

REDES NEURAIS ARTIFICIAIS NA ESTIMATIVA DE VOLUME EM UM PLANTIO DE EUCALIPTO EM FUNÇÃO DE FOTOGRAFIAS HEMISFÉRICAS E NÚMERO DE ÁRVORES

Fabrcio Assis LEAL¹

Eder Pereira MIGUEL¹

Eraldo Aparecido Trondoli MATRICARDI¹

Reginaldo Srgio PEREIRA¹

- RESUMO: Objetivou-se neste trabalho comparar e testar o desempenho de redes neurais artificiais e método de regressão linear na estimativa de volume em plantio de *Eucalyptus urophylla*, utilizando como variáveis apenas o número de árvores por parcela e abertura de dossel. A área está localizada no município de Rio Verde/Goiás, onde foram delineadas 23 parcelas de 400m², envolvendo área total de 116 hectares. Em cada parcela contou-se o número de árvores e foram tiradas 3 fotos hemisféricas na diagonal. O volume de cada árvore foi obtido pelo método de Hohenadl. Os ajustes foram por meio de regressão linear simples e redes neurais artificiais com dados de 18 parcelas para treinamento e 5 parcelas para teste. Embora os métodos testados tenham sido satisfatórios em relação às estimativas de volume, todos tiveram tendência em superestimar o volume, porém as redes neurais artificiais sempre foram superiores no desempenho, mesmo não diferindo estatisticamente da regressão.
- PALAVRAS-CHAVE: Abertura de dossel; inventário florestal; modelos volumétricos.

1 Introdução

O conhecimento do volume de um povoamento florestal é determinado por meio do inventário quantitativo. Essa operação consiste na mensuração de amostras representativas da população, nomeadas de parcelas (BINOTTI et al., 2013a). Dessa forma, o volume de madeira em cada parcela advém da aplicação de modelos volumétricos que, em sua maioria, possuem dependência de árvores cubadas previamente. Atividade essa bastante onerosa e morosa no setor florestal.

Todavia, a quantificação volumétrica dos povoamentos florestais comerciais depende altos custos às empresas. Estes custos são ainda maiores quando se deseja obter menores erros nas estimativas das variáveis dendrométricas de interesse. Neste contexto, técnicas e ferramentas que melhoram a precisão nas estimativas e minimizam os elevados custos dos inventários florestais, bem como a utilização de variáveis de fácil obtenção na floresta nos ajustes dos modelos, são bem quistas.

¹ Universidade de Brasília - UnB, Faculdade de Tecnologia - FT, Departamento de Engenharia Florestal, CEP: 70.910-900, Brasília, DF, Brasil. E-mail. fabricaoassisleall@gmail.com; miguelederpereira@gmail.com; ematricardi@gmail.com; reginaldosergiopereira@gmail.com.

No tocante ao uso de novas técnicas, uma das alternativas consiste no uso de redes neurais artificiais (RNAs). Segundo Binoti et al. (2013b), as redes são sistemas computadorizados formados por unidades de simples processamento, que se conectam de maneira particular para cumprir alguma tarefa.

Dada sua flexibilidade e por terem demonstrado sucesso em grande variedade de aplicações empíricas, as redes neurais artificiais têm se tornado foco de apreciável atenção como ferramenta para o estudo de modelagem e predição, em diversas áreas do conhecimento (KUAN e WHITE, 1994).

As RNAs apresentam um dos mais modernos artifícios no auxílio à tomada de decisão. Baseia-se em simulações matemáticas idênticas à lógica do raciocínio humano. Dessa forma, as RNAs são obtidas pela modelagem das variáveis relevantes em determinada análise, para resolver ou prever determinado problema. RNAs nada mais é que um modelo matemático cuja grande inspiração foi o cérebro humano (OSÓRIO, 1999).

Segundo Binoti et al. (2012a), a utilização de redes neurais artificiais (RNA) tem se apresentado alternativa promissora em relação às técnicas de regressão. Fato esse corroborado por vários trabalhos que visaram adaptar e aplicá-las em diversas situações e abordagens (BINOTI et al., 2013a; BINOTI et al., 2012a; BINOTI et al., 2012b; BINOTI, 2010; LEITE et al., 2010; SILVA et al., 2009a; DIAMANTOPOULOU, 2005).

Outra vantagem apresentada pelas RNAs é o fato de conseguirem trabalhar e desempenhar bons ajustes utilizando variáveis pouco tradicionais nos ajustes como, por exemplo, a possibilidade de modelagem com variáveis categóricas, também chamadas de qualitativas (HAYKIN, 2001).

Nessa vertente de variáveis qualitativas, as informações e dados obtidos por meio de fotografias hemisféricas vêm ganhando destaque com ampla utilização no meio florestal. As fotografias hemisféricas apresentam baixo custo e versatilidade em seu processamento. Dessa forma, diversos foram os trabalhos que utilizaram as fotografias hemisféricas. Por exemplo, citam-se as seguintes abordagens: métodos indiretos para medição de abertura do dossel (SILVA et al., 2009b; MONTE et al., 2007); dinâmica da cobertura de dossel (CHAVES et al., 2007); abertura de dossel, arquitetura de copa e índice de área foliar (HALE, 2003; WANG e MILLER, 1987); aplicação de fotografias hemisféricas coletadas no infravermelho próximo (OSMOND, 2009; CHAPMAN, 2007; MILTON, 2002), entre outras.

Diante do exposto o objetivo desse trabalho foi comparar e testar o desempenho das redes neurais artificiais e o método de regressão tradicional na estimativa de volume, em um plantio florestal comercial, utilizando como variáveis independentes apenas o número de árvores por parcela e a abertura de dossel (fotografias hemisféricas).

2 Material e métodos

2.1 Caracterização da área de estudo

O presente trabalho foi realizado em um plantio de *Eucalyptus urophylla* S. T. Blake, com seis anos e sete meses de idade, plantado em espaçamento 3x2 m. A área de estudo envolve 116 hectares, pertencente à Cooperativa Agroindustrial dos Produtores Rurais do Sudoeste Goiano (COMIGO), município de Rio Verde/Goiás. O local de estudo

apresenta altitude média de 685 metros, limitado entre os paralelos 18° 00' 45'' a 18° 01' 45'' de latitude Sul, e entre os meridianos 50° 52' 45'' a 50° 53' 15'' de longitude Oeste.

