

MAYRA LUIZA MARQUES DA SILVA BINOTI

**EMPREGO DE REDES NEURAS ARTIFICIAIS EM MENSURAÇÃO E
MANEJO FLORESTAL**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência Florestal, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2012

Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV

T

B614e
2012

Binoti, Mayra Luiza Marques da Silva, 1986-
Emprego de redes neurais artificiais em mensuração e
manejo florestal / Mayra Luiza Marques da Silva Binoti.
- Viçosa, MG, 2012.
vi, 130f. : il. ; 29cm.

Orientador: Helio Garcia Leite.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Florestas - Administração. 2. Eucalipto - Populações.
3. Levantamentos florestais. 4. Redes neurais (Computação).
5. Software - Desenvolvimento. I. Universidade Federal de
Viçosa. II. Título.

CDO adapt. CDD 634.9524

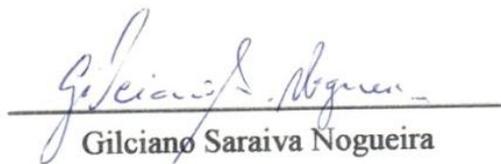
MAYRA LUIZA MARQUES DA SILVA BINOTI

**EMPREGO DE REDES NEURAIS ARTIFICIAIS EM MENSURAÇÃO E
MANEJO FLORESTAL**

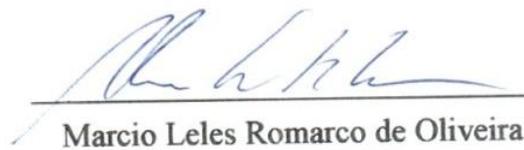
Tese apresentada à Universidade Federal
de Viçosa, como parte das exigências do
Programa de Pós-Graduação em Ciência
Florestal, para obtenção do título de
Doctor Scientiae.

APROVADA: 30 de julho de 2012.


José Marinaldo Gleriani
(Coorientador)


Gilciano Saraiva Nogueira


Agostinho Lopes de Souza


Marcio Leles Romarco de Oliveira


Helio Garcia Leite
(Orientador)

AGRADECIMENTOS

À Deus.

Ao Daniel pelo apoio incondicional na execução deste trabalho e em todos os momentos dessa nossa caminhada.

Ao professor Helio e à Silvana pela amizade, pelos conselhos, pelos preciosos exemplos de vida e profissionalismo.

Ao Departamento de Engenharia Florestal da Universidade Federal de Viçosa pela oportunidade de cursar o doutorado.

Ao Departamento de Engenharia Florestal da Universidade Federal dos Vales do Jequitinhonha e Mucuri pelo apoio nos momentos de ausência da universidade para resolver questões da tese.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro.

Aos professores Agostinho, Márcio, José Marinaldo e Gilciano pela participação na banca examinadora.

Aos bons amigos adquiridos durante a graduação, mestrado e doutorado.

BIOGRAFIA

Mayra Luiza Marques da Silva Binoti é natural de João Monlevade, Minas Gerais. Filha de Maurício de Oliveira da Silva e Neide Marques Pereira Oliveira da Silva, nascida em 29 de abril de 1986.

Iniciou o curso de engenharia florestal na Universidade Federal de Viçosa em 2004, tendo o concluído em 2008, com votos de louvor pelo desempenho acadêmico. No mesmo ano iniciou o mestrado, logo em seguida casou-se com Daniel Henrique Breda Binoti. Terminou o mestrado em fevereiro de 2010.

Em março de 2010, iniciou o Programa de Doutorado em Ciência Florestal, na Universidade Federal de Viçosa. Teve uma filha em março de 2012 e submeteu-se à defesa da tese em julho do mesmo ano.

SUMÁRIO

RESUMO.....	v
ABSTRACT	vi
INTRODUÇÃO.....	1
CAPÍTULO 1 - Aplicação de redes neurais artificiais para estimação da altura de árvores em povoamentos de eucalipto	5
CAPÍTULO 2 - Redução dos custos em inventário de povoamentos equiâneos com o emprego de redes neurais	19
CAPÍTULO 3 - Emprego de redes neurais artificiais para a estimação do volume de árvores	32
CAPÍTULO 4 - Projeção da distribuição diamétrica de povoamento equiâneos utilizando redes neurais artificiais.....	44
CAPÍTULO 5 - Modelagem da distribuição de diâmetros utilizando autômatos celulares e redes neurais artificiais.....	60
CAPÍTULO 6 - Modelagem da distribuição diamétrica de povoamentos eucalipto submetidos a desbaste utilizando autômatos celulares.....	79
CAPÍTULO 7 - Parametrização de redes neurais artificiais para estimação do volume de árvores	95
CAPÍTULO 8 - Estimação do volume de árvores utilizando o sistema <i>NeuroForest - volumetric</i>	107
CONCLUSÕES GERAIS	129

RESUMO

BINOTI, Mayra Luiza Marques da Silva, D.Sc., Universidade Federal de Viçosa, julho de 2012. **Emprego de Redes Neurais Artificiais em Mensuração e Manejo Florestal**. Orientador: Helio Garcia Leite. Coorientadores: Carlos Pedro Boechat Soares e José Marinaldo Gleriani.

