

ESTIMATIVA DA ALTURA DE ÁRVORES POR MEIO DE EQUAÇÕES HIPSOMÉTRICAS E POR REDES NEURAIS ARTIFICIAIS

Anderson Lopes Barreiros¹; Christiene Karine Ferreira²; Márcio Leles Romarco de Oliveira³;

1 Anderson Lopes Barreiros, Bolsista (CNPq), Engenharia Florestal, UFVJM Campus JK, Diamantina - MG; andersonlb16@gmail.com

2 Christiene Karine Ferreira, Engenharia Florestal, UFVJM, Diamantina – MG

3 Márcio Leles Romarco de Oliveira: Pesquisador da UFVJM, Campus JK; marciolormarco@ufvjm.edu.br

RESUMO

Objetivou-se com esse estudo, estimar a altura de árvores em povoamentos de eucalipto por meio de modelos de regressão e por redes neurais artificiais (RNA). Foram utilizados dados de inventário florestal de clones comerciais de eucalipto com idade entre 5 e 10 anos, com diferentes materiais genéticos e espaçamentos, provenientes de uma empresa florestal localizada na região central do estado de Minas Gerais. Três modelos de regressão foram avaliados, diferindo em relação à inserção de variáveis independentes ao modelo. Para a análise de regressão avaliou-se, portanto, a estimativa da altura em função apenas do DAP (diâmetros com casca a 1,30m do solo), em função do DAP e do diâmetro dominante e em função do DAP e da altura dominante. Dessa forma, para os modelos de regressão também foram verificadas diferenças entre a estimativa da altura por meio do ajuste em nível de talhão e a estimativa em nível de parcela. Para fins de comparação, duas combinações de RNA foram testadas, uma considerando a inserção e outra a ausência da altura dominante dentre as variáveis de entrada do ajuste. Para verificar a qualidade do ajuste das metodologias avaliadas utilizou-se as estatísticas correlação de Pearson, coeficiente de determinação ajustado, erro padrão da estimativa, bias, e raiz quadrada do erro médio, bem como análise gráfica de resíduos. Concluiu-se que o ajuste em nível de parcela gerou resultados imprecisos por falta de quantidade de dados suficientes para ajuste de regressão, e que, em alguns casos, o diâmetro dominante não é capaz de explicar as variações na altura total. Verificou-se que, tanto as redes neurais quanto os modelos de regressão se mostraram eficientes na estimativa da altura ao utilizar a altura dominante, pois apresentaram resultados semelhantes. Dessa forma, não houve diferença estatística significativa entre a estimativa da altura por meio de redes neurais e a estimativa da altura por modelos de regressão.

INTRODUÇÃO:

O inventário de florestas plantadas na maioria das vezes tem como objetivo estimar a produção volumétrica. Sua estrutura compreende fundamentalmente a determinação de um método de amostragem, a alocação de unidades amostrais na área, a medição de variáveis dentro da unidade amostral e a obtenção de variáveis de interesse. Dentre as variáveis computadas destacam-se as medições de diâmetro e altura em nível de árvores, fundamentais para a exatidão das estimativas de volumes da unidade amostral de inventário florestal.

A altura é uma variável importante no inventário florestal de plantios comerciais, por ser aplicada na estimativa de volume, na identificação do estágio de desenvolvimento de uma floresta e na classificação dos locais quanto a sua produtividade (SANQUETTA et al., 2009).

A medição da altura é uma etapa onerosa e sujeita a erros, tanto pela natureza do aparelho utilizado quanto pela condição do povoamento. Isso porque, a variável altura, na maioria das situações requer um maior tempo para ser mensurada, principalmente em espaçamentos mais adensados ou em floresta com árvores de grande porte (SCOLFORO, 1993). Já a avaliação do diâmetro em campo é relativamente simples, pois a medição é de fácil acesso e as medidas são mais confiáveis (SOARES et al., 2007).

Para superar esse gargalo da medição da variável altura em inventário de florestas plantadas, comumente mede-se a altura de apenas algumas árvores da unidade amostral e utiliza-se métodos de estimação para estimar a altura das árvores não medidas em campo. Esses métodos são baseados na relação biológica existente entre as variáveis diâmetro e altura (relação hipsométrica).