O clima, segundo a classificação de KÖPPEN, é o tipo Aw, tropical úmido, caracterizado por duas estações bem definidas: uma seca, que corresponde ao outono e inverno, indo do mês de maio ao mês de setembro, e a outra úmida, com chuvas torrenciais, correspondendo ao período de primavera e verão. A precipitação média anual varia de 1.200 a 1.500 mm, com média anual em torno de 1.300 mm e temperaturas médias entre 20° e 25°C (SIQUEIRA NETO et al., 2011).

2.2 Coleta de dados

2.2.1 Inventário florestal

Foi realizado o inventário florestal na área sendo amostradas aleatoriamente 23 parcelas de 400m² (20x20m). Em todas as parcelas foram medidos os DAP's (diâmetro a altura do peito/1,3 m) de todos os indivíduos contidos dentro dos limites de cada unidade amostral. A variável altura total (HT) foi mensurada e obtida diretamente por meio da cubagem rigorosa. O erro amostral considerado no inventário florestal foi de 10%.

A escolha das árvores a serem cubadas rigorosamente foi realizada em áreas próximas as unidades amostrais, após caracterização quanto à estrutura diamétrica proposta por Sturges (FINGER, 1992). Ao todo foram cubadas noventa e oito (98) árvores, sendo que cada uma delas foi abatida ao nível do solo com o uso de motosserra e o diâmetro inicial tomado neste ponto.

A cubagem foi realizada pelo método relativo de Hohenadl, em 10 seções, sendo a: 5, 15, 25, 35, 45, 55, 65, 75, 85 e 95% da altura total da árvore. Nesses pontos foram obtidas as áreas seccionais para o cálculo do volume de cada árvore (CAMPOS e LEITE, 2013; MACHADO e FIGUEIREDO FILHO, 2009).

A altura total e o volume (m e m³) foram obtidos de forma individual para cada árvore, por meio de equações desenvolvidas e validadas na própria área (LEAL et al., 2011). A seguir as equações ajustadas para altura total (1) (PRODAN, 1965) e volume total (2) (SCHUMACHER e HALL, 1933), são apresentadas. Ambas as equações ajustadas vieram da base de dados das 98 árvores cubadas rigorosamente (LEAL, 2013; LEAL et al., 2011). As estatísticas de ajuste e precisão de ambas as equações são apresentadas na Tabela 1.

$$HT = \frac{DAP^2}{2,86349 - 0,0972117 * DAP + 0,0349208 * DAP^2}, \quad (1)$$

$$Vol = 0,00004153 * DAP^{1,69978} * HT^{1,21340}, \quad (2)$$

em que: DAP = diâmetro a altura do peito (cm); HT = altura total (m).

De posse das equações citadas na Tabela 1 foram obtidas a altura total e o volume individual para cada árvore, em cada parcela amostrada. O volume total por parcela foi obtido somando os volumes individuais das árvores. O número de árvores por parcela (NA), variável independente que compôs as análises adveio da contagem simples dos indivíduos em cada uma das 23 parcelas.

Tabela 1 - Estatísticas de ajuste e precisão para as equações volumétricas e hipsométricas supracitadas

Equações ajustadas	Syx	Syx (%)	R ² ajustado	R ² (%)	CV (%)
(1) Prodan	1,97	8,30	0,84	84,10	41,71
(2) Schumacher e Hall	0,02	7,14	0,99	98,90	3,79

Syx = erro padrão da estimativa absoluto (m e m³); Syx (%) = Erro padrão da estimativa em percentagem; R² Ajustado = coeficiente de determinação ajustado; R² (%) = coeficiente de determinação em percentagem; CV(%) = coeficiente de variação da regressão em percentagem.

2.2.2 Fotos hemisféricas

As fotos hemisféricas foram obtidas utilizando a câmera *Nikon Coolpix 995/3.34 megapixels* (38-152mm de zoom óptico), com lente hemisférica *Nikon Fisheye Converter FC-E8 0,21x (LC-ER1)*. Utilizou-se suporte apoiado e nivelado a um metro do nível do solo.

Foram adquiridas três (3) fotos hemisféricas na diagonal de cada uma das parcelas amostradas. Para a estimativa da porcentagem da abertura do dossel, as fotos hemisféricas foram processadas no software *Gap Light Analyzer 2.0* (FRAZER et al., 1999).

As fotos foram convertidas em preto e branco (binário) e recortadas no formato requisitado pelo *Gap Light Analyzer*, em seguida foram divididas radialmente em círculos concêntricos equidistantes, de acordo com a geometria óptica da lente, de maneira que, cada setor, represente proporção igual do hemisfério. Esse cuidado objetivou a eliminação do efeito de visada oblíqua crescente à medida que se afasta do centro da fotografia, também chamado de efeito *vignette*. E finalmente, a fotografia foi segmentada em bitmap binário usando limiar estimado visualmente (SILVA et al., 2009b). Posteriormente obteve-se a média dos valores obtidos (percentagem) a partir das três fotos, para cada parcela.

Dessa forma, a abertura de dossel em percentagem (AD) também foi utilizada como variável independente nos métodos aqui utilizados (regressão e redes neurais), para estimativa do volume por parcela. Variável essa estimada em função do número de árvores (NA) e abertura de dossel em percentagem (AD).

2.2.3 Modelo de regressão e redes neurais artificiais

O modelo de regressão foi ajustado a partir da análise de correlação entre as variáveis medidas na floresta (abertura de dossel e número de árvores por parcela), com o volume por parcela. Para estimar o estoque em volume total de cada parcela foram utilizados os procedimentos de regressão linear múltiplo (método dos mínimos quadrados ordinários/MQO), e redes neurais artificiais.

Porém, com intuito de validar todos os métodos supracitados e avaliar o desempenho de cada um, aleatoriamente foram sorteadas dezoito (18) parcelas que fizeram parte dos ajustes e cinco parcelas foram utilizadas para validação (teste). Essas cinco parcelas não fizeram parte da de dados dos ajustes, tanto para regressão quanto redes neurais.

O modelo de regressão adotado foi o seguinte:

$$Y = \beta_0 + \beta_1 \cdot AD + \beta_2 \cdot NA, \quad (3)$$

em que: Y = volume total por parcela; AD = abertura de dossel (%); NA = número de árvores por parcela; e β_i = estimadores dos parâmetros a serem ajustados.

