

## ESTIMATIVA DA ALTURA DE ESPÉCIES FLORESTAIS EM REGENERAÇÃO NATURAL UTILIZANDO REDES NEURAIS ARTIFICIAIS

Emanuel França Araújo<sup>1</sup>

Milton Marques Fernandes<sup>2</sup>

Jeferson Pereira Martins Silva<sup>1</sup>

Sustanis Horn Kunz<sup>3</sup>

Marcia Rodrigues de Moura Fernandes<sup>4</sup>

### RESUMO

O uso de redes neurais vem sendo utilizado em vários ramos da ciência. Em estudos de vegetação as redes neurais têm sido utilizadas principalmente para estimativa de biomassa aérea, altura e diâmetro de indivíduos arbóreos, porém, existem poucos estudos com a regeneração natural. Neste contexto, objetivou-se estimar, por meio de redes neurais artificiais, a altura (H) das espécies regenerantes em uma área de um fragmento de Floresta Ombrófila Semidecidual. A base de dados foi proveniente de indivíduos de espécies arbóreas da regeneração de 20 parcelas de 10 m<sup>2</sup> (2 x 5 m), no município de São Cristóvão, SE. Para estimativa da H total por redes neurais artificiais foram testadas redes do tipo *Multilayer Perceptron*. A RNA foi eficiente na estimativa da H do estrato regenerativo em uma área de um fragmento de Floresta Ombrófila Semidecidual. A RNA 12, com 14 neurônios foi mais eficiente para estimar a altura das espécies mais abundantes na área.

**Palavras-chave:** Floresta Atlântica; Sucessão Secundária; Inteligência Artificial.

### ABSTRACT

**Estimation of the height of forest species in natural regeneration using artificial neural networks.** The use of neural networks has been used in various branches of science. In vegetation studies, neural networks have been used mainly to estimate tree biomass, height and diameter of tree individuals, but there are few studies with natural regeneration. In this context, the objective was to estimate, through artificial neural networks, the height (H) of regenerating species in an area of a Semideciduous Ombrophilous Forest fragment. The database came from individuals of tree species from regeneration of 20 plots of 10 m<sup>2</sup> (2 x 5 m), in São Cristóvão, SE. To estimate total H by artificial neural networks, Multilayer Perceptron networks were tested. RNA was efficient in the estimation of H of the regenerative stratum in an area of a Semideciduous Ombrophilous Forest fragment. RNA 12 with 14 neurons was more efficient to estimate the height of the most abundant species in the area. The greater number of individuals with lower heights promoted an overestimation of the height.

**Keywords:** Atlantic Forest; Secondary Succession; Artificial Intelligence.

<sup>1</sup> PPG em Ciências Florestais, Universidade Federal do Espírito Santo – UFES, ES, Brasil.

<sup>2</sup> Depto. de Ciências Florestais, Universidade Federal de Sergipe – UFS, SE, E-mail para correspondência: miltonmf@gmail.com

<sup>3</sup> Depto. de Engenharia Florestal e Engenharia Industrial, Universidade Federal do Espírito Santo – UFES, ES, Brasil.

<sup>4</sup> Secretaria de Estado do Desenvolvimento Urbano e Sustentabilidade, Secretaria de Estado do Desenvolvimento Urbano e Sustentabilidade, Sergipe, Brasil.

## INTRODUÇÃO

As florestas tropicais primárias foram amplamente desmatadas e fragmentadas em todo o mundo (Arroyo-Rodríguez et al., 2017). A conversão sem precedentes das florestas tropicais em terras agrícolas no século XX, resultou em fragmentos com populações reduzidas e isoladamente reprodutivas, o que aumenta os riscos de endogamia, deriva genética e, risco de extinção (Gibson et al., 2011; Laurance et al., 2014).

Partindo do pressuposto que a perda de florestas tropicais primárias provavelmente continuará, particularmente em ecossistemas tropicais potencialmente produtivos, há um crescente interesse no potencial mitigador das florestas secundárias (oriundas do reflorestamento ou da regeneração florestal da natural) para conservação da biodiversidade, serviços ecossistêmicos e mitigação de mudanças climáticas (Chazdon, 2014; Arroyo-Rodríguez et al., 2017; Schwartz et al., 2017).

Assim, estudos sobre a regeneração natural em fragmentos florestais são de suma relevância para o entendimento do funcionamento ecológico desses ecossistemas (Edwards et al., 2017). A regeneração natural é de extrema importância para a restauração florestal, pois auxilia no processo de renovação da composição florística e na manutenção da diversidade de espécies de um local, sendo, portanto, um importante indicador de avaliação e monitoramento da restauração de ecossistemas degradados.

A expressão regeneração natural é definida como sendo o conjunto de descendentes das plantas arbóreas que se encontram entre 0,1 metro de altura (H) até o limite de diâmetro estabelecido no levantamento estrutural (Finol, 1971). Segundo Campos e Leite (2009), existe uma relação entre o diâmetro (D) e a H dos vegetais denominada relação hipsométrica. O emprego de modelos de regressão é usualmente utilizado para estimação dos parâmetros desta função. Recentemente, novas ferramentas de inteligência artificial, tais como, as redes neurais artificiais (RNAs) surgem alternativamente no setor florestal como um potente instrumento de análise de dados, aumentando a precisão das estimativas e diminuição dos custos (Leite et al., 2010).

