo não uso dos dados do último ano do inventário florestal no Conjunto 1 durante o treinamento. Nesse caso, os dados utilizados para validação possuem variação de valores concentrados nos maiores diâmetros, acima de 15 cm (Figura 1), ao passo que a maior concentração de dados de treinamento está abaixo de tal valor.

Já para os resultados do treinamento nota-se que os valores das estatísticas calculadas foram bem semelhantes entre os dois conjuntos. Isso significa que as RNAs são capazes de mapear os relacionamentos existentes entre a variável dependente (DAP) e as variáveis independentes (h, Ap e Id), independentemente do banco de dados considerado.

TABELA 1: Estatísticas calculadas para a modelagem adotada nas RNAs referente a estimativa do diâmetro a 1,30 m do solo, para cada conjunto avaliado.

TABLE 1: Calculated statistics for the modeling adopted in ANNs for the estimation of the diameter at 1.30 m above ground, for each evaluated set.

Conjunto	Etapa	Bias	RQME %	r	EMP
1	Treinamento	-		0,9	
	Validação	0,01	11,72	2	1,74
2	Treinamento	0,47	11,83	3	4,67
		0,01	11,65	2	1,70
	Validação			0,9	
		0,23	11,95	1	3,06

Em que: Bias = desvio médio dos valores estimados em relação aos observados; RQME% = raiz quadrada do erro médio quadrático; r = correlação entre valores estimados e observados; EMP = erro médio percentual.

Os resultados discutidos anteriormente são evidenciados e complementados pelos gráficos de dispersão e histograma de resíduos para a estimativa do diâmetro a 1,30 m do solo (Figura 1). Para o Conjunto 1, observa-se uma maior amplitude para os valores de erro na etapa de validação da rede neural treinada, com erros variando entre as classes de -25% e 55%, sendo que para o treinamento os valores variam entre -25% e 45%. Padrão semelhante ocorreu para o Conjunto 2, exceto pela menor amplitude para os erros de validação.