O objetivo deste estudo foi treinar, aplicar e validar redes neurais artificiais para diferentes abordagens que auxiliam o conhecimento da produção volumétrica de povoamentos florestais, bem como desenvolver, aplicar e validar um sistema computacional específico para a obtenção de estimativas do volume de árvores. Os casos contemplados foram: estimação da altura, projeção dos parâmetros da função Weibull, cubagem de árvores e modelos de distribuição de diâmetros. Além disso, foram testadas diferentes configurações de redes, buscando a parametrização para a estimação do volume de árvores. Os dados compreenderam medições de parcelas permanentes de inventários florestais contínuos e dados de árvores abatidas para cubagem, em área de povoamentos de eucalipto. Para gerar as estimativas das variáveis de estudo foram utilizados o software *Statistica 7.0* e o sistema computacional *NeuroForest – Volumetric*. As aplicações de redes neurais artificiais apresentaram resultados satisfatórios comprovando a eficiência da técnica em mensuração e manejo florestal.

ABSTRACT

BINOTI, Mayra Luiza Marques da Silva, D.Sc., Universidade Federal de Viçosa, July, 2012. **Use of Artificial Neural Networks in Measurement and Forest Management**. Adviser: Helio Garcia Leite. Co-advisers: Carlos Pedro Boechat Soares and José Marinaldo Gleriani.

The aim of this study was to train, implement and validate artificial neural networks for different approaches that assist the knowledge of the volume of forest stands, as well as develop, implement and validate a computer system to obtain estimates of the volume of trees. The cases considered were: estimation of height, projection of Weibull function parameters, scaling tree and diameter distribution models. In addition, tested different network configurations, seeking parameterization for estimating the volume of trees. The data comprised measurements of permanent plots in continuous forest inventories and from trees felled for scalling in stands of eucalyptus. To generate estimates of the variables were used Statistica 7.0 and computer system NeuroForest - Volumetric. Applications of artificial neural networks showed satisfactory results proving the efficiency of the technique in measurement and forest management.

INTRODUÇÃO

O manejo de florestas equiâneas tem como objetivo principal obter produtos madeireiros, além de serviços ambientais de forma sustentável. Para explorar de maneira eficiente o recurso florestal madeireiro, é necessário conhecer o estoque em crescimento e de colheita através da medição e estimação de características das árvores e da área na qual elas estão plantadas.

Muitos métodos são utilizados para quantificar o volume de madeira atual e futuro em povoamentos florestais, dentre estes, o emprego da análise de regressão com variáveis como diâmetro, altura e idade das árvores, considerando características do plantio como espaçamento, espécie ou clone e capacidade produtiva para fins de estratificação (CAMPOS e LEITE, 2009). Recentemente, resultados satisfatórios têm sido obtidos ao empregar Redes Neurais Artificiais (RNA) ao invés de modelos de regressão (GUAN e GERTNER, 1991a; 1991b; 1995; GORDON, 1998; DIAMANTOPOULOU, 2005; BINOTI et al., 2009; GÖRGENS et al., 2009; BINOTI, 2010; LEITE et al., 2010).

Redes neurais artificiais (RNA) são aproximações brutas das redes e neurônios encontrados no cérebro humano. Do ponto de vista prático, consiste em um sistema computacional paralelo (camadas) constituído de várias unidades de processamento simples (neurônios artificiais) conectadas entre si de uma maneira específica para executar uma determinada tarefa (BULLINARIA, 2010).

Os estudos na área de RNA se justificam por diversas propriedades úteis e capacidades como: tolerância a falhas e ruídos, aprendizagem e generalização a partir de dados de treinamento, paralelismo massivo que as tornam muito eficientes, uniformidade de análise e projeto, analogia neurobiológica, mapeamento entrada-

saída e não linearidade (BRAGA et al., 2000; HAYKIN, 2001; BULLINARIA, 2010).

Os problemas tratáveis através de redes neurais enquadram-se nas tarefas de aprendizagem: aproximação de função, classificação de padrões, agrupamento de dados, predição (séries temporais), otimização, recuperação de conteúdo e controle (JAIN et al., 1996; HAYKIN, 2001).

Os estudos desenvolvidos nesta tese são problemas de aproximação de função direcionada para a estimação de estoques volumétricos de plantios florestais equiâneos. A aproximação de função consiste em projetar uma rede neural que aproxime a função desconhecida $f(x)$ e que descreva o mapeamento dos pares de entrada-saída $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ de um conjunto de n padrões de treinamento. Dessa maneira, o objetivo da tese foi treinar, aplicar e validar redes neurais artificiais para diferentes abordagens em mensuração e manejo florestal e desenvolver, aplicar e validar um sistema computacional específico para a obtenção de estimativas do volume de árvores.

A estrutura da tese está organizada em oito capítulos. O capítulo 1 consiste na construção e validação de um modelo de estimação da altura de árvores com a utilização de RNA. No capítulo 2 foi proposto um modelo de estimação da altura de árvores empregando RNA. Neste mesmo capítulo foram simuladas diferentes percentagens de redução na medição das alturas nas parcelas, visando eficiência e redução de custos. O capítulo 3 propôs uma metodologia com a utilização de RNA, para a redução do número de árvores a serem cubadas para a geração de equações volumétricas. No capítulo 4 são apresentados resultados do desenvolvimento e avaliação de RNA para a projeção de parâmetros da distribuição Weibull. Um modelo de distribuição diamétrica empregando um modelo de autômatos celulares unidimensionais e RNA é apresentado no capítulo 5, enquanto que no capítulo 6 esse mesmo enfoque é utilizado contemplando a simulação de desbaste. No capítulo 7 é apresentado um projeto de construção do sistema computacional que utiliza RNA para a estimação do volume de árvores, o *NeuroForest – Volumetric* e no capítulo 8 são avaliadas diferentes configurações e processos de treinamento de RNA para a estimação do volume de árvores.

