

PREDIÇÃO RECURSIVA DE DIÂMETROS UTILIZANDO PERCEPTRON DE MÚLTIPLAS CAMADAS PARA O CÁLCULO DE VOLUME COMERCIAL DE CLONES DE EUCALIPTO

Fabrízzio A. A. M. N. Soares^{1*}, Edna L. Flores², Christian D. Cabacinha³,
Antônio C. P. Veiga², Wanderley S. Alencar¹ e Marcelo S. Freitas¹

1: Coordenação de Ciência da Computação
Campus Jataí
Universidade Federal de Goiás
Jataí - GO, Brasil
e-mail: fabrizzio.soares@gmail.com,
wanderleyalencar@hotmail.com,
msfreitas@hotmail.com

2: Faculdade de Engenharia Elétrica
Universidade Federal de Uberlândia
Uberlândia - MG, Brasil
e-mail: edna.gilberto,acpveiga@ufu.br

3: Departamento de Engenharia Florestal
Instituto de Ciências Agrárias
Universidade Federal de Minas Gerais
Montes Claros - MG, Brasil
e-mail: ccabacinha@yahoo.com.br

Palavras chave: Perceptron de Multiplas Camadas, Predição Recursiva, Clones de Eucalipto, Estimativa de diâmetro

Resumo. *Um método de predições de diâmetros de árvores clones de eucalypto com redes neurais artificiais Perceptron de Multiplas Camadas (MLP) foi apresentado neste trabalho. O objetivo foi otimizar o tempo e o custo de realização do processo de inventário florestal. As predições realizadas com o modelo utilizaram 3 diâmetros da base da árvore e os demais diâmetros foram estimados um a um de forma recursiva até atingir o diâmetro mínimo comercial de 4 cm. Com os diâmetros previstos pode-se utilizar o método de Smalian para o cálculo dos volumes comerciais das árvores. O treinamento da rede MLP foi realizado com apenas 10% do total de árvores do sítio plantado e foram testados com os 90% restante do total de árvores. Para avaliar o desempenho do modelo, os diâmetros foram comparados com os valores reais medidos. O modelo proposto foi satisfatório nos teste realizados não apresentando diferenças significativas quando comparados aos diâmetros estimados com os valores reais das árvores.*

1 INTRODUÇÃO

Em torno de 3 milhões de hectares, aproximadamente 62% das florestas plantadas no Brasil, são do gênero *Eucalyptus*. O grande investimento das indústrias de papel e celulose e os esforços de instituições de pesquisas e universidades, tornaram esse país um dos maiores produtores de madeira de eucalipto do mundo.

O eucalipto é uma espécie florestal de rápido crescimento, menor custo e produz muito mais madeira que outras espécies. A produtividade média de madeira no Brasil é de $45 - 50 m^3 ha^{-1} ano^{-1}$. No Chile, Estados Unidos, Canadá e Finlândia, esse número é de 20, 10, 7 e $4 m^3 ha^{-1} ano^{-1}$, respectivamente [32].

O volume de madeira plantado é uma informação essencial no direcionamento racional e sustentável dos recursos disponíveis nas florestas. Assim, é muito importante quantificá-lo da forma mais precisa possível. Para isto, são necessárias diversas medições em amostras representativas das florestas para se conhecer a dinâmica dos povoamentos.

As principais variáveis das árvores utilizadas nos levantamentos florestais são a altura total e os diâmetros medidos no tronco em diversas alturas em relação ao solo. Na realização de um inventário florestal para o levantamento desses dados podem ocorrer diversos erros. O erro de 1 metro na medição da altura total (H_t) da árvore corresponde a aproximadamente 14% do volume cilíndrico das árvores. Enquanto o erro de 1 cm no diâmetro na altura do peito (DAP) corresponde a um erro de aproximadamente 19% do volume total da árvore. O Diâmetro na Altura do Peito é o diâmetro a 1,30 metros do solo.

Defeitos nos instrumentos de medição e falhas dos profissionais que os operam podem ocasionar erros na medição da altura total e dos diâmetros. A medição dos diâmetros é realizada diretamente na árvore utilizando a suta ou similar. Devido às dimensões das árvores, as medições da altura total são realizadas indiretamente com a árvore em pé. Nas medições indiretas são utilizados instrumentos para visualizar a base e o topo das árvores e por princípios geométricos ou trigonométricos são calculadas essas alturas.

Para que o operador possa medir a altura total da árvore é necessário que o topo e a base estejam visíveis. Porém, a densidade do povoamento, a inclinação do solo ou da árvore e os movimentos dos ventos, podem prejudicar essa visualização.

As medições dos diâmetros nas árvores podem ser realizadas com a árvore em pé, quando essas medidas estão próximas ao DAP. Medidas acima do DAP poderão exigir o uso de escadas ou andaimes quando as alturas ainda forem acessíveis. Porém, para alturas muito grandes exigirá equipamentos especiais para escalada além de muita perícia do operador.

Para simplificar as medições acima do DAP pode ser realizada uma cubagem rigorosa que é o corte da árvore e medir os diâmetros e a altura total com ela deitada. Porém, esse método é destrutivo pois a madeira cortada deverá ser comercializada em curto prazo. Isto impede que os dados levantados auxiliem nas decisões a médio e longo prazo. Por esses motivos, geralmente as medidas de diâmetros acima do DAP são estimadas.

De posse dos diâmetros medidos ou estimados e da altura total são utilizadas as fórmulas de Smalian, Huber e de Newton para a obtenção do volume total da árvore [10][23][29].

A partir dos anos 90, muitos pesquisadores têm utilizado as redes neurais artificiais (RNAs) em aproximações de funções não-lineares. As RNAs não exigem qualquer suposição geométrica sobre a função a ser aproximada. Elas têm sido aplicadas em diversas áreas como finanças, previsão de séries temporais, classificação de padrões, entre outras.