RNA são definidas como sistemas massivos e paralelos, compostos por unidades de processamento simples, denominados neurônios artificiais ou nodos, que possui a capacidade de aprendizagem por meio de um conjunto reduzido de exemplos, armazenando e generalizando o conhecimento, gerando respostas para dados não conhecidos (Braga et al., 2007; Binoti et al., 2013).

Nesse sentido, este trabalho foi realizado sob a hipótese de que o uso da RNA pode fornecer uma estimativa eficiente e segura da H de indivíduos regenerantes. Assim, objetivou-se desenvolver uma RNA capaz de estimar a H de espécies do estrato regenerativo em uma área de um fragmento de Floresta Ombrófila Semidecidual, no município de São Cristóvão, SE.

## MATERIAL E MÉTODOS

### Descrição da Área de Estudo e dos Dados Utilizados

Os dados utilizados nesse estudo foram obtidos de um fragmento de 200 ha de uma Floresta Ombrófila Semidecidual (Sergipe, 2012), localizada no Instituto Federal de Educação, Ciência e Tecnologia de Sergipe, no município de São Cristóvão, sob as coordenadas geográficas de 10°55'S e 37°07'W, e altitude de 20 m (Figura 1).

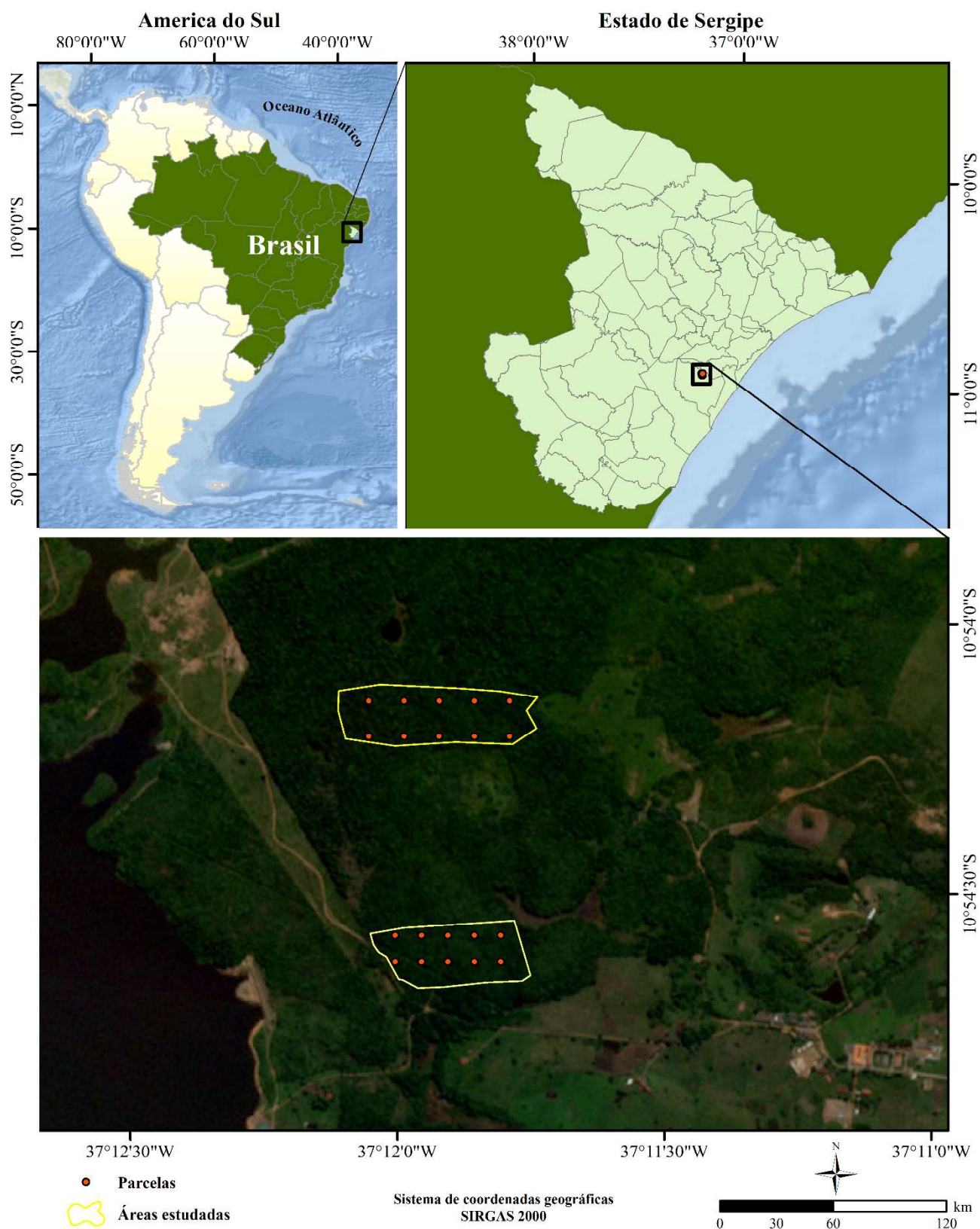


Figura 1. Localização da área de estudo e croqui das parcelas.