REFERÊNCIAS BIBLIOGRÁFICAS

BINOTI, M. L. M. S.; 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. **Revista Árvore**, v. 33, n. 6, p. 1133-1139, 2009.

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

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

BULLINARIA, J. A. Introduction to Neural Computation. **Notas de aula**. 2008. Disponível em: <<http://www.cs.bham.ac.uk/~jxb/inc.html>>. Acesso em: 20 de jun. 2008.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3 ed. Viçosa: UFV, 2009. 548 p.

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

GORDON, C. **Artificial Neural Network Modeling of Forest Tree Growth**. 1998. 76 f. Dissertação (Mestrado) - University of the Witwatersrand, 1998.

GÖRGENS, E. B.; LEITE, H. G.; SANTOS, H. N.; GLERIANI, J. M. Estimação do volume de árvores utilizando redes neurais artificiais. **Revista Árvore**, v.33, n.6, p. 1141-1147, 2009.

GUAN, B. T.; GERTNER G. Using a parallel distributed processing system to model individual tree mortality. **Forest Science**, v. 37, n. 3, p. 871-885, 1991a.

GUAN, B. T.; GERTNER G. Modeling red pine tree survival with an artificial neural network. **Forest Science**, v. 37, n. 5, p. 1429-1440, 1991b.

GUAN, B. T.; GERTNER G. Modeling individual tree survival probability with a random optimization procedure: An artificial neural network approach. **AI Application**, v. 9, n.1, p. 39-52, 1995.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 900 p.

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. M. Artificial Neural Networks: A Tutorial. **Computer**, v. 29, n. 3, p. 31-44, 1996.

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 *Tectona grandis* Linn. trees using artificial neural networks. **European Journal of Forest Research**, v.130, n.2, p.263-269, 2011.

CAPÍTULO 1

APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS PARA ESTIMAÇÃO DA ALTURA DE ÁRVORES EM POVOAMENTOS DE EUCALIPTO

1. INTRODUÇÃO

A estimação de estoques de crescimento e de colheita é um importante elemento no manejo florestal, uma vez que fornece informações quantitativas sobre a floresta, auxiliando na definição de planos de manejo e em análises econômicas de prescrições de manejo.

A quantificação do estoque volumétrico pode ser realizada por meio de inventário florestal contínuo ou temporário. Ambos consistem basicamente na medição de amostras representativas da população, constituída por unidades de amostra (parcelas). O volume de madeira existente em cada uma dessas parcelas é obtido pela aplicação de modelos volumétricos, de razão volumétrica ou de afilamento, que em sua maioria possuem como variáveis independentes o diâmetro com casca a 1,3 m de altura (*dap*) e a altura total da árvore.

Estudos sobre modelos, procedimentos e equipamentos para a estimação da altura de árvores são importantes, visto que, a estimação da altura das árvores no campo é um componente significativo no custo do inventário florestal (LEITE e ANDRADE, 2004). A redução deste custo pode ser obtida pelo uso da relação hipsométrica, como proposto por Ker e Smith em 1957, onde a partir da medição dos diâmetros (*dap*) e das alturas de algumas árvores da parcela, pode-se estimar a altura das demais. A medição da altura das árvores é outra grande dificuldade do inventário florestal, devido ao fato de sua aferição ser realizada indiretamente. Problemas como

a falta de visibilidade do topo da árvore em uma floresta densa e a ocorrência de ventos pode tornar a medição impraticável em muitas ocasiões, principalmente em plantios de eucalipto. Além disso, a relação hipsométrica é afetada por fatores ambientais e características do povoamento como: capacidade produtiva, idade, material genético, variações ambientais e características qualitativas (CURTIS, 1967; ZHANG et al., 1997; FANG e BAILEY, 1998; CAMPOS e LEITE, 2009).

A inclusão de características do povoamento nos modelos hipsométricos, como índice de local e idade, pode resultar em estimativas mais precisas e um maior realismo biológico, tornando a equação aplicável em diferentes locais. As principais dificuldades para a inclusão dessas características são a modelagem e quantificação das influências sobre a variável a ser estimada (altura total), pois as relações apresentam características não lineares ou valores qualitativos (categóricos) como, por exemplo, o tipo de solo que são incluídas em regressões como variáveis binárias, podendo ocasionar um aumento na complexidade de modelagem.

Uma opção é o emprego de modelos que incluem menor número de variáveis e o ajuste dos mesmos precedido de uma pré-estratificação com base nas variáveis categóricas. Ocorre que nem sempre é possível ter amostragem eficiente em todos os estratos, por questão de tempo e, principalmente, de custos.

Uma abordagem alternativa para a modelagem da relação hipsométrica, consiste no uso de redes neurais artificiais (RNA), que são sistemas computacionais paralelos constituídos por unidades de processamento simples, também denominadas neurônios artificiais ou nodos, conectadas entre si de uma maneira específica para desempenhar uma determinada tarefa (BINOTI, 2010). Os neurônios artificiais são modelos matemáticos simplificados dos neurônios biológicos que processam as informações recebidas, ponderadas por pesos sinápticos fornecendo uma única resposta (HAYKIN, 2001; BRAGA et al., 2007).