As redes neurais têm sido utilizadas na modelagem florestal para estimar diversos parâmetros das árvores tais como diâmetro, altura, volume e outros.

Foram utilizadas as RNAs em [13] para modelar o crescimento e a mortalidade das árvores.

Em [19] as RNAs foram utilizadas para gerar uma distribuição de probabilidade de DAP's para classificar as árvores. Como entrada foram utilizados idade, altura média das árvores dominantes, número de árvores por acre e identificador de sítio.

As RNAs foram utilizadas em [17] para encontrar uma distribuição de frequência de classes de diâmetros do tronco a partir dos diâmetros máximo relativo e médio e do coeficiente de variação.

Em [33][2][27][31] as alturas das árvores foram previstas a partir do DAP, idade e outras estatísticas usando redes neurais artificiais.

A previsão de diâmetros foi realizada para pinheiros por [5] utilizando redes neurais artificiais. Foram utilizados como entradas os diâmetros do toco e da altura do peito e a altura total das árvores. As saídas da rede foram todos os diâmetros entre 5.3 a 33.3 em intervalos 4 em 4 metros. Os resultados obtidos foram satisfatórios, pois as árvores fir são bastante regulares. Apesar disso, foi necessário utilizar em torno de 90% das amostras para treino, restando apenas 10% para teste.

Em [6][7][11][1][12][30][25][8] as redes neurais foram utilizadas para estimar os volumes de árvores. Nesses trabalhos, as entradas utilizadas foram os DAPs e as alturas totais das árvores. Um inconveniente nesses trabalhos é que árvores com o mesmo DAP e a mesma altura total, mas com características de afilamento diferentes, resultaram em volumes iguais.

A proposta deste trabalho é utilizar a rede neural artificial Perceptron de Múltiplas Camadas (MLP) na predição de diâmetros de árvores. Este modelo usa como entrada 3 medidas de diâmetro na base da árvore e de forma recursiva acompanha seu afilamento para prever as demais medidas de diâmetros. A contribuição desse trabalho é a estimativa dessas medidas com apenas 3 medidas da base e sem o conhecimento da altura total da árvore. Além disso, o modelo proposto necessita de apenas 10% das árvores como amostra para seu treinamento.

2 MÉTODOS DE PREDIÇÃO E REDES NEURAIIS

2.1 PREDIÇÃO

A predição de séries é basicamente um problema de modelagem. Nesta predição, um modelo é construído entre as entradas e as saídas. Este modelo utiliza valores prévios da série para prever os próximos valores. A predição direta e a recursiva são alguns dos métodos aplicados na predição de séries muito longas[18].

2.1.1 PREDIÇÃO DIRETA

Para prever os valores de uma série são construídos $M + 1$ modelos diferentes, conforme a Equação (1).

$$\hat{d}_{i+m} = f_m(d_{i-1}, d_{i-2}, \dots, d_{i-n}), \quad (1)$$

onde $m = 0, 1, \dots, M$; M é o horizonte máximo da predição, d e \hat{d} são os parâmetros de entrada e os preditos, respectivamente, i é o índice do parâmetro e n é o tamanho do regressor. As variáveis de entrada do lado direito da Equação (1) definem o regressor.

2.1.2 PREDIÇÃO RECURSIVA

Na predição recursiva os modelos são construídos utilizando valores prévios de parâmetros para estimar o parâmetro um passo a frente. A Equação (2) mostra esse modelo.

$$\hat{d}_i = f(d_{i-1}, d_{i-2}, \dots, d_{i-n}), \quad (2)$$

onde d e \hat{d} são os parâmetros de entrada e os preditos, respectivamente, i é o índice do parâmetro e n é o tamanho do regressor. As variáveis de entrada do lado direito da Equação (2) definem o regressor.

Para prever o próximo parâmetro, a predição recursiva utiliza novamente a Equação (2), porém o primeiro parâmetro de entrada é o resultado predito no passo anterior. Os demais parâmetros são deslocados uma entrada à direita e desconsidera-se o último parâmetro usado na predição anterior preservando assim o número de entradas. Esse modelo é mostrado na Equação (3).

$$\hat{d}_{i+1} = f(\hat{d}_i, d_{i-1}, d_{i-2}, \dots, d_{i-n+1}). \quad (3)$$

onde d e \hat{d} são os parâmetros de entrada e os preditos, respectivamente, i é o índice do parâmetro e n é o tamanho do regressor.

Neste trabalho são realizadas as predições de diâmetros de árvores da espécie *eucalyptus* de forma recursiva usando redes neurais artificiais. Assim, na predição são adotados como entrada os diâmetros $d_{1,3}$, $d_{0,7}$ e $d_{0,3}$ e na saída da rede é estimado o diâmetro $\hat{d}_{2,0}$. No próximo passo as entradas são $\hat{d}_{2,0}$, $d_{1,3}$ e $d_{0,7}$ e na saída da rede é estimado o diâmetro

$\hat{d}_{3,0}$. Esses passos são repetidos de metro em metro sucessivamente até que o diâmetro estimado seja menor ou igual a 4 cm.

2.2 REDE NEURAL ARTIFICIAL PERCEPTRON DE MÚLTIPLAS CAMADAS (MLP)

A rede neural Perceptron de Multiplas Camadas (MLP) é composta de uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Suas camadas são compostas de neurônios com várias entradas e uma saída. Cada entrada possui um peso associado e cada saída possui uma função de ativação. Os sinais de entradas dos neurônios são propagados para frente pela rede, camada por camada[16]. A Figura 1 mostra a MLP utilizada neste trabalho.