Normalmente a relação hipsométrica pode ser definida por meio de modelos de regressão que é o método tradicionalmente utilizado. E recentemente, tem ainda a utilização de técnicas de inteligência artificial, como as redes neurais artificiais (RNA), que têm sido cada vez mais empregadas na resolução de problemas complexos (HAYKIN, 2001).

METODOLOGIA:

Dados utilizados: Os dados utilizados para realização deste estudo foram obtidos a partir de inventário pré-corte realizado em 2015, oriundo de plantios de eucalipto (*Eucalyptus* sp.). O povoamento florestal, conduzido em regime de alto fuste, é pertencente a uma empresa florestal da região central de Minas Gerais.

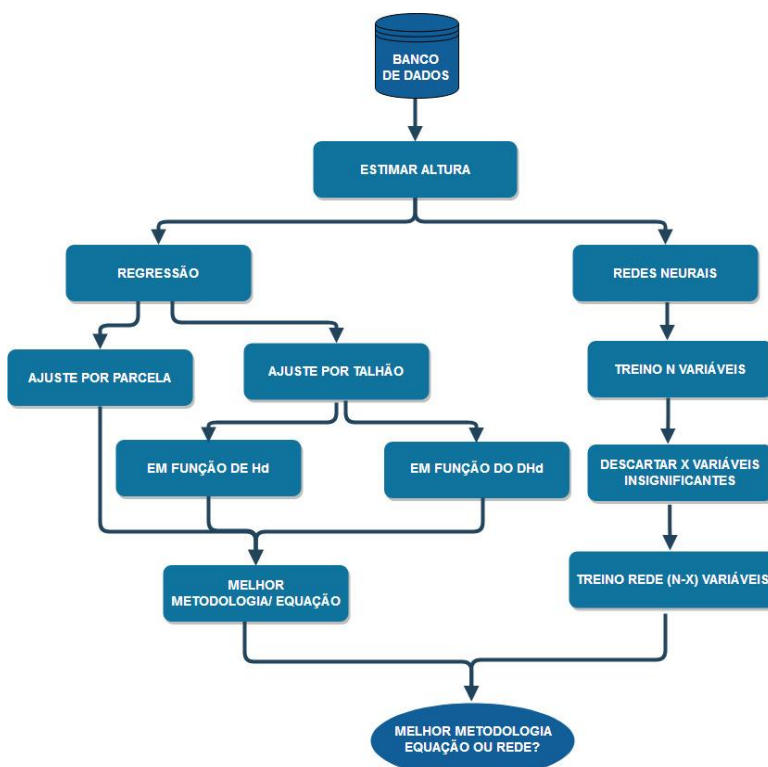
O povoamento possui oito materiais genéticos, em dois diferentes ciclos, distribuídos em 14 espaçamentos distintos. A área total plantada é de 1250,75 hectares, com idade entre cinco e dez anos.

Foram utilizados dados de 828 parcelas com área de, em média, 216,06 m² cada. Essas parcelas foram distribuídas aleatoriamente em 64 talhões.

Para cada parcela foram medidos os diâmetros com casca a 1,30m do solo (DAP) de todas as árvores. Foram mensurados também a altura de, em média, nove árvores normais e duas árvores dominantes. As variáveis altura dominante (Hd) e diâmetro dominante (D_{Hd}) foram obtidas do mesmo indivíduo. Sendo assim, os diâmetros com casca, mensurados a 1,30m do solo, das árvores dominantes foram considerados diâmetros dominantes (D_{Hd}).

Metodologias de ajuste avaliadas para estimativa de altura: Para avaliar a estimativa de altura foram utilizadas quatro diferentes alternativas. A primeira alternativa avaliou o ajuste de equação hipsométrica em nível de parcela. O ajuste de equação hipsométrica em nível de talhão, utilizando a altura dominante, foi a segunda alternativa. Para a terceira alternativa, utilizou-se também a equação hipsométrica em nível de talhão, porém, nesse caso utilizou-se o diâmetro dominante. Já a quarta alternativa considerou a estimativa da altura por meio de redes neurais. Estas etapas são apresentadas na Figura 1.