2.2.4 Redes neurais (RNAs)

As RNAs são modelos computacionais, composto por elementos de processamento simples, chamados neurônios artificiais, que aplicam determinada função matemática aos dados (função de ativação), gerando uma única resposta. Estes neurônios são ligados entre si por conexões, sendo estas conexões, geralmente, associadas a coeficientes denominados de pesos, os quais são ajustados por treinamentos/aprendizados, sendo responsáveis pela retirada das peculiaridades da base de dados e o armazenamento do conhecimento das redes. (BRAGA et al., 2007).

Segundo Binoti et al. (2010), um neurônio artificial é a unidade de processamento de informação de uma RNA, constituído de “n” entradas x_1, x_2, \dots, x_n (dendritos) e uma saída y (axônio). As entradas são associadas a pesos w_1, w_2, \dots, w_n que representam as sinapses. Essas podem ser negativas ou positivas. Atualmente, um modelo básico de neurônio artificial pode ser representado matematicamente como:

$$Y_k = \varphi(V_k), \quad (4)$$

em que: Y_k = saída do neurônio artificial; φ = função de ativação; V_k = resultado do combinador linear, ou seja:

$$V_k = \sum_o^m x_m \cdot w_m, \quad (5)$$

em que: V_k = combinador linear; x_m é o número de entradas; e w_m é o peso para cada entrada de m .

Para o ajuste das RNAs, as variáveis numéricas foram normalizadas linearmente no intervalo de 0 a 1. A camada de entrada foi constituída de dois (2) neurônios, sendo um neurônio para cada variável numérica preditora, em função da variável resposta/saída. Como saída, utilizou-se o volume total por parcela. As redes foram constituídas de apenas uma camada oculta, em que o número de neurônios nessa camada foi igual a dois (que corresponde ao número de neurônios na camada de entrada), logo a arquitetura da rede foi 2-2-1, e como função de ativação utilizou-se a *sigmoidal*.

A função de ativação sigmoide é a mais usual na elaboração de redes neurais artificiais (HAYKIN, 2001), e matematicamente é dada por:

$$\varphi(v) = \frac{1}{1 + \exp^{-\beta u}}, \quad (6)$$

em que: φ = função de ativação sigmoide; β = estimativa do parâmetro que determina inclinação da função sigmoide; u = potencial de ativação da função.

Aconselha-se que no uso de RNAs sejam indicadas configurações mais simples, com o menor número de neurônios possível na camada oculta, a fim de evitar o *overfitting*

(sobre ajuste). Isso consiste no aprendizado exagerado das informações contidas nos dados oferecidos às redes, pois as mesmas se tornam tão bem treinadas sobre o conjunto de dados, que acabam copiando, não apenas a semelhança estrutural entre as variáveis, mas, também, o ruído (erro da relação), conseqüentemente, estas redes com *overfitting* não podem ser utilizadas no conjunto de dados por inteiro, pois sua capacidade de generalização foi comprometida. Também, configurações mais simples facilitam o processo de busca e otimização da configuração para determinada tarefa (RUSSEL e NORVIG, 2010).

A opção pela escolha do algoritmo de treinamento interfere, especificamente, na fuga de mínimos locais, na atuação da tarefa desejada e na duração de treinamento. No presente estudo o algoritmo de treinamentos utilizado foi o *resilient propagation*, proposto por Riedmiller e Braun (1993), como alternativa mais eficiente e recomendada para RNAs do tipo *Multilayer Perceptron*.

Segundo Carvalho et al. (1998), denomina-se algoritmo de treinamento um conjunto de regras bem definidas para a solução de um problema de aprendizado. No *resilient propagation* os pesos são baseados em informações dos dados presentes, para isso é introduzido para cada peso o valor individual de atualização. Inicialmente os pesos de todas as redes foram gerados aleatoriamente (HEATON, 2010). Posteriormente, esse valor individual de atualização evolui durante o processo de aprendizado baseado na função do erro. As estimativas do volume foram simuladas com as combinações possíveis das variáveis de entrada, totalizando três combinações para a variável resposta.

Salienta-se que o aprendizado das redes foi do tipo supervisionado, logo foram dados para as redes dois conjuntos de valores: o conjunto de valores de entrada e o conjunto de valores de saída. Assim, o treinamento consiste em um problema de otimização dos parâmetros da rede (seus pesos sinápticos), para que possam responder às entradas conforme esperado, até que o erro entre os padrões de saída gerados pela rede alcance valor mínimo desejado (PORTUGAL, 1995).

Foram treinadas 100 redes do tipo *perceptrons* de múltiplas camadas, comumente conhecidas como MLP (*Multilayer Perceptron*). Esse tipo de rede é o mais usual na grande maioria dos trabalhos observados. A sua notoriedade fundamenta-se no fato de ser a generalização de formas, traçados e padrões a sua principal aplicação.

Existem vários métodos para a determinação do momento em que o treinamento de uma rede neural deve ser encerrado. O número exagerado de ciclos pode levar a rede à perda do poder de generalização (*overfitting*). Também, com pequeno número de ciclos a rede pode não chegar ao seu melhor desempenho (*underfitting*). Como critério de parada do algoritmo de treinamento utilizou-se o número total de ciclos igual a 3.000 ou erro quadrático médio inferior a 1%, conforme sugerido por Silva e Oliveira (2006).

Portanto, finalizou-se o treinamento quando um dos critérios foi atingido e a melhor rede para estimar o volume foi selecionada. Para o treinamento das redes neurais artificiais utilizou-se o programa *NeuroForest Star* (BINOTI et al., 2013b).

2.2.5 Comparação entre resultados da regressão e redes neurais artificiais

Os critérios escolhidos para comparação entre a técnica de regressão e redes neurais artificiais, na estimativa do volume, foram: coeficiente de determinação (R^2), erro padrão da estimativa absoluto e em porcentagem (S_{yx}) e análise gráfica dos resíduos (DRAPER e

SMITH, 1998).Adicionalmente, os modelos ajustados foram comparados por meio do *Critério de Informação de Akaike (AIC)*.

Segundo Akaike (1974), o AIC é uma medida de qualidade relativa do ajuste e é definido conforme Equação 7. Por fim, foi utilizado o *Critério de Informação Bayesiano (BIC)*, que segundo Schwarz (1978), sua determinação se dá pela expressão definida pela Equação 8.

$$AIC = -2 \ln(L_p) + 2p, \quad (7)$$

$$BIC = -2 \ln(L_p) + K_p, \text{ Para } K = \ln(n), \quad (8)$$

em que: L_p = valor que torna máxima a função de máxima verossimilhança do modelo estimado; p = número de parâmetros do modelo; n = número de observações.