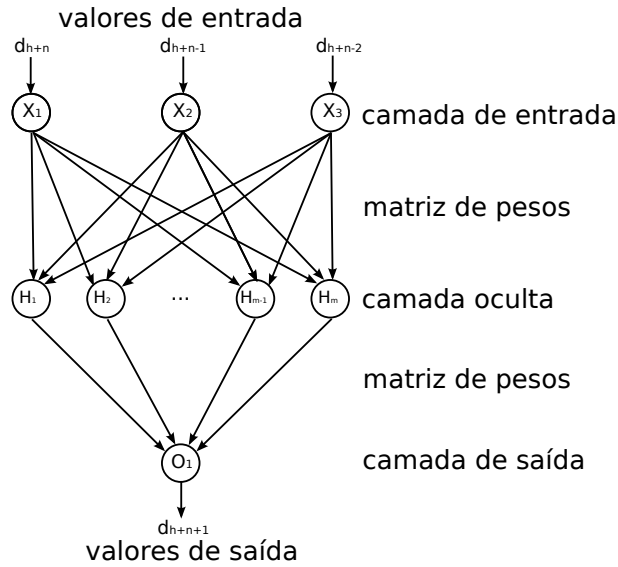


Figura 1: MLP com aArquitetura de três camadas

O número de neurônios da camada oculta das redes MLPs foi obtido pela Equação (4)[9].

$$N = \text{integer} \left(\frac{n}{2} + \sqrt{R} \right) \quad (4)$$

onde n é o número de neurônios de entrada e R é o tamanho do conjunto de dados.

Cada neurônio possui um conjunto de entradas. A finalidade da camada de entrada é repassar os valores recebidos para os neurônios da camada oculta. Os neurônios das camadas ocultas possuem a quantidade de entradas igual ao número de saídas da camada anterior. O número de neurônios de saída da rede é a quantidade de parâmetros que a RNA irá estimar.

O valor de saída de cada neurônio é calculado conforme a Equação (5)

$$a_j = w_0 + \sum_{i=1}^n x_i w_i \quad (5)$$

onde a_j é a saída do neurônio j , n é a quantidade de entradas, x_i e w_i são o valor e o peso da entrada i e w_0 é o bias do neurônio.

Diversas funções de ativação podem ser utilizadas nas redes MLPs [16]. Neste trabalho, a função de ativação utilizada nos neurônios das camadas oculta e de saída foi a tangente hiperbólica, também chamada de tan-sigmoide ou tansig. A Equação (6) mostra o cálculo dessa função.

$$\varphi_j(a_j) = \tanh(a_j) = \frac{2}{(1 + \exp(-2a_j))} - 1 \quad (6)$$

onde a_j é o valor de saída do neurônio j .

O treinamento da MLP é dividido em três etapas: propagação das entradas dos padrões de treino, cálculo do erro e retro-propagação desse erro para ajuste dos pesos dos neurônios. Neste trabalho foram utilizados 10% das árvores para treinamento da rede e os 90% restantes foram utilizados para teste.

Diversos algoritmos podem ser utilizados para o treinamento da MLP. Dentre eles pode-se destacar o gradiente descendente, o gradiente descendente com momentum e o gradiente conjugado[14], o quasi-Newton [21], o Levenberg-Marquardt[15] e outros. Neste trabalho é utilizado o algoritmo de treinamento Levenberg-Marquardt.

A técnica de Levenberg-Marquardt é conhecida por ser mais eficiente do que a técnica do gradiente descendente [15]. Este gradiente é um algoritmo de maior descida e envolve movimentos de pequenos passos abaixo do gradiente local do campo escalar. Uma vantagem desse método é a possibilidade do gradiente descendente encontrar um mínimo local antes que o mínimo global seja atingido. O algoritmo Levenberg-Marquardt é um aperfeiçoamento do método de Gauss-Newton, que é uma variante do método de Newton [20][24]. O método de Newton utiliza a informação da derivada parcial de segunda ordem do índice de desempenho utilizado para corrigir os pesos. Isso permite que além da informação do gradiente, seja utilizada a informação sobre a curvatura da superfície do erro.

2.3 CÁLCULO DOS ERROS DOS DIÂMETROS PREDITOS COM AS MLPS

Com a finalidade de verificar o comportamento das MLPs na predição de todos os diâmetros de cada árvore foram utilizados o root mean square error percentual ($RMSE_{\%}$) e a correlação linear (R), conforme as Equações (7) e (8), respectivamente.

$$RMSE_{\%} = \frac{1}{\bar{Y}} \left(\left(\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right) \frac{1}{n} \right)^{1/2} \times 100 \quad (7)$$

$$R = \frac{n \left(\sum_{i=1}^n \hat{Y}_i Y_i \right) - \left(\sum_{i=1}^n \hat{Y}_i \right) \left(\sum_{i=1}^n Y_i \right)}{\left(n \left(\sum_{i=1}^n Y^2 \right) - \left(\sum_{i=1}^n Y \right)^2 \right)^{1/2}} \quad (8)$$

onde \bar{Y} é a média dos diâmetros, \hat{Y}_i e Y_i são os diâmetros preditos e medidos, respectivamente e n é o número de medidas de diâmetros entre as medidas de entrada e a altura total.

3 RESULTADOS E DISCUSSÃO

Neste trabalho foram realizados vários testes para encontrar a estrutura da MLP de cada classe que mais se adequasse à predição de diâmetros das árvores. Inicialmente adotou-se 5 neurônios. Este valor foi incrementado de 5 em 5 para cada classe até serem atingidos os valores sugeridos pela Equação (4). A quantidade de neurônios calculada utilizando-se essa equação para cada classe sugeriu uma quantidade de neurônios muito grande. Porém, em nenhum das classes houveram melhorias significativas nos valores de $RMSE\%$ e R nas MLPs com mais de 20 neurônios na camada oculta.