A região apresenta clima do tipo tropical com estação seca de verão (As), segundo Köppen, com período chuvoso concentrado entre abril a agosto, com temperaturas elevadas durante todo o ano, com medias anuais em torno de 24 °C (Sergipe, 2012).

Foram alocadas duas áreas em diferentes estágios sucessionais (floresta aberta e floresta fechada). Para cada área foram instaladas 10 parcelas de 20 x 20 m, e em cada parcela foram demarcadas subparcelas de 2 x 5 m. Nas subparcelas foram mensuradas as alturas e diâmetros dos indivíduos vivos em um nível de inclusão menor ou igual a 15 cm de CAB (circunferência na altura da solo < 15 cm). As medições foram realizadas por meio de uma trena milimetrada. A identificação dos indivíduos foi realizada por meio da coleta do material botânico da regeneração natural e por comparação no Herbário da Universidade Federal de Sergipe (ASE), utilizando-se o sistema de classificação de APG III (2009).

No total foram observados 234 de indivíduos em 23 espécies, com Índice de diversidade de Shannon de 2,41. Os indivíduos de espécies arbóreas foram distribuídos em três classes de H: Classe I com indivíduos com H entre, 30 - 99 cm com 115 indivíduos, Classe II com H entre 100 a 199 cm, com 83 indivíduos e a Classe III com H entre 200 a 300 cm, com 36 indivíduos.

Para a estimação da H foram utilizadas as nove espécies com maior número de indivíduos em cada área de estudo, com um total de 136 indivíduos. As principais características descritivas das variáveis quantitativas das nove espécies mais observadas estão apresentadas na tabela 1.

**Tabela 1.** Valores médios e desvio padrão da circunferência à altura da base, diâmetro à altura da base, altura e índice diversidade de Shannon das espécies florestais regenerantes mais observadas nas parcelas. \* Número de indivíduos \*\*Desvio padrão

Espécies	Ni*	CAB (cm)	DAB (cm)	H (cm)	Shannon
Tapirira guianensis	35	3,01 (1,18)**	0,96 (0,37)	104,33 (56,44)	0,272
Protium heptaphyllum	24	3,14 (1,70)	1,00 (0,54)	145,01 (82,99)	0,246
Myrcia splendens	21	2,89 (1,48)	0,92 (0,47)	102,52 (52,64)	0,219
Ocotea bracteosa	18	2,73 (0,94)	0,87 (0,30)	124,82 (51,84)	0,264
Myrcia poliantha	10	2,76 (0,75)	0,88 (0,24)	112,7 (42,64)	0,237
Campomanesia aromatica	9	4,43 (0,88)	1,41 (0,28)	195,4 (43,04)	0,223
Xylopia frutencens	8	2,27 (0,64)	0,72 (0,20)	74,7 (24,56)	0,208
Eugenia florinbunda	6	2,78 (1,05)	0,89 (0,33)	96,33 (50,00)	0,135
Schoefia brasiliensis	5	2,52 (0,54)	0,80 (0,17)	98,62 (25,90)	0,119

### Processamento dos Dados com Redes Neurais Artificiais (ANN)

As RNAs usadas para estimar a altura foram do tipo *Multilayer Perceptron* (MLP), visto que, esse tipo de rede demonstrou grande potencial de aproximação de função em diversos estudos (Nunes e Görgens, 2016; Vieira et al., 2018; Silva et al., 2019).

A MLP (Figura 2) é formada por uma camada de entrada, onde as variáveis são apresentadas à rede; camada intermediária, onde é feito a maior parte do processamento, e a camada de saída, onde o resultado é apresentado. A camada intermediária é responsável por identificar os padrões não lineares dos dados com a utilização de funções de ativação (Braga, 2007). Neste estudo, a função tangente hiperbólica foi usada na camada intermediária e a função linear na camada de saída.