Figura 1 – Metodologias de ajuste avaliadas para estimativa da altura das árvores que não tiveram a altura total medidas em campo.



Fonte: O AUTOR, 2019.

Em que: N = total de variáveis utilizadas nesse estudo; X = número qualquer de variáveis insignificantes nesse estudo; Hd = altura dominante, em metros; D_{Hd} = diâmetro das árvores dominantes, em cm.

Modelos hipsométricos avaliados: Com a finalidade de estimar a altura das árvores que não tiveram a altura total medida em campo foram testados os modelos apresentados na Tabela 1. Trata-se de

adaptações de um modelo hipsométrico (modelo de Curtis) comumente empregado pelas empresas florestais, no Brasil, para estimação da altura em plantios comerciais. O uso desta adaptação buscou avaliar qual das variáveis utilizadas possui maior influência na estimativa da altura.

Tabela 1 – Modelos hipsométricos avaliados neste estudo.

Nº	Modelo
(1)	$\text{LnHT} = \beta_0 + \beta_1 \frac{1}{\text{DAP}} + \varepsilon$
(2)	$\text{LnHT} = \beta_0 + \beta_1 \frac{1}{\text{DAP}} + \beta_2 \ln \text{Hd} + \varepsilon$
(3)	$\text{LnHT} = \beta_0 + \beta_1 \frac{1}{\text{DAP}} + \beta_2 \ln D_{\text{Hd}} + \varepsilon$

Fonte: O AUTOR, 2019.

Em que: Ln = logaritmo neperiano; HT = altura total ou comercial, em metros; DAP = diâmetro com casca medido a 1,30m do solo, em cm; Hd = altura dominante, em metros; D_{Hd} = diâmetro das árvores dominantes, em cm; β_i = parâmetros de regressão ($i = 0, 1, 2$); e ε = erro aleatório.

O ajuste dos modelos foi realizado no software R Core Team (2018), versão 3.5.2, pela plataforma RStudio. O modelo (1) foi ajustado em nível de parcela resultando em 885 equações; e os modelos (2) e (3) foram ajustados em nível de talhão, resultando em 64 equações cada. Para o ajuste dos modelos foram utilizadas todas as árvores que tiveram as alturas medidas em campo.

Foram utilizadas sete variáveis de entrada, sendo elas as variáveis ciclo, material genético, espaçamento, DAP (cm), idade (meses), altura dominante (m) e categoria em relação à altura da árvore (bifurcada acima de 1,30m, dominante, normal, ponta seca, e, as demais árvores não classificadas nestas categorias foram consideradas de outras). Como se desejou estimar a altura, a variável de saída foi a altura total.

Para treinamento utilizou-se duas combinações de RNA. A primeira combinação considerou a inclusão de todas as sete variáveis de entrada acima mencionadas. Já a segunda combinação avaliou a ausência da variável altura dominante considerando apenas a combinação entre as seis demais variáveis de entrada.

RESULTADOS E DISCUSSÕES:

Coefficiente de determinação ajustado (\bar{R}^2) e erro padrão da estimativa (S_{yx}): Conforme descrito na Tabela 5, o modelo (1) apresentou tanto parcelas com a melhor precisão da estimativa (menor valor de S_{yx}), quanto parcelas com a pior precisão da estimativa (maior valor de S_{yx}). Da mesma forma, este modelo apresentou o maior valor de coeficiente de determinação ajustado (0,9935), porém também apresentou valores negativos para esse mesmo critério estatístico (8,76% dos valores do ajuste).

Tabela 5 – Valores mínimos, médios e máximos de coeficiente de determinação (R^2) e erro padrão da estimativa (S_{yx}), resultantes de cada um dos modelos hipsométricos ajustados.

	$R^2_{\text{mín}}$	$R^2_{\text{méd}}$	$R^2_{\text{máx}}$	$S_{yx_{\text{mín}}}$	$S_{yx_{\text{méd}}}$	$S_{yx_{\text{máx}}}$
Modelo 1	-0,1662	0,5023	0,9935	0,0042	0,0253	0,1272
Modelo 2	0,1927	0,5842	0,9010	0,0198	0,0328	0,0643
Modelo 3	0,0798	0,4882	0,8432	0,0208	0,0368	0,0734

Fonte: O AUTOR, 2019.