A estrutura adotada das MLPs, como ilustrado na Figura (1), possui uma camada de entrada com três variáveis que são medidas de diâmetros em alturas sucessivas, uma camada oculta com 20 neurônios e uma camada de saída com um neurônio que é o valor do diâmetro a ser estimado.

3.1 DADOS

Os dados utilizados neste estudo foram levantados em um plantio localizado no município de Aracruz, na região metropolitana do estado do Espírito Santo, no Brasil. Foram usadas 615 árvores, do gênero *Eucalyptus*, obtidas de um mesmo material genético clonal. Todas elas com a mesma idade, 6 anos e seis meses, plantadas com espaçamento $3 \times 3 m$, distribuídas em 16 parcelas circulares de $360 m^2$.

As árvores foram abatidas e cubadas rigorosamente pelo método Smalian [3]. Durante a cubagem foram medidos, com Suta, os diâmetros nas alturas 0, 10; 0, 30; 0, 50; 0, 70; 0, 90, 1, 10; 1, 30; 2, 00 m em relação ao solo e no restante do tronco em intervalos de 1 em 1 m. O diâmetro no topo da árvore é considerado igual a zero cm. Os diâmetros nas alturas $h_{0.3}$ e $h_{1.3}$ são conhecidos como diâmetro do toco e diâmetro na altura do peito (DAP), respectivamente. A Figura 2 ilustra o procedimento de como as árvores foram medidas.

Diversos estudos afirmam que as características das árvores variam de acordo com o DAP. Devido a isso, vários trabalhos têm separado as árvores em classes diamétricas baseadas no DAP em amplitudes de 4 e 5 cm [28][26]. Neste trabalho, as árvores foram separadas de acordo com a medida do DAP em classes com amplitudes de 3 em 3 cm. A classe 6 possuía apenas uma árvore. Devido a isso esta árvore foi incorporada à classe 5.

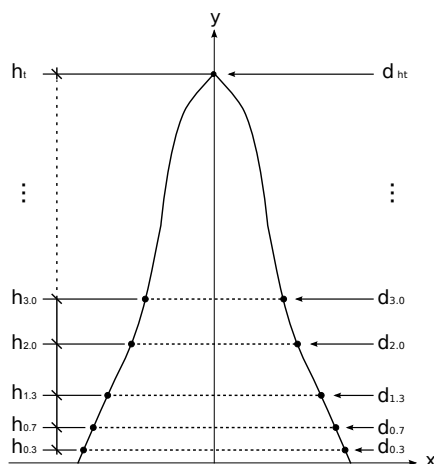


Figura 2: Alturas e diâmetros medidos nas árvores

E com isto o número total de classes utilizadas nos experimentos do algoritmo proposto foi de 5 classes.

A Tabela 1 mostra a quantidade das árvores, suas estatísticas de DAPs e alturas para as classes de 1 a 5 utilizadas neste trabalho.

Foi criada uma MLP para treino e predição para cada uma das classes de 1 a 5. O modelo de MLP proposto neste trabalho baseou-se na relação entre os 3 diâmetros vizinhos utilizados como entrada e o próximo diâmetro a ser previsto pela rede. Para verificar essa relação foi calculado o coeficiente de autocorrelação entre os diâmetros de cada árvore, utilizando a Equação (8). O cálculo desse coeficiente foi realizado com deslocamentos de 1 a 3 para cada árvore. A Tabela 1 mostra o menor e o maior coeficientes de autocorrelação para 1 deslocamento em cada classe.

Tabela 1: Estatísticas descritivas das árvores do gênero *Eucaplyptus*

Classe	n	Autocorrelação(d_i)		Diâmetro na Altura do Peito - DAP(cm)			Altura Total (H_t) (m)						
		Min	Max	Média	Var	DP	Min	Max	Média	Var	DP	Min	Max
1	14	0,63	0,84	8,68	0,95	0,97	7,15	10,10	17,60	3,49	1,87	14,20	19,70
2	106	0,77	0,87	12,01	0,54	0,73	10,32	13,14	22,83	2,75	1,66	18,00	26,70
3	302	0,78	0,89	14,62	0,68	0,82	13,15	16,14	25,74	1,57	1,25	18,40	29,40
4	178	0,83	0,90	17,23	0,64	0,80	16,15	19,14	27,52	1,39	1,18	24,20	30,10
5	15	0,82	0,90	20,21	1,7	1,31	19,31	24,55	29,23	1,06	1,03	27,20	30,60

Os coeficientes de autocorrelação encontrados para 1 deslocamento apresentaram valores bastante altos, como mostrado na Tabela 1. Pode-se verificar nessa tabela que o menor e o maior coeficientes de autocorrelação foram de 0,63 e 0,90, respectivamente. Os menores e os maiores valores encontrados de coeficiente de autocorrelação para 3 deslocamentos foram de 0,55 e 0,81, respectivamente. Estes resultados mostram que os diâmetros de cada árvore são fortemente correlacionados.

Para uma maior confiabilidade e rapidez no treinamento e predição com as MLPs, os dados dos diâmetros das árvores são normalizados para o intervalo de $[-1, 1]$, usando a

Equação (9)[34]. Neste intervalo o -1 indica os diâmetros iguais a 0. O valor 1 representa o maior diâmetro medido de todas as árvores. E entre esses valores encontram-se os demais diâmetros. A normalização é feita para todas as árvores independente de sua classe.

$$X_{NORM_i} = (b - a) \cdot \left(\frac{X_i - \min(X)}{\max(X) - \min(X)} \right) + a \quad (9)$$

onde X são os diâmetros de todas as árvores, $\min(X)$ e $\max(X)$ são o menor e o maior diâmetro das árvores, respectivamente, X_i são os valores dos diâmetros da árvore i , X_{NORM_i} são os valores dos diâmetros normalizados da árvore i e a e b são os intervalos de normalização com valores -1 e 1 , respectivamente.