Em alguns casos, as RNA têm apresentado desempenho superior aos modelos de regressão devido a diversos fatores como: estrutura maciça e paralelamente distribuída (camadas); habilidade de aprender e generalizar, que as tornam capazes de resolver problemas complexos; tolerância a falhas e ruídos; capacidade de modelar diversas variáveis e suas relações não-lineares; capacidade de modelagem com variáveis categóricas (qualitativas), além das numéricas (quantitativas); e analogia neurobiológica (HAYKIN, 2001; BULLINARIA, 2010).

A modelagem da altura de povoamentos florestais com RNA se enquadra na tarefa de aprendizagem denominada aproximação de funções. Esta tarefa consiste em treinar uma rede neural capaz de aproximar a função desconhecida $f(x)$ que descreve o mapeamento dos pares de entrada-saída $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ de um conjunto de n exemplos de treinamento (JAIN et al., 1996). O treinamento, também denominado de aprendizagem, consiste no ajuste dos parâmetros da rede a partir de dados de treinamento (exemplos) a fim de desempenhar uma determinada tarefa (HAYKIN, 2001; BRAGA et al, 2007).

Este trabalho teve como objetivo propor, construir e validar um modelo de estimação da altura de árvores com a utilização de redes neurais artificiais, visando aumentar a precisão das estimativas usualmente obtidas por meio de regressão e reduzir a necessidade de aferição das mesmas em campo, bem como os custos do inventário florestal.

2. MATERIAL E MÉTODOS

2.1 Dados

Os dados utilizados neste estudo foram obtidos em cerca de 3000 árvores mensuradas em 145 parcelas permanentes de 215 m², medidas em seis ocasiões, em povoamentos de clones de eucalipto localizados na Bahia, em uma área de cerca de 2.000 ha, contemplando cinco tipos de solo e três clones (C1, C2 e C3).

As variáveis utilizadas para estimar a altura total das árvores foram: (a) quantitativas: idade (meses), diâmetro com casca a 1,30 m de altura (*dap*) e altura dominante da parcela; e (b) qualitativa: tipo de solo com suas respectivas classes.

2.2 Cenários

Para validação e aplicação da metodologia proposta foram consideradas duas situações: (a) quando há a introdução de um novo material genético e não existem informações sobre sua relação hipsométrica; (b) quando já se conhece a tendência de crescimento em altura dos povoamentos implantados, obtida pela existência de medições em parcelas de Inventário Florestal Contínuo (IFC).

Para o cenário (a), utilizou-se a primeira medição do clone C1. As informações deste clone foram estratificadas em três classes em função da amplitude da altura dominante (*Hd*). O treinamento das redes foi feito com dados de uma parcela selecionada aleatoriamente em cada tipo de solo e classe de altura dominante, totalizando 15 parcelas (5 tipos de solo x 3 classes de *Hd*). Nessas parcelas foram medidos os *dap* e altura de todas as árvores.

Na simulação do cenário (b), duas abordagens foram utilizadas: (1) considerando uma rede para cada clone e (2) empregando uma rede genérica para

todo o povoamento tendo a inclusão do clone como variável categórica. Os dados foram estratificados em três classes em função da dispersão da altura dominante, com o intuito de abranger todas as classes de capacidade produtiva. Em ambas as abordagens, selecionou-se 80 % dos dados para o treinamento das redes, considerando parcelas em todas as combinações de variáveis quantitativas (clone x tipo de solo). Os demais dados (20%) foram utilizados no processo de generalização.

2.3 Treinamento e Avaliação das RNA

A estimativa da altura foi simulada com todas as combinações possíveis das variáveis de entrada (idade, *dap*, altura dominante da parcela e tipo de solo). As redes treinadas foram do tipo *perceptrons* de múltiplas camadas, comumente conhecidas como MLP (*Multilayer Perceptron*), consistindo em duas camadas de neurônios artificiais que processam os dados (camada intermediária e camada de saída) e uma camada de neurônios artificiais que apenas recebem os dados (camada de entrada) e direciona-os à camada intermediária. Para obtenção das redes neurais artificiais, utilizou-se o software Statistica 7 (STATSOFT, 2007).

As melhores RNA foram selecionadas com base na correlação entre a altura observada e a altura estimada pelas redes ($r_{H\hat{H}}$) e na estabilidade dos índices de treinamento das redes, fornecidos pelo software nas fases de treinamento, seleção e avaliação. As redes selecionadas foram aplicadas aos dados das demais parcelas permanentes para estimar a altura total das árvores individuais.

As estimativas das alturas totais através das RNA foram avaliadas com base nos seguintes critérios: dispersão dos erros percentuais, histograma de frequência dos erros percentuais e a raiz quadrada do erro médio (RMSE - *Root Mean Square Error*). Sendo o erro percentual obtido por:

$$erro_{\%} = \frac{(\hat{H} - H)}{H} \times 100$$

H é a altura total observada nos inventários, \hat{H} é a altura total estimada pela RNA.