Tanto o AIC quanto o BIC aumentam conforme a soma dos quadrados dos erros aumenta. Desta forma, quanto menor o valor do AIC e BIC de determinado modelo, menor sua discrepância geral estimada e, por consequência, menor sua distância relativa para os valores reais observados (FLORIANO et al., 2006).

2.2.6 Validação do desempenho da regressão e redes neurais artificiais.

Na validação do desempenho da estimativa do volume de ambos os métodos foram utilizadas as cinco (5) parcelas não contempladas nos ajustes iniciais. Para verificar a precisão e a eficiência dos métodos utilizaram-se as seguintes estatísticas: Teste *Qui Quadrado*, erro padrão da estimativa absoluto e em porcentagem (S_{yx}), diferença agregada em porcentagem ($Da\%$), erro médio absoluto (E_i) e gráfico de dispersão de resíduos.

3 Resultados e discussão

3.1 Inventário Florestal

Após aplicação da equação 1 à base de dados foram obtidas as alturas totais de cada árvore na parcela, posteriormente, foi aplicada a equação 2 e o volume individual das árvores nas parcelas foi obtido. O volume individual de cada árvore foi somado, em sua respectiva parcela, e o volume total por parcela foi conhecido. Já o número de árvores por parcela adveio da contagem simples dos indivíduos contemplados em cada uma delas. O resultado do volume para cada parcela amostrada em campo, bem como o número de árvores são apresentados na Tabela 2.

Com os dados da Tabela 2, o erro amostral relativo do inventário florestal foi de 7,9%, ficando abaixo do erro máximo admissível no inventário florestal (10%), e corresponde ao erro amostral absoluto de 0,8067 m³/parcela. Esse resultado não indica nenhum impedimento para proceder as análises sugeridas nessa abordagem (SCOLFORO e MELLO, 2006).

Tabela 2 - Volume total de madeira (m³) e número de árvores para todas as 23 parcelas amostradas na área de estudo

Parc.	Árv.	Vol.	Parc.	Árv.	Vol.	Parc.	Árv.	Vol.	Parc.	Árv.	Vol.
1	59	12,4046	7	59	9,8879	13	54	9,6791	19	39	7,9867
2	57	11,7842	8	65	9,6880	14	52	8,6597	20	50	9,1372
3	48	8,8002	9	58	8,8266	15	61	11,4998	21	45	7,9530
4	55	11,5512	10	63	10,2612	16	42	8,5599	22	52	10,1026
5	48	10,3860	11	66	10,7197	17	45	8,6483	23	49	9,2956
6	58	11,1312	12	66	11,1831	18	36	8,5444	-	-	-

Parc. = parcela; Árv. = número de árvore por parcela; Vol. = volume (m³/400m²).

3.2 Fotografias hemisféricas

A Figura 1 exemplifica o processo para obtenção do valor de abertura de dossel (%) e também apresenta, como exemplo, as fotografias hemisféricas oriundas da parcela 15. Dessa forma, os valores da abertura de dossel de todas as parcelas foram obtidos.

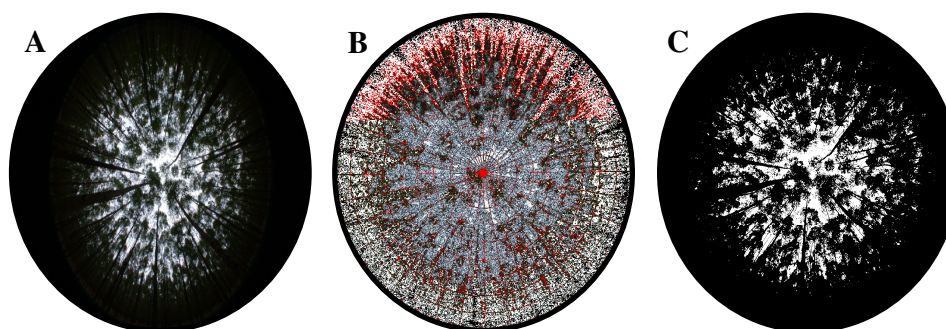


Figura 1 - Fotografias hemisféricas da parcela 15 que exemplificam o processamento utilizado em todas as outras parcelas. A saber: fotografia hemisférica original (A); fotografia hemisférica pré-processada em círculos concêntricos equidistantes (B); fotografia hemisférica binária processada (C).

Após o processamento das fotografias hemisféricas calculou-se a abertura de dossel (%) por meio dos arquivos binários. A Tabela 3 apresenta os valores médios encontrados para cada uma das 23 parcelas amostradas em campo.

As fotografias hemisféricas são alternativas menos onerosas e acessíveis em relação aos métodos tradicionais para obtenção de volume e apresenta elevado potencial de aplicação no meio florestal, conforme demonstrado por Chaves et al. (2007), na obtenção do índice de cobertura de dossel (ICD) em povoamentos de eucalipto.

Tabela 3 - Abertura de dossel em porcentagem para cada uma das 23 parcelas amostradas em campo

Parcela	Dossel	Parcela	Dossel	Parcela	Dossel	Parcela	Dossel
1	19,5	7	16,9	13	17,2	19	20,1
2	20,3	8	16,7	14	18,4	20	22,7
3	17,7	9	17,1	15	17,2	21	25,7
4	16,5	10	18,1	16	25,2	22	23,1
5	18,5	11	16,0	17	25,2	23	25,7
6	17,0	12	16,2	18	19,9	-	-

Em média a abertura de dossel foi de 19,6% (Tabela 3). Em trabalho com dinâmica de cobertura de dossel Chaves et al. (2007), calcularam que em média o índice de cobertura de dossel foi de 82%, sendo a abertura de dossel igual a 18%(sem desbaste e desrama).

3.3 Abordagem descritiva sobre as variáveis utilizadas nos ajustes

Os valores médios, mínimos e máximos do volume, número de árvores e abertura de dossel por parcela e também suas medidas de dispersão são apresentadas na Tabela 4. Essa Tabela também mostra o grau de correlação entre as variáveis analisadas.