Após as MLPs preverem os diâmetros das árvores, estes diâmetros são denormalizados usando os valores de $\min(X)$ e $\max(X)$ utilizados na normalização.

O algoritmo de treino utilizado nas MLPs foi o Levenberg-Marquardt [24] com no máximo 1.000 épocas. Todos os treinos das MLPs foram realizados com pesos iniciais aleatórios e não foram usados parâmetros adaptativos de taxa de aprendizado e momentum.

Neste trabalho foi utilizada uma adaptação do método *early stopping* para evitar o *overfit* [4][9][22]. Nesta adaptação adotou-se a seguinte heurística: estabelecer um erro médio quadrático (*MSE*) mínimo de 1×10^{-10} e treinar as MLPs até que por 5 épocas seguidas o *MSE* de cada MLP não apresente redução significativa ou aumente. Nos testes realizados, os treinamentos das MLPs convergiram abaixo de 150 épocas. A Tabela 2 mostra os valores de *MSE* e o número médio de épocas necessários para a convergência em cada uma das classes de 1 a 5 das árvores.

Tabela 2: Amostras de treino

Classe	Total		Amostras		Treino	
	Árvores	Conjuntos	Árvores	Conjuntos	Épocas	MSE
1	14	231	10	155	35	$3,0 \times 10^{-4}$
2	106	2405	11	239	45	$3,0 \times 10^{-4}$
3	302	8025	31	792	120	$3,5 \times 10^{-4}$
4	178	5105	18	498	110	$3,5 \times 10^{-4}$
5	15	462	10	298	130	$5,0 \times 10^{-4}$
Total	615	16228	80	1982	-	-

Para o treino das MLPs, em cada classe foram selecionadas aleatoriamente 10% das árvores ou no mínimo 10 árvores para a geração dos conjuntos de amostras de entrada e saída. Os 90% restantes das árvores foram utilizados para treino. Para cada uma das árvores selecionadas para treino foram geradas as amostras de entradas $[d_{i-2}, d_{i-1}, d_i]$ e a saída $[d_{i+1}]$. Por exemplo, no treinamento da rede, o par de entrada e saída foi $[d_{0.3}, d_{0.7}, d_{1.3}]$ e $[d_{2.0}]$, respectivamente, onde $d_{0.3}$, $d_{0.7}$, $d_{1.3}$, $d_{2.0}$ representam os diâmetros nas distância do solo até as medidas 0.3, 0.7, 1.3 e 2.0 metros. O próximo par de entrada e saída foi $[d_{0.7}, d_{1.3}, d_{2.0}]$ e $[d_{3.0}]$, respectivamente, onde $d_{3.0}$ representa o diâmetro na distância do solo até a medida de 3.0 metros, e assim sucessivamente.

Na etapa de teste das MLPs são utilizadas apenas as medidas iniciais de diâmetro $d_{0.3}$, $d_{0.7}$ e $d_{1.3}$ para prever as demais medidas de diâmetros de cada árvore. Essa etapa é

Tabela 3: Definição formal do algoritmo de predição

X	Vetor de entradas da MLP
X_1	Entrada 1 da MLP
X_2	Entrada 2 da MLP
X_3	Entrada 3 da MLP
i	A altura que está sendo predita
$d_{0.3}$	Diâmetro na altura 0,3 m (Diâmetro do toco)
$d_{0.7}$	Diâmetro na altura 0,7 m
$d_{1.3}$	Diâmetro na altura 1,3 m (Diâmetro na Altura do Peito - DAP)
D	Conjunto de diâmetros preditos O diâmetro na ponta da árvore é considerada como zero
D_i	Diâmetro na altura i
D_{min}	Diâmetro mínimo comercial, 4.0 cm
\hat{H}_t	Altura total estimada da árvore

Tabela 4: O algoritmo de predição

function PREDICTTREE($d_{0.3}, d_{0.7}, d_{1.3}, D_{min}$)	
$X \leftarrow [d_{0.3}, d_{0.7}, d_{1.3}]$	
$D \leftarrow []$	
$i \leftarrow 1.0$	
repeat	
$i \leftarrow i + 1.0$	
$D_i \leftarrow \text{NET.PREDICTDIAMET}(X)$	
$X_1 \leftarrow X_2$	
$X_2 \leftarrow X_3$	
$X_3 \leftarrow D_i$	
until ($D_i \leq D_{min}$)	
$\hat{H}_t \leftarrow i$	
$D_{\hat{H}_t} \leftarrow 0$	
return [D, \hat{H}_t]	
end function	

realizada para cada árvore individualmente, conforme o procedimento de predição recursivo descrito na Equação (3). Portanto, essas medidas iniciais são utilizadas para prever o diâmetro $\hat{d}_{2.0}$. Em seguida, as medidas $d_{0.7}$, $d_{1.3}$ e a $\hat{d}_{2.0}$ estimada são utilizadas para prever a medida $\hat{d}_{3.0}$, e assim sucessivamente até que o diâmetro previsto seja menor ou igual a 4 cm. A altura total da árvore é assumida como sendo a altura do último diâmetro predito. As Tabelas 3 e 4 ilustram as definições utilizadas e o algoritmo de predição com

base na Equação (3), respectivamente.

3.2 PREDIÇÃO DE DIÂMETROS DAS ÁRVORES

A Tabela 5 mostra de forma resumida os valores de $RMSE_{\%}$ e R dos testes realizados para a predição dos diâmetros das árvores para as classes de 1 a 5.