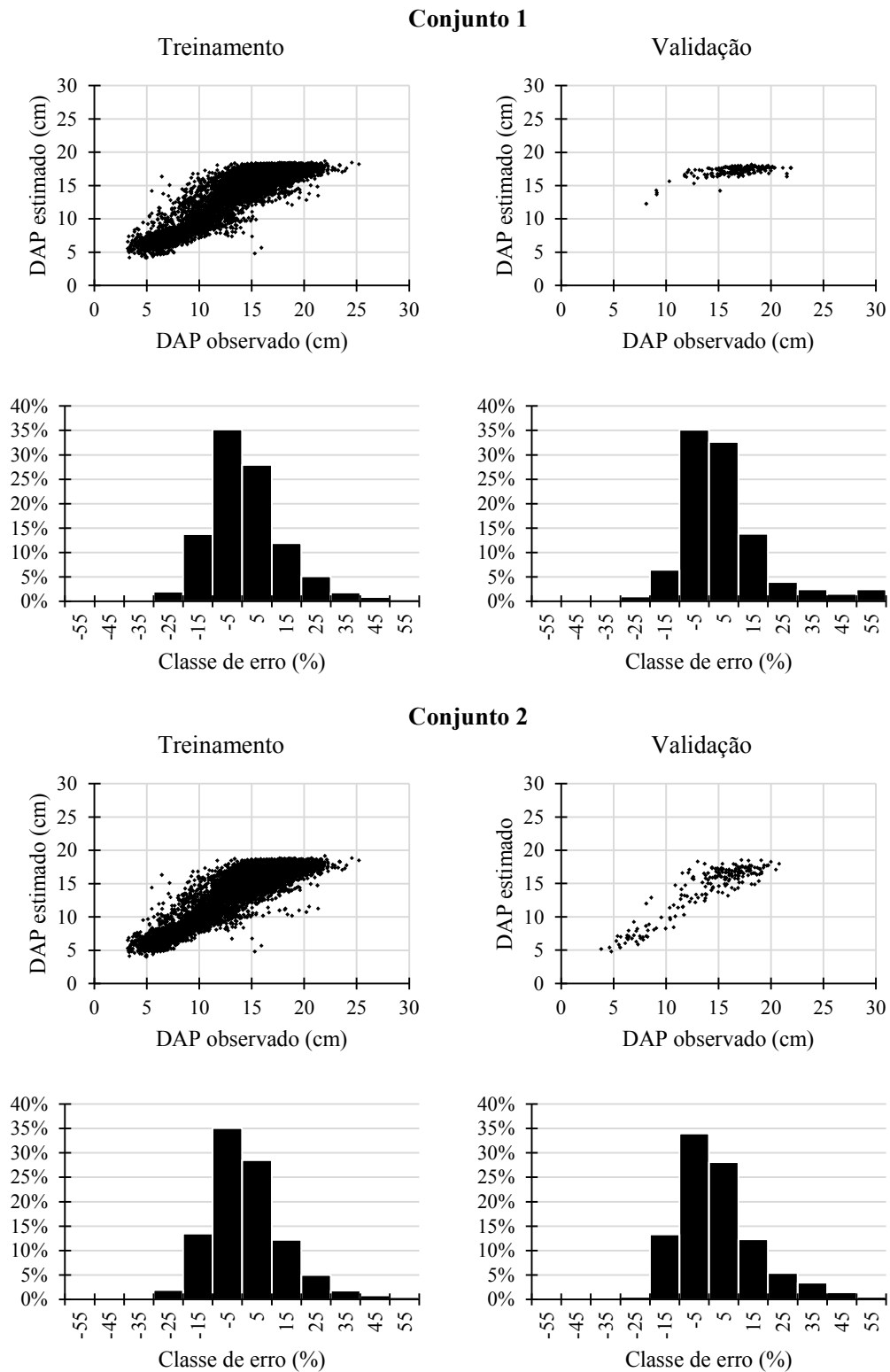


FIGURA 1: Gráficos de dispersão considerando os valores estimados e observados e histogramas de resíduos, para cada conjunto avaliado.

FIGURE 1: Scatter plots considering the estimated and observed values and histograms of residues, for each evaluated set.

Os resultados demonstram que, mesmo que as RNAs tenham apresentado bons resultados para o treinamento, o conjunto de dados para o qual a mesma será aplicada deverá ter sido representado durante a fase anterior, fazendo com que o aprendizado das RNAs seja de fato eficaz. Para que a generalização seja a mais ampla possível, é necessário que os dados de treinamento sejam o mais representativo do problema, possuindo grande variabilidade em termos de exemplos da relação existente entre a variável dependente e as variáveis independentes. Caso contrário, haverá sempre o risco de que haja *overfitting*, como abordado no estudo de Choudhury, Hosseinzadeh e Berndt (2012), o qual discute a capacidade de generalização das RNAs para previsão de determinada variável.

Assim, o maior número de dados representativos para treinamento possibilita o aumento da variabilidade de informações fornecidas a RNA melhorando a acurácia da estimativa (PEREIRA e CENTENO, 2017). O conjunto 2 obteve melhores resultados, tanto para treinamento como para validação. Em ambos, houve menor dispersão, sendo que para o treinamento os pontos se mostraram mais concentrados e erros variando entre as classes de -25% e 45%, enquanto para a validação os erros variaram entre -15% e 45%.

CONCLUSÕES

As redes neurais artificiais se mostraram uma ferramenta eficaz para estimativa do variável diâmetro, desde que os dados fornecidos a mesma sejam representativos do problema abordado para que o treinamento seja eficiente de tal forma que as RNAs possam ser aplicadas a outros dados.

AGRADECIMENTOS

Os autores expressam seus agradecimentos à FAPEMIG e ao CNPq pelo auxílio financeiro e à Universidade Federal de Minas Gerais pelo suporte científico.

REFERÊNCIAS BIBLIOGRÁFICAS

- BINOTI, D. H. B. et al. Redução dos custos em inventário de povoamentos equiâneos. *Agrária*, [s.l.], v. 8, n. 1, p. 125-129, 2013.
- BINOTI, M.L.M.S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. 2010. 54 f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa. Viçosa, 2010.
- CABOA, C. et al. Automatic dendrometry: Tree detection, tree height and diameter estimation using terrestrial laser scanning. *International Journal of Applied Earth Observation and Geoinformation*, [s.l.], v. 69, n. 1, p. 164-174, 2018.
- CHEN, C.-S.; LIN, J.-M. Applying Rprop Neural Network for the Prediction of the Mobile Station Location. *Sensors*, [s.l.], v. 11, n. 4, p. 4207-4230, 2011.
- CHOUHURY, T. A.; HOSSEINZADEH, N.; BERNDT, C. C. Improving the Generalization Ability of an Artificial Neural Network in Predicting In-Flight Particle Characteristics of an Atmospheric Plasma Spray Process. *Journal Of Thermal Spray Technology*, [s.l.], v. 21, n. 5, p.935-949, mar. 2012.
- COSENZA, D. N. **Aplicação de escaneamento a laser aerotransportado no inventário a nível de árvores individuais de povoamentos de eucalipto**. 2016. 45 f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa. Viçosa, 2016.
- KOREŇ, M.; MOKROŠ, M.; BUCHA, T. Accuracy of tree diameter estimation from terrestrial laser scanning by circle-fitting methods. *International Journal of Applied Earth Observation and Geoinformation*, [s.l.], v. 63, n. 1, p. 122-128, 2017.
- LEITE, H. G. et al. Estimation of inside-bark diameter and heart wood diameter for *Tectona grandis* Linn. trees using artificial neural networks. *European Journal of Forest Research*, [s.l.], v. 130, n. 2, p. 263-269, 2011.
- PEREIRA, G. H. A.; CENTENO, J. A. S. Avaliação do tamanho de amostras de treinamento para redes neurais artificiais na classificação supervisionada de imagens utilizando dados espectrais e laser scanner. *Boletim de Ciências Geodésicas*, [s.l.], v. 23, n. 2, p.268-283, jun. 2017.
- PRASAD, N.; SINGH, R.; LAL, S. P. Comparison of back propagation and resilient Propagation algorithm for spam classification. *Fifth International Conference On Computational Intelligence, Modelling And Simulation*, p. 29-34, 2013.