O modelo (2) apresentou melhor desempenho, por meio dos valores de coeficientes de determinação ajustado maiores que o modelo (3). Indicando que as variáveis independentes do modelo (2) explicam mais a variação na altura, do que as variáveis independentes do modelo (3).

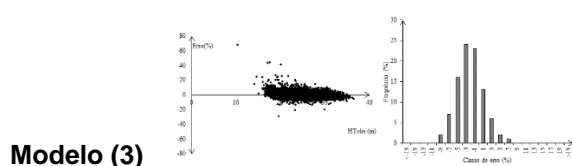
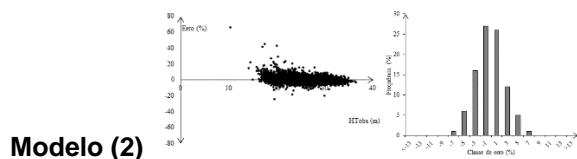
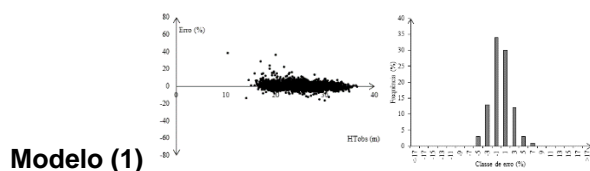
O mesmo ocorreu para a precisão da estimativa, em que o modelo (2) apresentou melhor precisão da estimativa (menor valor de S_{yx}), em relação ao modelo (3).

Comportamento gráfico dos resíduos

A Figura 5 apresenta o comportamento da distribuição gráfica dos resíduos, mostrando que todos os três modelos apresentam valores distribuídos adequadamente. Verificou-se resíduos positivos e negativos balanceados e inexistência de tendências ou heterocedasticidade.

Por intermédio dos histogramas de frequência dos erros pôde-se confirmar que todos os três modelos apresentaram resultados adequados de distribuição dos resíduos.

Figura 5 – Distribuição dos resíduos em função da altura observada e histograma do erro percentual para cada um dos modelos ajustados.



Fonte: O AUTOR, 2019.

Qualidade do ajuste por redes neurais: Conforme as estatísticas da amostra de validação descritas na Tabela 9 foi possível inferir que, de modo geral, as RNA que consideram a combinação entre todas as 7 variáveis de entrada (RNA 1, RNA 2, RNA 3, RNA 4 e RNA 5) apresentam tendência a superestimativas da altura, pois resultaram em valores negativos de bias (variando de -0,0029 a -0,0097).

Já as RNA que consideram a retirada da variável de entrada altura dominante (RNA 6, RNA 7, RNA 8, RNA 9 e RNA 10), tendem a subestimativas da altura, pois resultaram em valores positivos de bias (variando de 0,0148 a 0,0181).

No entanto, as RNA que consideraram a inserção da variável altura dominante apresentaram valores de bias mais próximos de zero. Logo, a inclusão da variável altura dominante demonstrou menor viés nas estimativas.

Em que: RQEM = raiz quadrada do erro médio; SQR = soma dos quadrados dos resíduos; r_{HT} = coeficiente de correlação de Pearson entre os valores observados e estimados; HT = altura total, em metros.

Analisando a RQEM verificou-se que as combinações que consideram a inserção da variável altura dominante apresentam melhor precisão da estimativa, pois obtiveram menores valores para a RQEM (variando de 0,8852 a 0,9166).