Tabela 4 - Média, valor máximo e mínimo, desvio padrão e coeficiente de variação para o volume por parcela, número de árvores e abertura de dossel (%), bem como correlação de *Pearson*, entre variáveis respostas e preditoras

	Volume (m ³ /parcela)	Número de árvores	Abertura de dossel (%)
Média	9,8561	53	19,6
Máximo	12,4046	66	7,1
Mínimo	7,9530	36	16,0
Desvio Padrão	1,2951	9	25,7
CV (%)	13,1%	16,0%	16,9%
Correlação entre variáveis resposta e preditoras			
	Volume	Número de árvores	Abertura de dossel
Volume	1	-	-
Número de árvores	0,69*	1	-
Abertura de dossel	-0,48*	-0,64*	1

*Correlação de *Pearson* significativa a 5%; CV = Coeficiente de variação.

A utilização do número de árvores como variável preditora se justifica pela sua correlação significativa com volume (0,69). O aumento do número de árvores por unidade de área resulta no aumento do volume, caracterizando assim, correlação positiva. A abertura de dossel também apresentou correlação significativa (-0,48). Todavia, o seu aumento caracteriza diminuição no volume, conferindo assim valores negativos para a correlação.

Tais variáveis são de fácil obtenção e de custo relativamente baixo, em detrimento aos métodos tradicionais de inventário para obtenção do volume (MELLO, 2004). A variação da abertura de dossel em relação ao número de árvore em todas as parcelas amostradas pode ser visualizada na Figura 2.

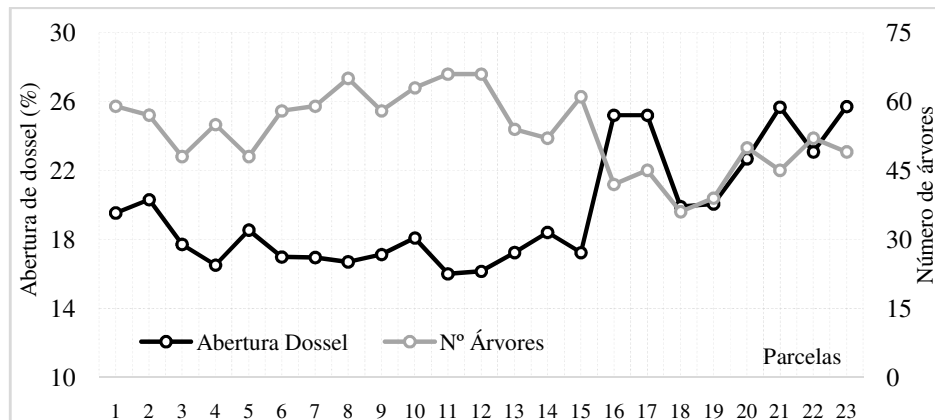


Figura 2 - Comportamento da abertura de dossel (%) em relação ao número de árvores, ambas por parcela, em todas as unidades amostrais.

De acordo com a Figura 2 é notória a percepção de que quando há mais árvores na parcela há menos abertura de dossel. Essa correlação inversa (-0,64) foi constatada na Tabela 4 e esse aparente contraste reflete na escolha das duas variáveis para a estimativa do volume.

3.4 Comparação da regressão e redes neurais na estimativa do volume

Em geral, os modelos ajustados (regressão e redes) apresentaram bom desempenho na estimativa do volume, com valores baixos para o erro padrão da estimativa (Syx) e critérios AIC e BIC. E de moderado a alto para o coeficiente de determinação (R^2), conforme especificados na Tabela 5.

Tabela 5 - Estatísticas de ajuste e precisão para redes neurais artificiais e regressão, na estimativa do volume em função da abertura de dossel e número de árvores

Método	R^2	Syx (m ³)	Syx (%)	AIC	BIC
Redes	0,99	0,1909	1,91	153,2	164,3
Regressão	0,60	0,8680	8,66	384,2	398,3

R^2 = coeficiente de determinação; Syx = erro padrão da estimativa absoluto e percentual; AIC = estatística Akaike; BIC = critério de informação bayesiano.

Ao analisar a Tabela 5 é possível observar que ambos os métodos apresentaram estatísticas de ajuste e precisão aceitáveis. Destaca-se o desempenho das redes neurais

artificiais por apresentar erro padrão da estimativa de 1,91% e alto coeficiente de determinação (0,99). No entanto, os resultados obtidos pelo procedimento de regressão também são considerados satisfatórios, com erro padrão da estimativa inferior a 10%.

Mesmo que os estimadores de ajuste sejam bons indicadores para a seleção de modelos, a análise gráfica de resíduo é quesito fundamental na escolha de uma equação, pois erros de tendência podem ocorrer em determinada amplitude de classe da variável resposta, sem ser detectadas pelas estatísticas que medem a exatidão (DRAPPER e SMITH, 1998). Na Figura 3 pode ser observada a análise gráfica do resíduo em relação ao comportamento do modelo de regressão e redes neurais artificiais.

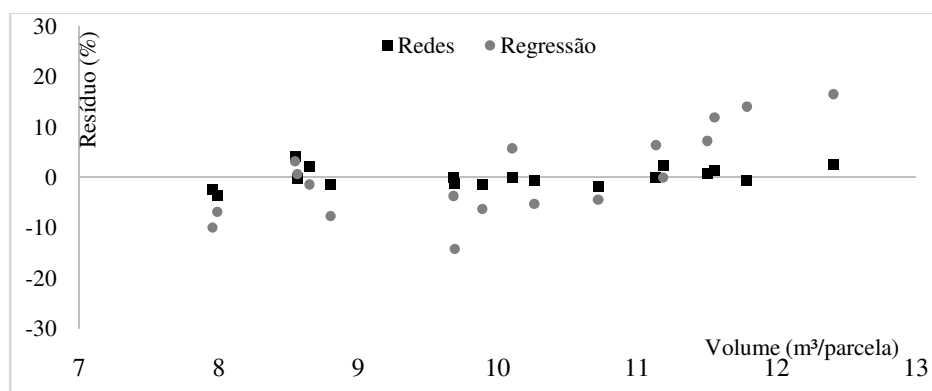


Figura 3 - Distribuição dos resíduos em relação às estimativas do modelo de regressão e redes neurais artificiais, para as 18 parcelas utilizadas no treinamento (ajuste).

Por meio da análise gráfica de resíduos fica evidenciado que o desempenho das redes neurais artificiais foi superior ao modelo de regressão adotado. Percebe-se que, ao longo do eixo de erro zero (0), concentra-se a maioria da dispersão das redes neurais, em relação ao modelo de regressão que variou em aproximadamente 20%.