Tabela 5: Erros de predição de diâmetros por classe - $RMSE_{\%}$ (Menor, Maior, Médio e Desvio Padrão) e Correlação - R (Menor, Maior)

Classe	Árvores	RMSE%				Correlação (R)	
		Min	Max	Media	DP	Min	Max
1	4	0,99	20,57	7,17	9,13	0,96	0,99
2	95	0,52	7,50	2,69	1,49	0,97	0,99
3	271	0,41	6,14	2,20	1,07	0,98	0,99
4	160	0,43	6,01	1,87	0,89	0,99	0,99
5	5	1,27	2,02	1,74	0,30	0,99	0,99

Pode-se observar na Tabela 5 que nas classes 1 e 5, onde a quantidade de árvores é muito pequena, o maior valor de $RMSE_{\%}$ é aproximadamente 21% e 2%, respectivamente. Verifica-se também que nessas classes a média dos valores de $RMSE_{\%}$ ficaram em torno de 7.17% para a classe 1 e 1.74% para a classe 5. E os desvios padrões das classes 1 e 5 foram 9.1% e 0.29%, respectivamente. É importante salientar que na classe 5, embora a quantidade de árvores seja tão pequena quanto na classe 1, os valores de $RMSE_{\%}$ foram bem menores. Como pode ser verificado na Tabela 2, isto ocorre porque as árvores da classe 5 são bem mais altas do que as da classe 1 e conseqüentemente fornecem mais amostras para treino das MLPs.

Verifica-se ainda na Tabela 5 que nas classes 2, 3 e 4, onde a quantidade de árvores é bem maior do que nas classes 1 e 5, o maior valor de $RMSE_{\%}$ é aproximadamente 7,5%. Nessas classes as médias dos valores de $RMSE_{\%}$ ficaram entre 1.87% e 2.69% e os desvios padrões foram entre 0.88% e 1.49%. Esses valores mostram que os erros de predição das MLPs nas classes de 2 a 5 foram significativamente menores do que na classe 1 e portanto as MLPs para essas classes foram mais precisas.

A Tabela 5 mostra que o R ficou no intervalo de 0.9564 a 0.9991 mostrando uma fortíssima correlação linear entre os diâmetros reais medidos e os preditos pelas MLPs.

As Figuras de 3(a) a 3(f) ilustram alguns exemplos de valores de diâmetros reais e estimados pelas MLPs. E as Figuras de 4(a) a 4(c) mostram os histogramas das classes exemplificadas dos valores de $RMSE_{\%}$ calculados para cada árvore.

As Figuras 3(a), 3(b) e 3(c) mostram para as classes 2, 3 e 4, árvores com variação de diâmetros bastante regulares. Pode-se observar nessas figuras que os diâmetros reais e preditos foram bastante coincidentes, conseqüentemente seus valores de $RMSE_{\%}$ foram muito pequenos.

Árvores das classes 2, 3 e 4 com variação de diâmetros muito irregulares são ilustradas nas Figuras 3(d), 3(e) e 3(f). Verifica-se nessas figuras que os diâmetros preditos pelas

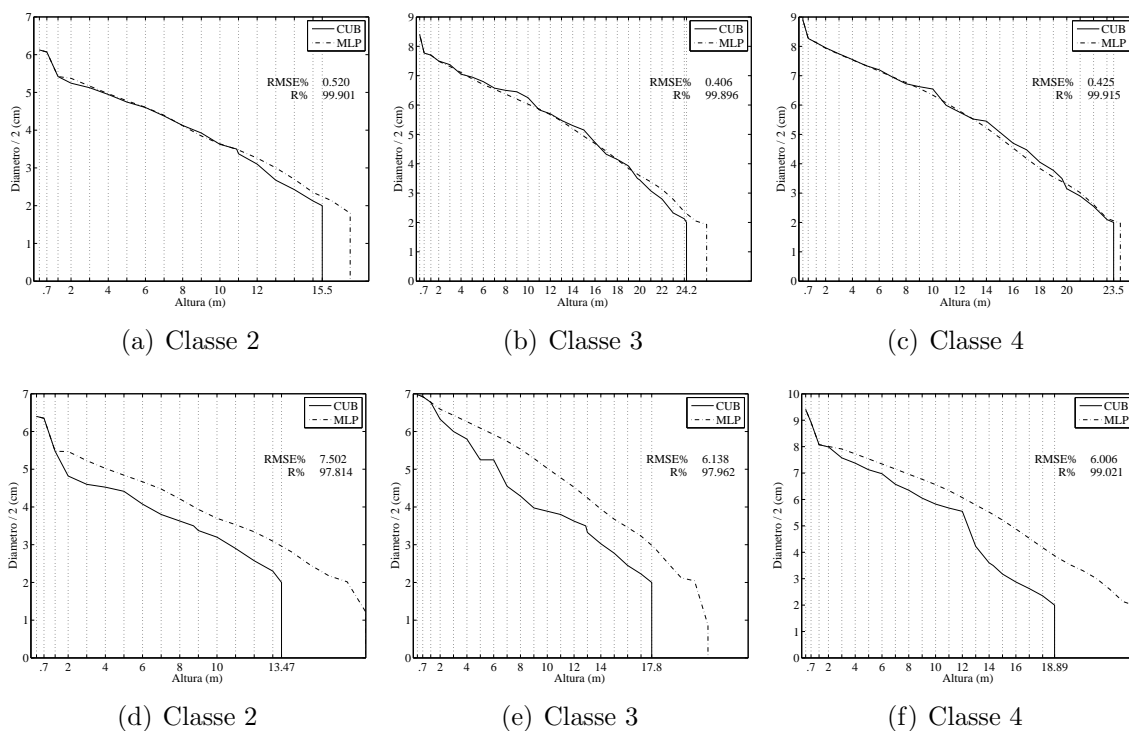


Figura 3: Exemplos de predição de Árvores com diâmetros regulares e irregulares.