A raiz quadrada do erro médio avalia a diferença quadrática média entre os valores observados e estimados. Quanto menor o RMSE, melhor a precisão da estimativa, sendo ótimo quando é igual a zero (MEHTÄTALO et al., 2006):

$$RMSE(\%) = \frac{100}{\bar{H}} \sqrt{\frac{\sum_{i=1}^n (H_i - \hat{H}_i)^2}{n}}$$

\bar{H} é a média das alturas totais observadas e n é o número total de observações.

3. RESULTADOS

3.1 Cenário (a)

No cenário (a), simulou-se a introdução de um material genético do qual não se conhece a relação hipsométrica. A estratificação em função da altura dominante resultou em classes com amplitude de 5 m e centro de classes com valores de 10, 15, e 20 m. Todas as possíveis combinações de variáveis de entrada foram testadas sendo que as combinações que melhor expressaram a relação hipsométrica existente foram: $H = f(Hd, dap)$, $H = f(Hd, dap, I)$, $H = f(Hd, dap, I, solo)$ e $H = f(Hd, dap, solo)$. As características das redes selecionadas com base nos índice de treinamento e na correlação entre as variáveis são apresentadas na Tabela 1.

Tabela 1 – Características das redes neurais artificiais selecionadas para estimar a altura total das árvores na situação (a).

RNA	Arquitetura*	Variáveis de entrada	$r_{H\hat{H}}$	Índice do treino**	Índice da seleção**	Índice da avaliação**
1	2-5-1	Hd, dap	0,9986	0,044	0,068	0,056
2	3-2-1	Hd, dap, I	0,9987	0,043	0,065	0,054
3	7-3-1	Hd, dap, solo	0,9980	0,059	0,069	0,066
4	8-6-1	Hd, dap, I, solo	0,9986	0,044	0,069	0,055

* Número de neurônios em cada camada. ** Índices do treino (obtenção da rede), seleção (parada do treino) e avaliação (qualidade da rede treinada) que devem ser estáveis, ou seja, pouca variação entre eles. Hd = altura dominante (m), dap = diâmetro com casca a 1,30 m de altura (cm), I = idade (meses). $r_{H\hat{H}}$ = correlação entre altura total estimada e observada.

A dispersão e a frequência relativa dos erros percentuais mostraram alta precisão das estimativas da altura nas demais parcelas permanentes a partir das redes selecionadas, sempre acima de 90% dos casos com erro de $\pm 2,5\%$, que é uma exatidão extraordinária no caso da altura (Figura 1).

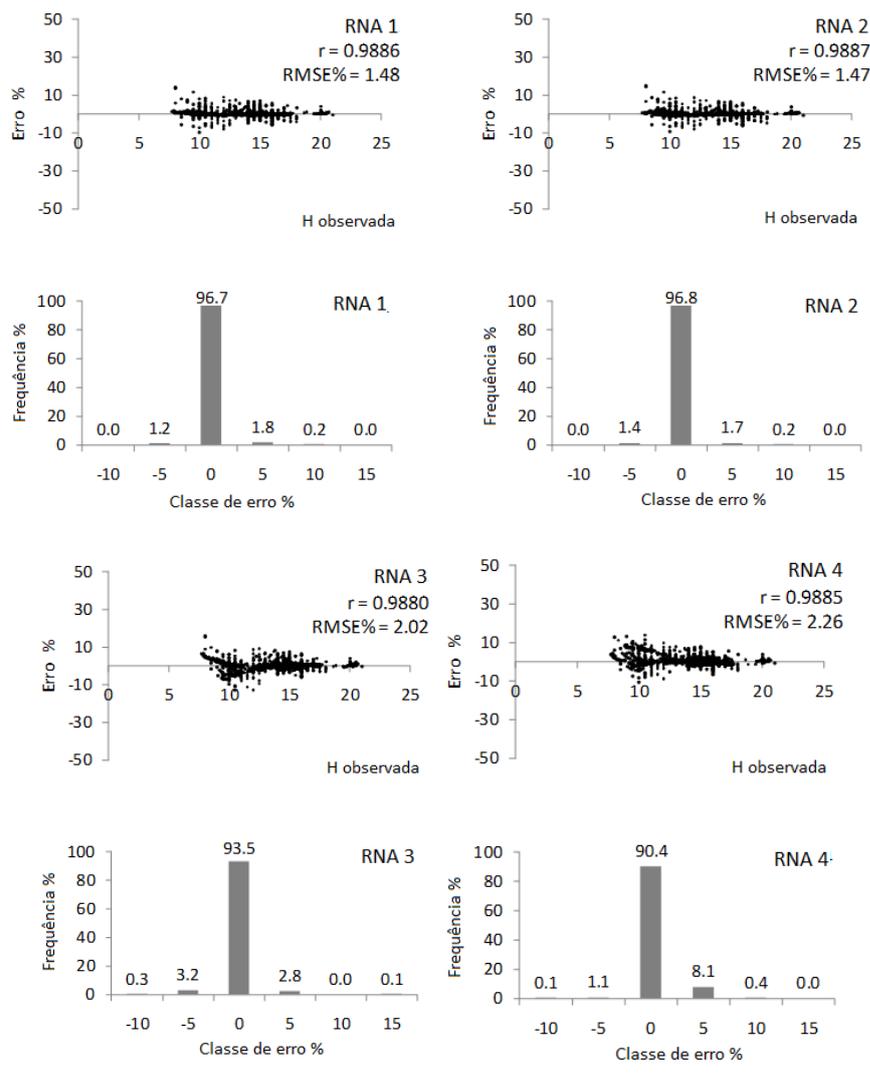


Figura 1 – Dispersão e histograma de frequência dos erros percentuais da generalização das RNA para estimar a altura total das árvores no cenário (a).