A superioridade das RNAs frente aos modelos de regressão, na maioria das vezes é devido a fatores de estrutura e distribuição paralela, celeridade no aprendizado e na habilidade de generalizar. Dessa forma, as RNAs resolvem problemas mais complexos e conseguem modelar uma gama de variáveis, bem como entre relações não lineares (HAYKIN, 2001).

Alguns trabalhos têm destacado o potencial das RNAs em relação a procedimentos tradicionais no meio florestal. Silva et al. (2009a), avaliaram o desempenho das RNAs na estimativa do volume de madeira de eucalipto, onde concluíram que as RNAs tiveram boa adequação as situações testadas e recomendaram o uso para previsão volumétrica.

Por sua vez, Binoti et al. (2012a), propuseram, construíram e validaram modelo para estimar a altura de povoamentos equiâneos de eucalipto, visando redução dos custos do inventário florestal. Os autores concluíram que o uso das RNAs foi eficiente e permitiu reduzir o número de medições da variável altura, sem nenhum prejuízo a precisão requerida.

O reflexo do bom desempenho das redes neurais artificiais em relação ao modelo de regressão na estimativa do volume, também ressaltado nos trabalhos supracitados, pode ser visualizado na Figura 4, onde é possível notar pequenos desvios nas estimativas das redes neurais em relação aos dados observados, nas 18 parcelas utilizadas no treinamento.

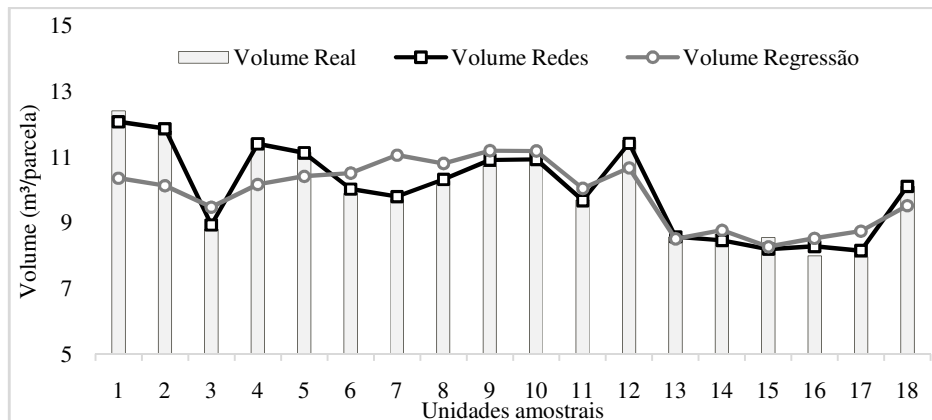


Figura 4 - Desempenho das redes neurais e regressão na estimativa do volume nas 18 parcelas utilizadas no ajuste.

De acordo com a Figura 4, ambas as categorias de modelos se ajustaram ao banco de dados, todavia, as redes neurais foram mais eficientes na estimativa do volume por parcela, conforme demonstradas nas estatísticas que medem exatidão (Syx, AIC, BIC) e corroborada na Figura 4.

3.5 Validação do desempenho da regressão e redes neurais artificiais

Para verificar a confiabilidade dos modelos ajustados quando comparado aos valores reais, do total de unidades amostrais (23) aproximadamente 25% foram separadas de forma aleatória para a validação (5). Assim, foram obtidas as estatísticas de precisão conforme apresenta a Tabela 6. Essas estatísticas mediram o desempenho da regressão e da rede neural na estimativa de volume por parcela, quando comparados com os valores de referência.

Tabela 6 - Estatísticas de precisão para redes neurais artificiais e regressão, após o processo de validação utilizando cinco (5) parcelas

Validação	Syx (m³)	Syx (%)	Ei	Da (m³)	Da (%)	χ^2
Redes	0,3343	3,6	-0,0108	-0,4686	-5,1	0,06
Regressão	1,5469	16,7	-0,0338	-1,7890	-19,3	0,48

Syx = erro padrão da estimativa absoluto e percentual; Ei = erro médio absoluto; Da = diferença agregada absoluta e em porcentagem; χ^2 = qui quadrado calculado.

Conforme valores de Qui Quadrado calculado, tanto para regressão quanto para redes neurais, é possível notar que não houve diferença estatística entre os dois métodos,

ao nível de 5% de significância. Onde o valor de Qui Quadrado tabelado foi 9,49, para o mesmo nível de significância e 4 graus de liberdade do resíduo. Ambas as técnicas podem ser utilizadas na estimativa do volume por unidade de área, tendo a abertura de dossel e o número de árvores como variáveis preditoras.

Embora o número de parcelas para a validação não tenha sido abundante (aproximadamente 25% do total), e o teste de Qui-Quadrado não apontar diferença entre as duas categorias de ajustes, os resultados obtidos foram satisfatórios e revelaram que as redes neurais se sobressaíram em relação à regressão quando analisadas as demais estatísticas.

As redes apresentaram erro padrão da estimativa relativo de 3,6%, ligeiramente superior ao processo de ajuste inicial com 18 parcelas. O aumento do erro padrão da estimativa também foi verificado na equação de regressão, porém de forma mais acentuada com valores de 16,7%, para referida estatística.

A diferença agregada (D_a) é um parâmetro estatístico usado como índice que mede o ajuste de modelos e corresponde à diferença entre o somatório dos valores observados e o somatório dos valores estimados. Esse índice serve como critério indicador de sub ou superestimativa e no presente trabalho, para melhor visualização, a mesma foi obtida na forma percentual. Tanto a equação de regressão como as redes apresentaram valores percentuais negativos de D_a , para a estimativa do volume por unidade de área (-19,3% e -5,1%, respectivamente). Tais valores caracterizam superestimativas na variável estimada.

Os resíduos médios (E_i) gerados pelas duas categorias de ajustes foram próximos de zero, fato esse desejável. Pois demonstraram capacidade dos modelos em estimar as variáveis de interesse com acurácia. Por regressão, o valor médio do resíduo foi de -0,0338 e da RNA de -0,0108 m^3 /parcela.

Para melhor compreensão do comportamento de ambos, a análise gráfica mais detalhada dos resíduos foi então utilizada na validação. A Figura 5 mostra o desempenho dos métodos avaliados (redes e regressão) no processo de validação.