MLPs superestimaram os reais, e conseqüentemente seus valores de $RMSE\%$ foram mais altos.

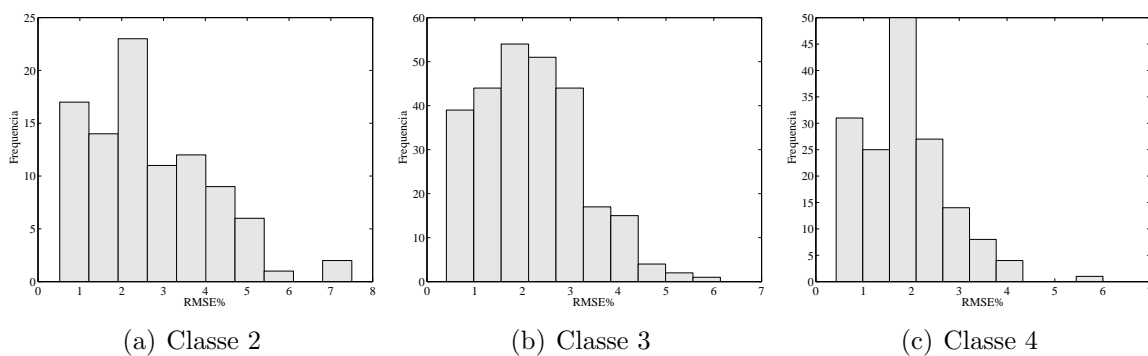


Figura 4: Histograma de $RMSE\%$ entre os diâmetros estimado e real.

As Figuras 4(a), 4(b) e 4(c) mostram os histogramas de $RMSE\%$ entre os diâmetros estimado e real para as classes de 2 a 4. Observa-se nessas figuras que a maior parte dos valores de $RMSE\%$ da classe 2 estão abaixo de 5% e nas classes 3 e 4 a maior parte dos valores de $RMSE\%$ estão abaixo de 3%. Estes valores de $RMSE\%$ dessas classes mostram que o modelo proposto estimou de maneira satisfatória os diâmetros das árvores. Como a

quantidade de árvores de teste nas classes 1 e 5 é muito pequena não existe um acúmulo de $RMSE\%$ em torno da média e portanto não foram mostrados seus respectivos histogramas.

Tabela 6: Diferenças entre o diâmetro mínimo esperado (4 cm) e os menores diâmetros estimados

Classe	Min	Max	Média
1	0,02	2,14	0,65
2	0,00	1,57	0,78
3	0,00	3,03	1,28
4	0,02	3,66	1,86
5	0,16	2,95	1,67

A Tabela 6 apresenta a menor, maior e a média da diferença entre o diâmetro mínimo adotado de 4 cm e os menores diâmetros estimados para cada árvore. Verifica-se que a menor e a maior diferença é inferior a 0.2 cm e 3.7 cm, respectivamente. Pode-se verificar também que as médias estão abaixo 1.9. Embora, em algumas classes a diferença tenha sido alta, esses valores são muito pequenos, representam uma porção muito pequena das árvores e portanto, não influenciam no volume final de cada árvore.

4 CONCLUSÕES

Em todos os testes realizados pôde-se observar que independente da variação dos diâmetros em cada árvore ser regular ou não, os valores dos diâmetros preditos pelas MLPs foram bastante satisfatórios em relação aos valores dos diâmetros reais. Porém, as MLPs tiveram dificuldade para prever os diâmetros mais próximos do topo das árvores. Isso já era esperado pois o modelo proposto neste trabalho realiza predições recursivas e erros em um determinado diâmetro são repassados para predição dos próximos diâmetros.

Nos testes realizados verificou-se que as diferenças entre os diâmetros estimados pelas MLPs e os diâmetros reais foram muito pequenas, portanto, os resultados obtidos com o método proposto foram bastante satisfatórios.

Todos os experimentos realizados foram utilizados para calcular volumes para fins de inventário florestal e os resultados serão publicados em revistas especializadas na área de ciências agrárias.

REFERÊNCIAS

- [1] Gleimar Botelho Baleeiro. Modelagem neuro-fuzzy na estimativa de volume de *eucalyptus sp.* Master's thesis, Universidade Federal de Lavras, Lavras - MG, 2007.
- [2] Fillipe Gustavo Brandão. Estimativa da altura total de *eucalyptus sp.* utilizando lógica fuzzy e neuro fuzzy. Master's thesis, Universidade Federal de Lavras, 2007.
- [3] Christian Dias Cabacinha. Um método para a realização do inventário florestal suprimindo a cubagem rigorosa. Master's thesis, Universidade Federal de Lavras, 2003.