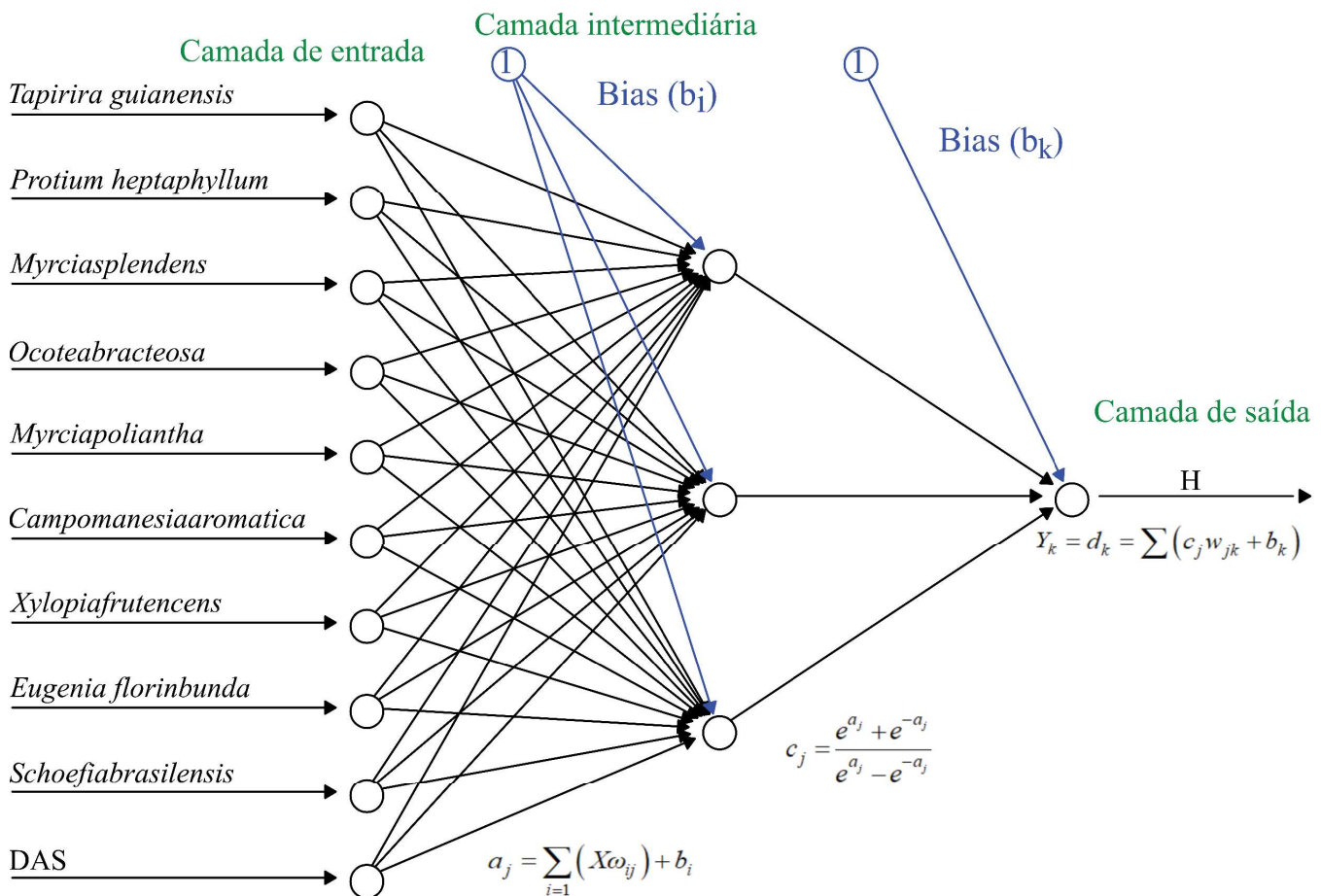


Figura 2. Estrutura de uma rede neural artificial do tipo *Multilayer Perceptron*.

As entradas foram normalizadas em intervalos de -1 a 1, correspondendo à função tangente hiperbólica. Após a normalização, as entradas ( $x_i$ ) foram inseridas na rede e passadas para a camada intermediária ( $a_j$ ), onde a soma ponderada foi calculada de acordo com a Eq. (1), que é o valor da variável ( $x_i$ ) vezes o peso da variável ( $w_{ij}$ ) mais o bias da variável ( $b_i$ ). O produto deste cálculo foi transmitido às funções de ativação tangente hiperbólica [Eq. (2)]. Por fim, os cálculos dos neurônios intermediários serviram de entrada, para os neurônios de saída ( $Y_k$ ) [Eqs. (3)], obtendo a altura estimada.

$$a_j = \sum_{i=1} (x_i w_{ij}) + b_i \quad (1)$$

$$C_j = \frac{e^{a_j} + e^{-a_j}}{e^{a_j} - e^{-a_j}} \quad (2)$$

$$Y_K = d_k = \sum_{j=1} (c_j w_{jk}) + b_k \quad (3)$$

Ao usar o MLP (Figura 2) na estimativa da altura das espécies nativas, foi necessário ajustar os pesos sinápticos das conexões entre as unidades de processamento. Estes foram ajustados em um processo iterativo comumente chamado de aprendizagem (Braga, 2007). O algoritmo de aprendizagem usado foi o Levenberg-Marquardt (Haykin, 2003; Braga, 2007).

As RNAs foram treinadas variando de 3 a 15 o número de neurônios na camada intermediária. As funções de ativação usadas foram a tangente hiperbólica na camada intermediária e a função linear na

cada de saída. As variáveis utilizadas como entrada das RNA para a estimação da H dos indivíduos foram numéricas, diâmetro à altura do Solo (DAS); e categóricas, as espécies.