3.2 Cenário (b)

A simulação do cenário (b) foi feita com dados que possuíam seis medições em idades variando de 25 a 84 meses. Todas as possíveis combinações de variáveis de entrada foram testadas, considerando a abordagem individual por clone e para todo o povoamento, sendo selecionadas as redes que melhor expressaram a relação hipsométrica com base nos índices de treinamento e na correlação entre as variáveis (Tabela 2). Com estas redes realizou-se a generalização para comprovar a aplicação da metodologia, sendo que os gráficos de resíduos são apresentados na Figura 2 e os histogramas percentuais de resíduos na Figura 3.

Tabela 2 – Características das redes neurais artificiais (RNA) selecionadas para estimar a altura total das árvores na situação (b).

RNA	Clone	Arquitetura*	Variáveis de entrada	Treinamento			Generalização		
				r_{HH}	Índice do treino**	Índice da seleção**	Índice da avaliação**	r_{HH}	RMSE%
1	C ₁	2-5-1	Hd, dap	0,9939	0,11	0,11	0,11	0,9924	3,14
2	C ₁	6-5-1	Hd, dap, S	0,9943	0,11	0,11	0,11	0,9926	3,06
3	C ₁	7-8-1	Hd, dap, I, S	0,9958	0,09	0,09	0,09	0,9948	2,55
4	C ₁	3-6-1	Hd, dap, I	0,9961	0,09	0,09	0,09	0,9953	2,43
5	C ₂	10-3-1	Hd, dap, S	0,9846	0,178	0,178	0,164	0,9816	4,90
6	C ₂	2-3-1	Hd, dap	0,9902	0,155	0,126	0,120	0,9860	4,42
7	C ₂	11-3-1	Hd, dap, I, S	0,9922	0,132	0,119	0,113	0,9921	3,20
8	C ₂	3-5-1	Hd, dap, I	0,9919	0,142	0,115	0,105	0,9935	2,82
9	C ₃	2-4-1	Hd, dap	0,9853	0,169	0,172	0,174	0,9856	5,00
10	C ₃	8-2-1	Hd, dap, S	0,9870	0,157	0,164	0,164	0,9846	4,84
11	C ₁ , C ₂ , C ₃	15-1-1	Hd, dap, C, S	0,9855	0,170	0,171	0,168	0,9833	4,8
12	C ₁ , C ₂ , C ₃	6-5-1	Hd, dap, I, C	0,9902	0,142	0,137	0,137	0,9897	3,7
13	C ₁ , C ₂ , C ₃	2-2-1	Hd, dap	0,9916	0,131	0,126	0,129	0,9895	3,8
14	C ₁ , C ₂ , C ₃	12-4-1	Hd, dap, S	0,9921	0,127	0,122	0,125	0,9899	3,7
15	C ₁ , C ₂ , C ₃	13-4-1	Hd, dap, I, S	0,9943	0,110	0,104	0,105	0,9934	3,0

* Número de neurônios em cada camada. ** Índices do treino (obtenção da rede), seleção (parada do treino) e avaliação (qualidade da rede treinada) que devem ser estáveis, ou seja, pouca variação entre eles. Hd = altura dominante (m), dap = diâmetro com casca a 1,30 m de altura (cm), I = idade (meses), S = solo, C = clone. r_{HH} = correlação entre altura total estimada e observada.

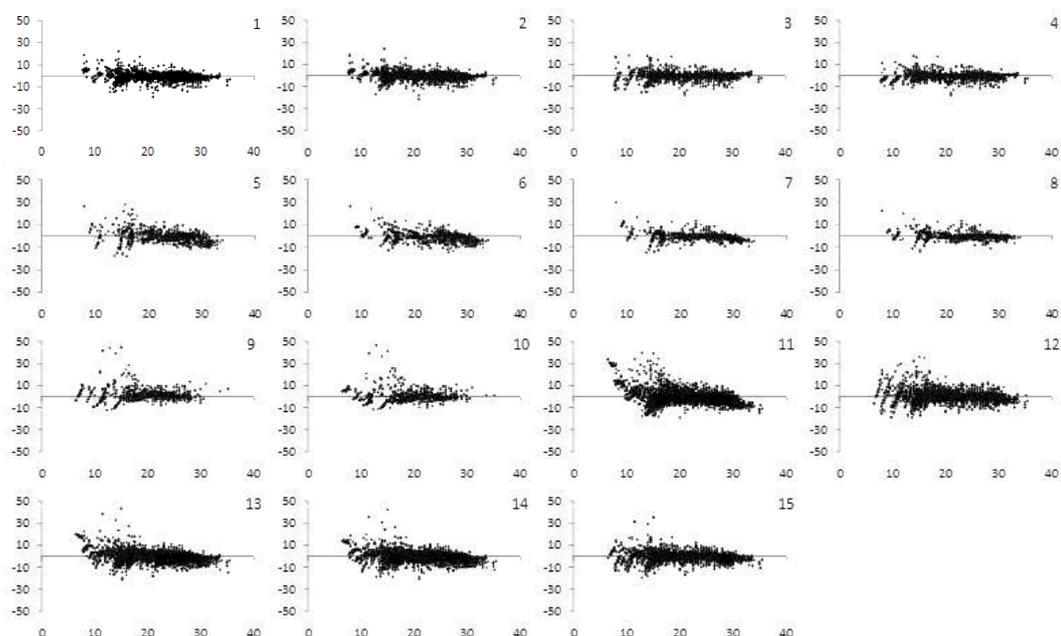


Figura 2 – Dispersão dos erros percentuais das estimativas da altura total pelas RNA (y) em função dos valores observados da altura total (x).