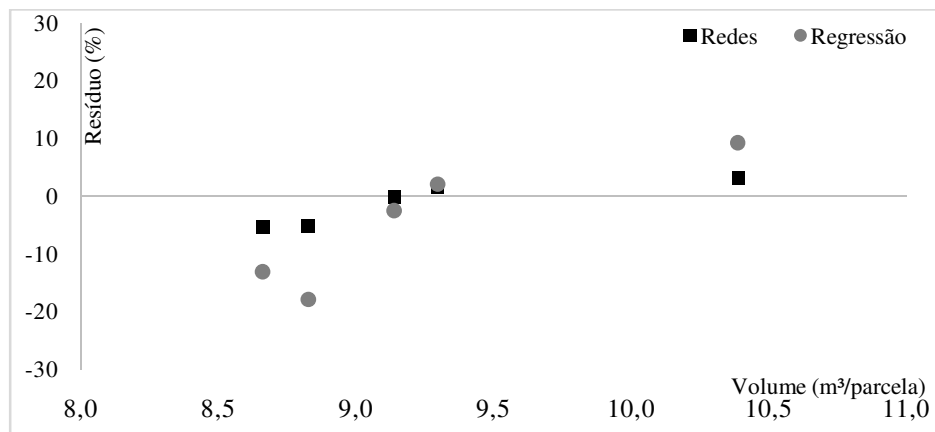


Figura 5 - Distribuição dos resíduos em relação às estimativas do modelo de regressão e redes neurais artificiais no processo de validação.

Embora todos os métodos testados nesse trabalho tenham sido satisfatórios em relação as estimativas de volume por parcela, todos tiveram certa tendência em superestimar a variável resposta, de forma mais branda no procedimento de redes e mais acentuada via regressão. Entretanto, tal fato não reflete em prejuízos na aplicação e adoção dessas práticas no meio florestal. Ressalta-se apenas que o desempenho das redes neurais artificiais foi superior à regressão.

Considerações Finais

O teste qui-quadrado demonstrou que os procedimentos de regressão e RNAs associados ao número de árvores por parcela e abertura de dossel são precisos nas estimativas de volume. Também, ambos os procedimentos de ajuste não diferiram entre si.

Embora ambas as técnicas sejam consistentes notou-se que o desempenho das RNAs na predição volumétrica apresentou os critérios estatísticos analisados superiores aos advindos da equação de regressão.

O desempenho das redes neurais artificiais foi superior ao método de regressão, pois apresentou boa adequação aos dados de entrada.

A utilização do número de árvores por parcela e abertura de dossel, independente da técnica utilizada (regressão ou RNA), reduz os custos dos inventários florestais, pois são de fácil obtenção.

Por fim, amparado nos resultados obtidos recomenda-se a utilização da técnica de RNA, associada as variáveis número de árvores por parcela e abertura de dossel para estimativa de volume por unidade de área.

Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior(CAPES), pela concessão da bolsa de doutorado.

LEAL, F. A.; MIGUEL, E. P.; MATRICARDI, E. A. T.; PEREIRA, R. S. Artificial neural networks in volume estimation at eucalyptus plantation using hemispherical photographs and number of trees. *Rev. Bras. Biom.*, São Paulo, v.33, n.2, p.234-250, 2015.

- *ABSTRACT: The objective of this work was to compare and evaluate the performance of artificial neural networks and regression method of the volume estimated in a plantation of Eucalyptus urophylla using as variables only the number of trees per plot and canopy openness. The areas located in the municipality of Rio Verde, Goiás, where they were released 23 plots of 400 m² in total area of 116 hectares. In each plot was counted the number of trees and was taken 3 hemispheric photos diagonally. The volume of each tree was obtained by the method Hohenadl. Adjustments were using simple linear regression and artificial neural networks with data from 18 plots, the other 5 were for validating. Although tested methods have been satisfactory in relation to the volume estimates, all data tendencies to overestimate the response variable, however the neural networks were always higher in performance even not differing significantly from the regression.*

▪ **KEYWORDS:** *Opening canopy; forest inventory; volumetric models.*

Referências

AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, Boston, v.19, n.6, p.716-723, 1974.

BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. *Revista Árvore*, Viçosa-MG, v.37, n.4, p.639-645, 2013a.

BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G. *NeuroForest Star*. 2013. Patente: Programa de Computador. Número do registro: 13410-5, data de registro: 30/04/2013, título: "NeuroForest Star", Instituição de registro: INPI - Instituto Nacional da Propriedade Industrial. 2013b.

BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G.; SILVA, A. Redução dos custos em inventário de povoamentos equiâneos utilizando redes neurais artificiais. *Agrária*, v.8, p.125-129, 2012a.

BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G.; SILVA, A.; SANTOS, A. C. A. Modelagem da distribuição diamétrica em povoamentos de eucalipto submetidos a desbaste utilizando autômatos celulares. *Revista Árvore*, v.36, n.5, p.931-939, 2012b.

BINOTI, M. L. M. S. *Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto*. 2010. 54f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, MG, 2010.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDEMIR, T. B. *Redes Neurais Artificiais: Teoria e Aplicações*. 2 ed. Rio de Janeiro: LTC, 2007. 260p.

CAMPOS, J. C. C.; LEITE, H. G. *Mensuração florestal: perguntas e respostas*. 4.ed., Viçosa: Editora UFV, 2013. 605p.

CARVALHO, A. P. L. F.; BRAGA, A. P.; LUDERMIR, T. B. *Fundamentos de redes neurais*. Publicação da COPPE/UFRJ para a XI Escola de Computação Brasileira, Rio de Janeiro, 1998.

CHAPMAN, L. Potential applications of near infrared hemispherical imagery in forest environments. *Agricultural and Forest Meteorology*. v.143, p.151-156, 2007.

CHAVES, R. A.; REIS, M. G. F.; REIS, G. G.; PEZZOPANE, J. E. M.; XAVIER, A.; MONTE, M. A. Dinâmica de cobertura de dossel de povoamentos de clone de *Eucalyptus grandis* W. Hill ex-Maiden submetidos a desrama artificial e desbaste. *R.Árvore*, Viçosa-MG, v.31, n.6, p.989-998, 2007.

DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. *Computers and Electronics in Agriculture*, v.10, n.3, p.235-244, 2005.

DRAPER, N.; SMITH, H. *Applied regression analysis*. John Wiley e Sons: New York, 706p. 3ª edição. 1998.

FINGER, C. A. G. *Fundamentos de Biometria Florestal*. Santa Maria: UFSM/CEPEF/FATEC. 1992. 269p.