Para o ajuste dos parâmetros das redes neurais artificiais, utilizou-se o *Toolbox* redes neurais artificiais do *software* MATLAB R2016. Foram treinadas dez redes alterando o número de neurônios na camada intermediária, que variou de 3 a 15, buscando investigar desde a aplicação de uma estrutura mais simples até uma mais complexa. Segundo Braga (2007), quanto mais simples for a estrutura utilizada melhor, desde que forneça estimativas acuradas. Foram estabelecidos quatro critérios de parada do treinamento da RNA. Durante o treinamento, o critério que foi alcançado primeiro estabeleceu o término do processamento. Os critérios analisados foram: número máximo de épocas igual a 1,000, tempo máximo de treinamento igual a 300 segundos, erro máximo de 0.01 e *early stopping*. Para obtenção das RNAs, a base de dados foi dividida aleatoriamente em dois conjuntos: 70% para treinamento e 30% para generalização.

### Métodos para Avaliação da Exatidão das Estimativas

A avaliação das estimativas das H por meio das RNA foi realizada por meio dos seguintes critérios estatísticos: coeficiente de correlação linear simples entre as H observadas e estimadas ( $r_{y\hat{y}}$ ), raiz quadrada do erro quadrático médio (RQEM), análise gráfica dos resíduos (Tabela 2) e histogramas de frequência dos resíduos percentuais.

**Tabela 2.** Estatísticas para avaliar o desempenho das técnicas.  $Y_i$  = variável dependente observada;  $\hat{Y}_i$  = variável dependente estimada;  $\bar{Y}$  = média da variável dependente observada;  $\hat{Y}_m$  = média da variável dependente estimada; e  $n$  = número de observações.

Estatísticas	Fórmulas	Nº
Coeficiente de correlação	$r_{y\hat{y}} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(\hat{Y}_i - \hat{Y}_m)}{\sqrt{\left[ \sum_{i=1}^n (Y_i - \bar{Y})^2 \right] \left[ \sum_{i=1}^n (\hat{Y}_i - \hat{Y}_m)^2 \right]}}$	(01)
Raiz quadrada do erro médio	$RMSE(\%) = \frac{100}{\bar{Y}} \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}$	(02)
Resíduo	$\text{Resíduo}(\%) = \frac{(Y_i - \hat{Y}_i)}{\bar{Y}} \cdot 100$	(03)

## RESULTADOS E DISCUSSÃO

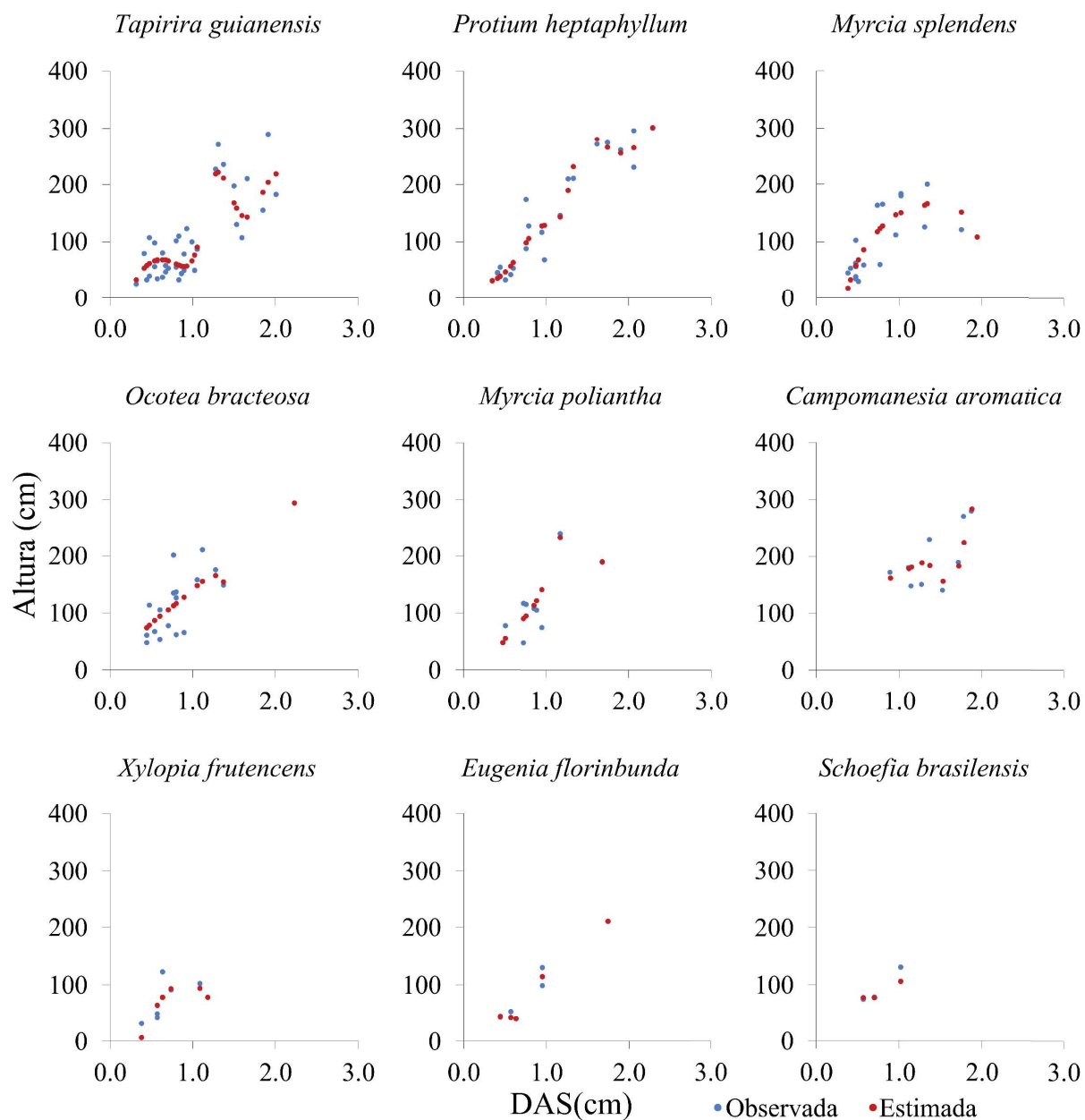
Para a avaliação das RNAs, considerou-se apenas a rede de melhor desempenho estatístico, dentre as testadas de acordo com a quantidade de neurônios na camada intermediária. Na tabela 3 estão apresentadas as estatísticas obtidas para estimar a H das espécies em regeneração.