- [4] Rich Caruana, Steve Lawrence, and Lee Giles. Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In *Advances in Neural Information Processing Systems 13*, 2001.
- [5] Maria J. Diamantopoulou. Artificial neural networks as an alternative tool in pine bark volume estimation. *Computers and Electronics in Agriculture*, 48(3):235–244, September 2005.
- [6] Maria J. Diamantopoulou. Predicting fir trees stem diameters using artificial neural network models. *Southern African Forestry Journal*, 205(1):39–44, November 2005.
- [7] Maria J. Diamantopoulou. Tree-bole volume estimation on standing pine trees using cascade correlation artificial neural network models. *Agricultural Engineering International: The CIGR EJournal*, VIII, 2006.
- [8] Maria J. Diamantopoulou and Elias Milios. Modelling total volume of dominant pine trees in reforestations via multivariate analysis and artificial neural network models. *biosystems engineering*, 105(3):306–315, March 2010.
- [9] Chi Dung Doan and Shie-Yui Liong. Generalization for multilayer neural network bayesian regularization or early stopping. In *APHW2004 Conference*, 2004.
- [10] César Augusto Guimarães Finger. *Fundamentos de Biometria Florestal*. UFSM/CEPEF/FATEC, Santa Maria, 1992.
- [11] Eric Bastos Görgens. Estimaco do volume de rvores utilizando redes neurais artificiais. Master’s thesis, Universidade Federal de Viosa, 2006.
- [12] Eric Bastos Görgens, Helio Garcia Leite, Heleno do Nascimento, and Jos Marinaldo Gleriani. Estimaco do volume de rvores utilizando redes neurais artificiais. *Revista rvore*, 33(6), December 2009.
- [13] Biing T. Guan and Geore Gertner. Modeling red pine tree survival with an artificial neural network. *Forest Science*, 37(5):1429–1440, November 1991.
- [14] M.T. Hagan, H.B. Demuth, and M.H. Beale. *Neural Network Design*. PWS Publishing Boston, 1996.
- [15] MT Hagan and MB Menhaj. Training feed forward networks with the marquardt algorithm. In *IEEE Transactions on Neural Networks*, volume 6, pages 861–867. IEEE, 1994.
- [16] Simon Haykin. *Neural Networks - A comprehensive Foundation*. Prentice Hall: Upper Saddle River, 2nd edition, 1998.

- [17] Jiarong Huang, Junhui Zhao, Guanggin Gao, Xianyu Meng, and Yuxiu Guan. Prediction of stand diameter distribution with artificial neural network. In *2009 Fifth International Conference on Natural Computation*, 2009.
- [18] Yongnan Ji, Jin Hao, Nima Reyhani, and Amaury Lendasse. Direct and recursive prediction of time series using mutual information selection. In *IWANN 2005, Lecture Notes in Computer Science*, volume 3512, pages 1010–1017. Springer Verlag, 2005.
- [19] Daniel J. Leduc, Thomas G. Matney, Keith L. Belli, and V. Clark Baldwin Jr. *Predicting diameter distributions of longleaf pine plantations [electronic resource] : a comparison between artificial neural networks and other accepted methodologies / Daniel J. Leduc ... [et al.]*. U.S. Dept. of Agriculture, Forest Service, Southern Research Station, Asheville, NC :, September 2001.
- [20] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly Journal of Applied Mathematics*, II(2):164–168, 1944.
- [21] Aristidis Likas and Andreas Stafylopatis. Training the random neural network using quasi-newton methods. *European Journal of Operational Research*, 126(2):331–339, October 2000.
- [22] Aleksander Lodwich, Yves Rangoni, and Thomas Breuel. Evaluation of robustness and performance of early stopping rules with multi layer perceptrons. In *2009 international joint conference on Neural Networks*, 2009.
- [23] S.A. Machado and A. Figueiredo Filho. *Dendrometria*. UFPR, Curitiba, 2003.
- [24] D. Marquardt. An algorithm for least squares estimation of non-linear parameters. *Journal of the society of Industrial and Applied Mathematics*, pages 431–441, 1963.
- [25] Ramazan Özçelik, Maria J. Diamantopoulou, John R. Brooks, and Harry V. Wiant Jr. Estimating tree bole volume using artificial neural network models for four species in turkey. *Journal of Environmental Management*, 91(3):742–753, 2010.
- [26] Leonardo Machado Pires and Natalino Calegario. Ajuste de modelos estocásticos lineares e não-lineares para a descrição do perfil longitudinal de árvores. *Revista Árvore*, 31(5):845–852, 2007.
- [27] K. J. Ranson, J. Masek, A. Fu, and D. Wang. Predicting tree height and biomass from glas data. In *10th International Symposium on Physical Measurements and Signatures in Remote Sensing - ISPMSRS'07*, 2007.
- [28] Paulo Renato Schneider, César Augusto Guimarães Finger, Jorge Euclides Mayer Klein, José Artemio Totti, and José Luiz Bazzo. Forma de tronco e sortimentos de madeira de *Eucalyptus grandis* maiden para o estado do rio grande do sul. *Ciência Florestal*, 6(1):79–88, 1996.

- [29] J.R.S. Scolforo. *Biometria florestal: Parte I: modelos de regressão linear e não-linear; Parte II: modelos para relação hipsométrica, volume, afilamento, e peso de matéria seca*. UFLA/FAEPE, Lavras, 2005.
- [30] Mayra Luiz Marques Silva, Daniel Henrique Breda Binoti, José Marinaldo Gleriani, and Helio Garcia Leite. Ajuste do modelo de schumacher e hall e aplicação de redes neurais artificiais para estimar volume de árvores de eucaplito. *Revista Árvore*, 33(6):1133–1139, 2009.
- [31] Ricardo M. A. Silva, Fillipe Gustavo Brandão, Geismar B. Baleeiro, Felipe L. Valentin, Adriano R. Mendonça, and Danilo M. Pires. Fuzzy and neuro-fuzzy estimates of the total height of eucalyptus trees. In *2008 ACM symposium on Applied computing*, pages 1772–1776, New York, NY, USA, 2008. ACM.
- [32] Votorantim Celulose e Papel - VCP. Sustainable growth with value creation and solid returns. Technical report, VCP, 2004.
- [33] Rongrui Xiao, Richard Carande, and Dennis Ghiglia. A neural network approach for tree height estimation using ifsar data. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5):345–348, 1998.
- [34] Cleber Zanchettin and Teresa Bernarda Ludermir. Sistemas neurais híbridos para reconhecimento de padrões em narízes artificiais. *SBA Controle & Automação*, 16(2), 2005.