FLORIANO, E. P.; MÜLLER, I.; FINGER, C. A. G.; SCHNEIDER, P. R. Ajuste e seleção de modelos tradicionais para série temporal de dados de altura de árvores. *Ciência Florestal*, Santa Maria, v.16, n.2, p.177-199. 2006.

FRAZER, G. W.; CANHAM, C. D.; LERTZMAN, K. P. *Manual do Gap Light Analyzer (GLA)*, Version 2.0: Imaging software to extract canopy structure and gap light transmission indices from true-color fisheye photographs, user's manual and program documentation. Simon Fraser University, Burnaby, British Columbia, and the Institute of Ecosystem Studies, Millbrook, New York. 1999. 36p.

HALE, S. E. The effect of thinning intensity on the below-canopy light environment in a Sitka spruce plantation. *Forest Ecology and Management*, v.179, p.341-349, 2003.

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

HEATON, J. *Programming Neural Networks with Encog3 in Java*. 2ed. Heaton Research, Incorporated, 2010. 240p.

KUAN, C. M.; WHITE, H. Artificial neural networks: An econometric perspective. *Econometric Reviews*, vol.13, p.1-91, 1994.

LEAL, F. A. *Técnicas de sensoriamento remoto e geoestatística aplicadas em um povoamento de Eucalyptus urophylla S. T. Blake para estimativa de volume*. Dissertação de mestrado em Ciências Florestais. Publicação PPGEFL.DM-190/2013. Departamento de Engenharia Florestal, Universidade de Brasília - UnB, Brasília/DF. 2013, 96p.

LEAL, F. A.; MIGUEL, E. P.; MATRICARDI, E. A. T. Mapeamento de unidades produtivas utilizando a interpolação geoespacial krigagem a partir do inventário florestal em um povoamento de *Eucalyptus urophylla* S. T. Blake. *Enciclopédia Biosfera*, v.7, n.13, p.727-745, 2011.

LEITE, H. G.; SILVA, M. L. M.; BINOTI, D. H. B.; FARDIN, L.; TAKIZAWA, F. H. Estimation of inside-bark diameter and heartwood diameter for *Tectonagrandis* Linn. trees using artificial neural networks. *European Journal of Forest Research*, v.130, n.2, p.263-269, 2010.

MACHADO, S. A.; FIGUEIREDO FILHO, A. *Dendrometria*. Sebastião do Amaral Machado, Afonso Figueiredo Filho. - 2ª.ed. - 1ª reimpressão - Guarapuava : UNICENTRO, 2009. 309p.

MELLO, J. M. *Geoestatística aplicada ao inventário florestal*. 122p. Tese (Doutorado em Recursos Florestais) Universidade de São Paulo campus Piracicaba (ESALQ), Piracicaba/SP, 2004.

MILTON, E. J. Low-cost ground-based digital infrared photography. *International Journal of Remote Sensing*. vol. 23, 1001-1007, 2002.

MONTE, M. A.; REIS, M. G. F.; REIS, G. G.; LEITE, H. G.; STOCKS, J. J. Métodos indiretos de estimação da cobertura de dossel em povoamentos de clone de eucalipto. *Pesquisa Agropecuária Brasileira*, Brasília, v.42, n.6, p.769-775, 2007.

OSMOND, P. *Application of near-infrared hemispherical photography to estimate leaf area index of urban vegetation*. The seventh International Conference on Urban Climate. 29.June-3.July 2009, Yokohama, Japan. Disponível em: <http://www.ide.titech.ac.jp/~icuc7/extended_abstracts/pdf/375978-1-090513173310-002.pdf>.

- OSÓRIO, F. S. A hybrid system for constructive machine learning. *Neurocomputing*, n.28, v.2, p.191-205, 1999.
- PRODAN, M. Holzmesslehre. *Frankfurt A. M* :Sauerländers Verlag, 1965. 644p.
- PORTUGAL, M. S. Neural networks versus time series methods: a forecasting exercise. *Revista Brasileira de Economia*, v. 49, n. 4, p. 611-629, 1995.
- RIEDMILLER, M.; BRAUN, H. A direct adaptive method for faster back propagation learning: the rprop algorithm. *IEEE International Conference on Neural Networks*, v.1, p.586-591, 1993.
- RUSSEL, S; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 3 ed. Editora Pearson, 2010, 1.152p.
- SCHUMACHER, F. X.; HALL, F. S. Logarithmic expression of timber-tree volume. *Journal of Agricultural Research*, v.47, n.9, p.719-734, 1933.
- SCOLFORO, J. R. S.; MELLO, J. M. *Inventário Florestal / José Roberto Soares Scolforo, José Márcio de Mello*. Lavras: UFLA/FAEPE, 2006.
- SCHWARZ, G. Estimating the dimensional of a model. *Annals Statistics*, Hayward, v.6, n.2, p.461-464, 1978.
- SILVA, E.; OLIVEIRA, A. C. *Dicas para Configuração de Redes Neurais*. Disponível em: <http://equipe.nce.ufrj.br/thome/grad/nn/mat_didatico/dicas_configuracao_rna.pdf>. Acesso 08 mar. 2015.
- SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Ajuste do modelo de Schumacher e Hall e aplicação de redes neurais artificiais para estimar volume de árvores de eucalipto. *R. Árvore*, Viçosa/MG, v.33, n.6, p.1133-1139, 2009a.
- SILVA, J. R.; MAILARD, F.; COSTA-PEREIRA, P. S.; ANDRADE, C. F. Comparação de métodos indiretos para medição de abertura do dossel no cerrado, utilizando observações obtidas nos trabalhos de campo e fotos hemisféricas digitais. *Anais... XIV Simpósio Brasileiro de Sensoriamento Remoto*, Natal, Brasil, INPE, p. 3059-3064. 2009b.
- SIQUEIRA NETO, M.; PICCOLO, M. C.; COSTA JUNIOR, C.; CERRI, C. C.; BERNOUX, M. Emissão de gases do efeito estufa em diferentes usos da terra no bioma Cerrado. *Revista Brasileira de Ciência do Solo*, v.35, p. 63-76, 2011.
- WANG, Y.S.; MILLER, D.R. Calibration of the hemispherical photography technique to measure leaf area index distributions in hardwood forests. *Forest Science*, v.33, p.210-216, 1987.

Recebido em 03.02.2015

Aprovado após revisão em 16.03.2015