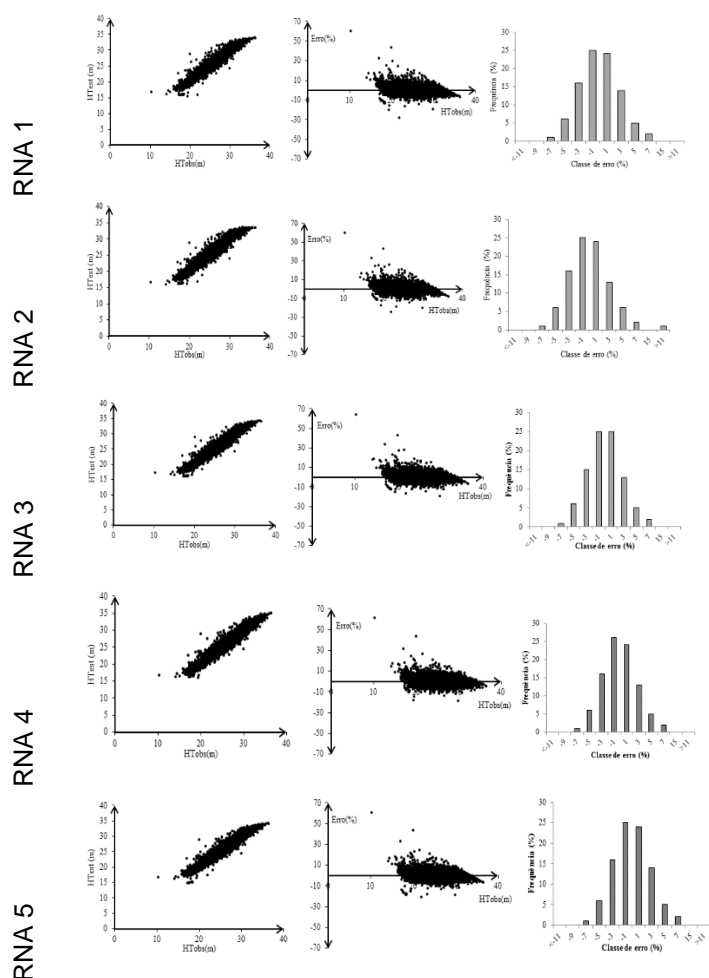
Da mesma forma, a RNA 1, a RNA 2, a RNA 3, a RNA 4 e a RNA 5 apresentaram maior correlação de Pearson (valores variando de 0,9679 a 0,9701), quando comparadas às RNA 6, RNA 7, RNA 8, RNA 9 e RNA 10. Isso indica que a combinação de variáveis independentes incluindo a variável altura dominante explica mais a variação na altura total, quando comparada à combinação que não utiliza a altura dominante.

Na Figura 6 é possível notar que, em geral, todas as RNA apresentaram tendência linear positiva entre os valores estimados e os valores observados.

Por meio da análise do comportamento gráfico dos resíduos é possível perceber homogeneidade da variância adequada para todas as cinco redes treinadas com a combinação de variáveis escolhidas (FIG. 6).

Embora as cinco RNA apresentem pouca variação quanto aos critérios estatísticos e ao comportamento gráfico dos resíduos, o melhor desempenho foi da RNA 4 pois, esta rede apresentou menor valor para a RQEM (0,8852) e maior valor para a r_{HH} (0,9701).

Figura 6 – Tendência linear entre os valores estimados e os valores observados; distribuição dos resíduos, em porcentagem, em função da altura real; e histograma do resíduo para os ajustes por redes neurais com a melhor combinação de variáveis, evidenciando o comportamento das estimativas e desvios encontrados entre a altura observada e a altura estimada.



Fonte: O AUTOR, 2019.

Seleção da melhor metodologia: Por método de comparação, os resultados estatísticos do ajuste por meio do modelo (2) e do ajuste por meio da RNA 4 foram sucintamente descritos na Tabela 10.

Nesta tabela é possível notar que, embora o ajuste do modelo (2) tenha resultado em um menor RQEM e um maior r_{HH} , as duas metodologias avaliadas apresentaram resultados satisfatórios (RQEM maior que 0,8), o que evidencia precisão adequada no ajuste.

Tanto para a regressão quanto para a RNA os valores de bias tenderam a zero. No entanto, o ajuste do modelo (2) apresentou bias positivo, indicando leve tendência em subestimar a altura das árvores, já a RNA 4 apresentou bias negativo, indicando leve tendência em superestimar a altura das árvores (TAB. 10).

Tabela 10 – Coeficientes e parâmetros estatísticos para ambas as metodologias avaliadas.