Todas as RNA apresentaram forte correlação ( $> 0.80$ ) entre as alturas observadas e estimadas, sendo o maior valor encontrado para a RNA 12. Com exceção das RNA 1 e 2 para os dados de treinamento, todas as configurações geraram modelos com RMSE% abaixo de 38%. Nota-se que a RNA 12 também apresentou maior exatidão nas estimativas de altura para os dados de validação, seguida pela RNA 6, que teve oito neurônios na camada intermediária. Isso demonstra que a RNA conseguiu captar a relação biológica em que a altura pode ser explicada pelo diâmetro das diferentes espécies. Além disso, percebe-se que não ocorreu overfitting durante o treinamento da RNA, já que os valores de RMSE% da validação foram próximos aos do treinamento. Dessa forma, com base nos resultados obtidos, deve ser dada preferência para a RNA 12 para a estimativa das alturas das espécies estudadas.

**Tabela 3.** Estatísticas obtidas para a estimativa da altura da regeneração.

RNA	n° neurônios	Treino		Validação	
		Correlação (r)	RMSE (cm)	Correlação (r)	RMSE (cm)
RNA 1	3	0,80	42,96	0,80	47,63
RNA 2	4	0,61	58,69	0,62	57,40
RNA 3	5	0,83	37,97	0,84	44,80
RNA 4	6	0,87	33,75	0,86	40,76
RNA 5	7	0,87	36,39	0,82	42,95
RNA 6	8	0,88	35,32	0,88	33,58
RNA 7	9	0,87	36,95	0,57	64,32
RNA 8	10	0,86	37,47	0,82	40,76
RNA 9	11	0,84	36,02	0,92	36,93
RNA 10	12	0,88	34,78	0,78	42,90
RNA 11	13	0,86	36,91	0,89	36,58
RNA 12	14	0,92	28,23	0,89	33,49
RNA 13	15	0,85	37,59	0,84	40,58