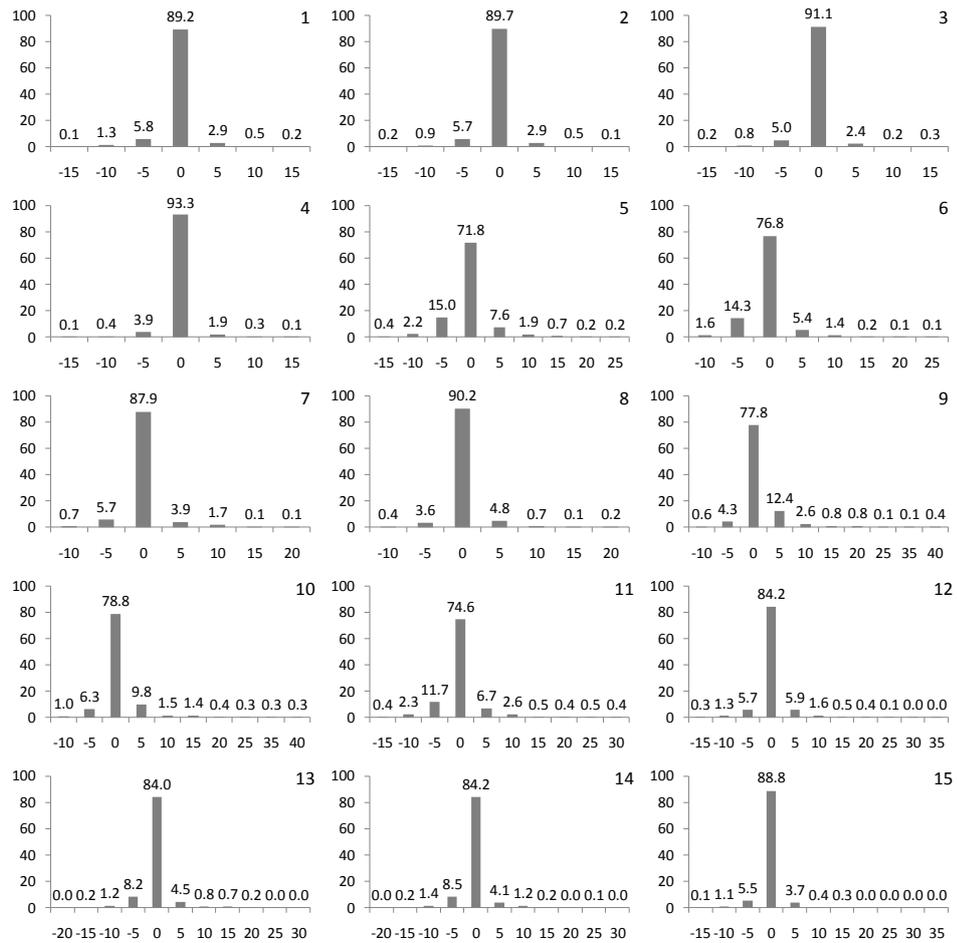


Figura 3 – Frequência percentual dos erros percentuais das estimativas da altura total pelas redes neurais artificiais (y) em função das classes de erros percentuais (x).

4. DISCUSSÃO

O diâmetro a 1,3 m de altura (*dap*) é uma das variáveis mais utilizadas na mensuração florestal, isso se deve ao fato de ser medido facilmente e com alta precisão nos povoamentos florestais. Por outro lado, a altura total de todas as árvores das parcelas de inventário, não é tão comumente medida por várias razões, entre as quais: (1) tempo necessário para concluir as medições (2); erro de observação, e (3) obstruções visual (COLBERT e LOOTENS, 2002; CAMPOS e LEITE, 2009). Consequentemente, a altura é medida somente em algumas árvores, sendo as demais estimadas por equações hipsométricas. A utilização da relação hipsométrica é responsável por uma redução significativa no tempo e nos custos de medições de parcelas de inventário florestal. Apesar disso, o custo com a medição de altura é ainda muito significativo. Diversos modelos para a predição da altura de povoamentos são encontrados na literatura (CURTIS, 1967; MONSERUD, 1975; EK et al., 1984; LARSEN e HANN, 1987; PARRESOL, 1992; FLEWELLING e DE JONG, 1994; COLBERT e LOOTENS, 2002; INOUE e YOSHIDA, 2004; CAMPOS e LEITE, 2009).

O presente estudo objetivou aumentar a precisão das estimativas de altura, reduzindo a necessidade de medição e o custo do inventário. Realizou-se a simulação de dois cenários comumente encontrados em empresas florestais. No primeiro, foi simulada a introdução de um novo material genético, com apenas uma medição. Um segundo cenário foi feito com base em informações que estão disponíveis na maioria das vezes, ou seja, uma série histórica de medições das parcelas de inventário contínuo.