Metodologia	Bias	r_{HH}	RQEM
Modelo (2)	0,0259	0,9759	0,8094
RNA 4	- 0,0035	0,9701	0,8852

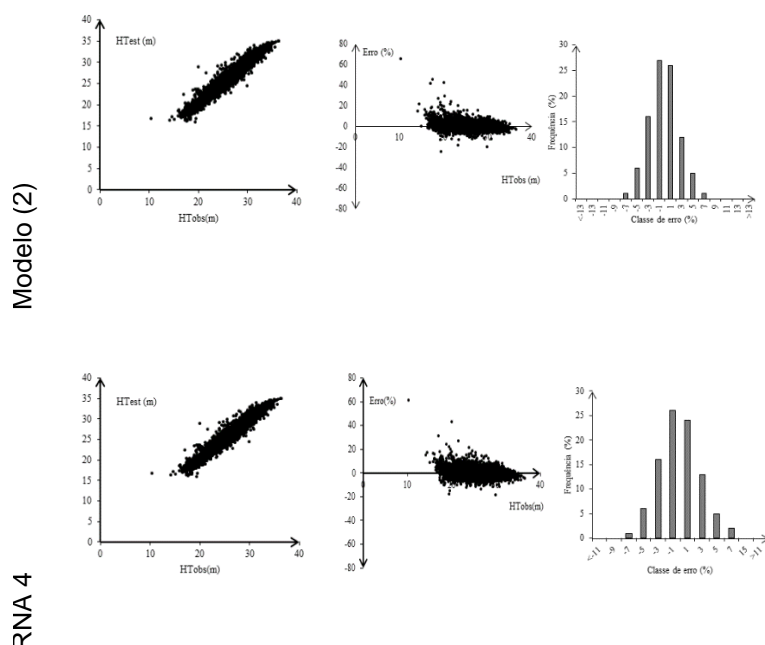
Fonte: O AUTOR, 2019.

Em que: r_{HH} = coeficiente de correlação de Pearson entre os valores observados e estimados; RQEM = raiz quadrada do erro médio.

A Figura 7 evidencia que tanto o modelo de regressão quanto a rede neural apresentam pouca variação quanto à homogeneidade da variância. É possível visualizar também, que ambas as metodologias obtiveram tendência linear adequada entre os valores estimados e os valores observados.

Portanto, considerando todas as observações anteriores, verificou-se que as metodologias avaliadas apresentam coeficientes estatísticos e comportamento gráfico dos resíduos muito próximos.

Figura 7 – Tendência linear entre os valores estimados e os valores observados; e distribuição dos resíduos, em porcentagem, em função da altura real; e histograma do resíduo para os melhores ajustes de cada uma das metodologias avaliadas.



Fonte: O AUTOR, 2019.

CONCLUSÕES:

Dependendo da homogeneidade do povoamento e da quantidade de dados utilizados no ajuste existe diferença estatística entre a estimativa da altura por parcela e/ou por talhão, por meio de equações hipsométricas;

Com base nas equações hipsométricas obtidas, estatisticamente a variável altura dominante explicou melhor a variação da altura total, quando comparada à variável diâmetro dominante;

A estimativa da altura por meio de modelos hipsométricos ou por redes neurais geraram resultados semelhantes, utilizando-se a altura dominante.

REFERÊNCIAS BIBLIOGRÁFICAS:

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

SANQUETA, C.R.; WATZLAWICK, L.F.; DALLA CORTE, A.P.; FERNANDES, L.A.V. **Inventários Florestais: planejamento e execução**. Curitiba: Editora Multi-Graphic, 2006. 270 p.

SCOLFORO, J. R. S. **Mensuração Florestal 3: Relações quantitativas em volume, peso e a relação hipsométrica**. Lavras: ESAL/FAEPE, 1993. 292 p.

SOARES, C. P. B.; PAULA NETO, F.; SOUZA, A. L. **Dendrometria e Inventário Florestal**. Viçosa-MG: Editora UFV, 2007. 276 p.

Participação em Congressos, publicações e/ou pedidos de proteção intelectual:

Bolsa do PROGRAMA INSTITUCIONAL DE INICIAÇÃO CIENTÍFICA E TECNOLÓGICA
PIBIC/CNPq.

Este trabalho teve participação no VII SINTEGRA/UFVJM.