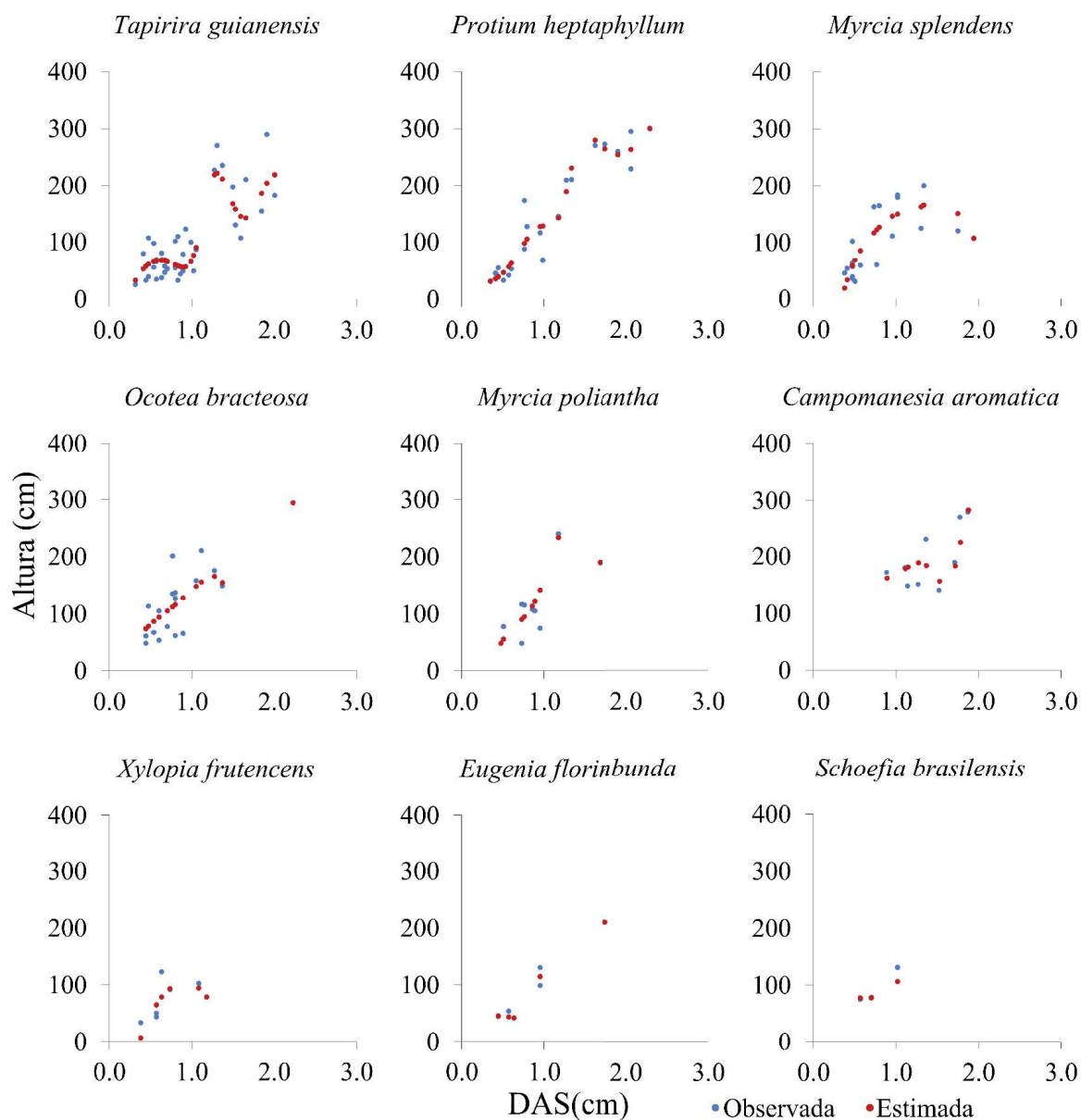
Na figura 3 são apresentadas as análises gráficas para os dados de ajuste e validação da RNA 12 para estimativa das alturas das espécies estudadas. Na análise gráfica dos resíduos em porcentagem (Figura 3c e 3d), verifica-se que os valores não foram igualmente distribuídos, apresentando, assim, uma tendência de superestimar a H em menores valores de Diâmetro a Altura do Solo - DAS (cm). Esses erros mais expressivos nos menores diâmetros podem ser devido à área ter sofrido algum raleamento ou corte seletivo dos indivíduos arbóreos adultos recentemente, resultando em maior número de indivíduos e espécies da regeneração natural nas menores classes de H (Fernandes et al., 2019). Outro fator que pode ter provocado essa maior variação nos menores valores de DAS é a competição. É razoável afirmar que indivíduos mais novos não estão totalmente estabelecidos na área, ou seja, sofrem mais com o efeito da competição.



**Figura 3.** Alturas observada e estimada, distribuição gráfica dos resíduos e histogramas de frequência dos erros percentuais por redes neurais artificiais, treinamento (2a, 2c e 2e) e validação (2b, 2d e 2f) para altura do estrato regenerante.

Na Figura 4 estão apresentadas as estimativas de H por RNA e H a observada em função do DAS para cada espécie. As estimativas de H pela RNA foram satisfatórias para as espécies, ficando mais evidente para as *Tapirira guianensis*, *Protium heptaphyllum*, *Myrcia splendens*, *Ocotea bracteosa* e *Myrcia poliantha*, que são as espécies mais frequente na área de estudo.





**Figura 4.** Alturas observada e estimada em função do DAS para cada espécie.

A RNA do tipo *Multilayer Perceptron* se mostrou uma técnica eficiente na predição da altura das espécies em regeneração no fragmento de Floresta Ombrófila Semidecidual, confirmando a hipótese testada. A estimação do crescimento das árvores é essencial para avaliação das reservas de madeira bem como o potencial de regeneração de regeneração. Isso evidencia a relevância desse estudo, o uso de RNAs para estimação das alturas de espécies nativas em habitat pode fomentar ações de manejo, conservação e também em futuras pesquisas em fragmentos florestais e seus entornos.

De forma geral, as RNAs treinadas demonstraram comportamento semelhante entre si. Contudo, a RNA 12 apresentou as melhores estatísticas para o treinamento e validação (Tabela 2). Além disso, apresentou bom desempenho nas análises gráficas, demonstrando proximidade entre os valores observados e estimados (Figura 3a e 3b) e uma boa distribuição dos resíduos (Figura 3c e 3d), por isso deve ser dada preferência para esta RNA.

Quando se compara o treinamento e a validação em termos de distribuição dos erros (Figura 3c e 3d) nota-se que o comportamento do erro foi semelhante nestas duas situações. Este é um resultado satisfatório, pois, quando a validação reproduz um comportamento semelhante do erro com o encontrado para o ajuste, significa que o modelo tem boa capacidade generalização, o que permite obter alturas fidedignas para um outro conjunto de dados. Complementarmente as distribuições residuais, os histogramas de frequência dos erros apresentaram uma distribuição adequada, com as maiores frequência (40% dos erros) nas classes próximas a -10% e 10% de erro (Figura 3e e 3f).

Os maiores erros concentraram-se nos menores valores de *DAS*, evidenciando maior dificuldade em gerar estimativas exatas para as alturas reais às esses diâmetros. Isso pode ser atribuído a maior heterogeneidade observada nas taxas de crescimento das menores árvores, uma vez que estas são as que mais sofrem com os efeitos da competição. Além disso, o ritmo de crescimento de uma árvore é influenciado por suas características genéticas e sua interação com os fatores ambientais como o clima, o solo e a topografia (Prodan et al., 1997). Árvores mais grossas, que já estão estabelecidas há mais tempo, apresentam menor variabilidade diamétrica (Rozendaal et al., 2015).

## CONCLUSÕES

A rede neural artificial do tipo *Multilayer Perceptron* se mostrou uma técnica eficiente na predição da altura das espécies regenerantes em uma área de um fragmento de Floresta Ombrófila Semidecidual. A configuração da rede neural artificial com 14 neurônios na camada intermediária apresentou maior exatidão nas estimativas geradas para a altura total da espécie em estudo.

## REFERÊNCIAS

- APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. **Botanical Journal of the Linnean Society**, 161(2):105–121.
- ARROYO-RODRÍGUEZ, V. et al. 2017. Multiple successional path ways in human-modified tropical landscapes: new insights from forest succession, forest fragmentation and landscape ecology research. **Biological Reviews**, 92(1):326-340.
- BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. 2013. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, 37(4):639-645.
- BRAGA, A. de P. 2000. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: Livros Técnicos e Científicos, 262p.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. 2007. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 248p.
- CAMPOS, J. C. C.; LEITE, H. G. 2009. **Mensuração Florestal: perguntas e respostas**. 3. ed. Viçosa: UFV, 548p.
- CHAZDON, R. L. 2014. **Second Growth**. Chicago: Chicago University Press, 472p.
- EDWARDS, D. P. et al. 2017. Tropical secondary forest regeneration conserves high levels of avian phylogenetic diversity. **Biological Conservation**, 209(1):432-439.
- FERNANDES, M.M. et al. 2019. Regeneração natural em fragmento de Floresta Ombrófila Semidecidual em Sergipe, Brasil. **Ecologia e Nutrição Florestal**, 7(1):1-10.

- FINOL, U. H. 1971. Nuevos parametros a considerar se en el analisis estrutural de las selvas virgenes tropicales. **Revista Florestal Venezolana**, **14**(21):29-42.
- GIBSON, L. et al. 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. **Nature**, **478**(1):378–381.
- HAYKIN, S. 2007. **Redes neurais: princípios e prática**. Porto Alegre: Bookman Editora, 849p.
- LAURANCE, W. F.; SAYER, J.; CASSMAN, K. G. 2014. Agricultural expansion and its impacts on tropical nature. **Trends in Ecology & Evolution**, **29**(2):107-116.
- LEITE, H. G. et al. 2011. Estimation of inside-bark diameter and heart wood diameter for *Tectona grandis* Linn. trees using artificial neural networks. **European Journal of Forest Research**, **130**(2):263-269.
- MATHWORKS, INC. 2017. **Neural Network toolbox: for use with MATLAB®: user's guide**. Natick, USA: MathWorks, 846p.
- NUNES, M. H.; GÖRGENS, E. B. 2016. Artificial intelligence procedures for tree taper estimation within a complex vegetation mosaic in Brazil. **PloS one**, **11**(5):e0154738.
- PRODAN, M. et al. 1997. **Mensura Forestal**. San José: IICA, 586p.
- ROZENDAAL, D. M et al. 2015. Long-term growth patterns of juvenile trees from a Bolivian tropical moist forest: shifting investments in diameter growth and height growth. **Journal of Tropical Ecology**, **31**(6):519.
- SCHWARTZ, N. B. et al. 2017. Fragmentation increases wind disturbance impacts on forest structure and carbon stocks in a western Amazonian landscape. **Ecological Applications**, **27**(6):1901-1915.
- SERGIPE. 2012. **Florestas em Sergipe: construindo uma Política Florestal**. Sergipe: SEMARH, 360p.
- SILVA, J. P. M. et al. 2019. Computational techniques applied to volume and biomass estimation of trees in Brazilian savanna. **Journal of Environmental Management**, **249**:109-368.
- VIEIRA, G. C. et al. 2018 Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. **Science of the Total Environment**, **619**:1473-1481.