Para a aplicação da metodologia considerando o cenário (a) obteve-se resultados satisfatórios para a estimação de alturas de clones introduzidos pela empresa. As redes 1, 2 e 3, não apresentaram tendenciosidade na dispersão dos resíduos, enquanto que a rede 4 apresentou pequena tendência (8,1% dos casos) de superestimação dos valores estimados (Figura 1). Contudo, as relações funcionais demonstraram-se satisfatórias para a estimação da altura, para o cenário em questão, com um percentual de casos com erro de $\pm 2,5\%$ acima de 90%. Isto dificilmente será conseguido com o emprego de relações hipsométricas.

Para a aplicação da metodologia considerando o cenário (b) obteve-se resultados satisfatórios para a construção de redes estratificando para cada clone e a construção de uma rede genérica considerando todos os clones. Pelo fato de se utilizar medições em todas as idades do povoamento, todas as redes e em todas as relações funcionais utilizadas não apresentaram tendenciosidade, com alta precisão, que são comprovadas pela pequena dispersão de resíduos apresentadas nas Figuras 2 e 3. Contudo recomenda-se a utilização de apenas uma rede para todo o povoamento para facilitar o processo de modelagem. Para as redes em questão dispunha-se somente de informações de solo e clone, sendo que a introdução de informações adicionais como precipitação, tipo de preparo do terreno, arranjo espacial e radiação, dentre outras, pode aumentar a precisão das estimativas.

A aplicação das metodologias (a) e (b) em conjunto permite reduzir o tempo para a realização do inventário. A aplicação da metodologia (a) reduz significativamente medições em clones recém introduzidos no povoamento, pois se mede apenas um número mínimo de parcelas por classe de local. Com a obtenção das redes com o emprego da metodologia (b), a realização do próximo inventário é feita somente com a mensuração do *dap* e das alturas das árvores dominantes do povoamento, que são necessárias também para a classificação da capacidade produtiva. Recomenda-se a mensuração da altura em somente algumas parcelas para o acompanhamento das estimativas geradas pelas redes e possíveis calibrações das mesmas. Além disso, conforme Leite et al. (2011), em alguns casos, o diâmetro das árvores dominantes pode ser empregado em substituição à altura.

5. CONCLUSÃO

O modelo proposto para estimação da altura de árvores por meio de redes neurais artificiais é eficiente e pode ser empregado, com vantagens, em substituição aos usuais modelos hipsométricos.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- BINOTI, M. L. M. DA S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. 2010. 54 f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2010.
- BRAGA, A. DE P.; CARVALHO, A. P. DE L. F. DE; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. 2 ed. Rio de Janeiro: LTC, 2007.
- BULLINARIA, J. A. **Introduction to Neural Computation**. Notas de aula. Disponível em: <http://www.cs.bham.ac.uk/~jxb/inc.html>. 2010.
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3 ed. Viçosa: UFV, 2009. 548 p.
- COLBERT, K. C.; LARSEN, D. R.; LOOTENS, J. R. Height-diameter equations for thirteen midwestern bottomland hardwood species. **Northern Journal of Applied Forestry**. v.19, n.4, p.171-176. 2002.
- CURTIS, R.O. Height-diameter and height-diameter-age equations for second growth Douglas-fir. **Forest Science**. v.13, p.365-375. 1967.
- EK, A. R.; BIRDSALL, E. T.; SPEARS, R. J. **A simple model for estimating total and merchantable tree heights**. USDA Forest Service Research Note NC-309. 5 p. 1984.
- FANG, Z.; BAILEY, R. L. Height-diameter models for tropical forests on Hainan Island in southern China. **Forest Ecology and Management**, v.110, p. 315-327, 1998.
- FLEWELLING, J. R.; DE JONG, R. Considerations in simultaneous curve fitting for repeated height-diameter measurements. **Canadian Journal Forest Research**. v.24, p.1408-1414, 1994.
- HAYKIN, S. **Redes Neurais: princípios e prática**. Porto Alegre, 2001. 900 p.

INOUE, A.; YOSHIDA, S. Allometric model of the height–diameter curve for even-aged pure stands of Japanese cedar (*Cryptomeria japonica*). **Journal Forest Research**, v.9, p.325-331, 2004.

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. M. Artificial Neural Networks: A Tutorial. **Computer**, v. 29, n. 3, p. 31-44, 1996.

LARSEN, D. R.; HANN, D. W. **Height-diameter equations for seventeen tree species in southwest Oregon**. Res. Pap. 49, Forest Research Laboratory, Oregon State Univ., Corvallis, OR. 1987. 16 p.

LEITE, H.G.; ANDRADE, V. C. L. Uso do método da altura relativa em inventário florestal de um povoamento de Pinus. **Revista Árvore**, v.28, n.6, p. 865-873, 2004.

LEITE, H. G.; CASTRO, R. V. O.; SILVA, A. L.; JÚNIOR, C. A. A.; BINOTI, D. H. B.; CASTRO, A. F. N. M.; BINOTI, M. L. M. S. Classificação da capacidade produtiva de povoamentos de eucalipto utilizando diâmetro dominante. **Silva Lusitana**, v. 19, p. 169-183, 2011.

MONSERUD, R. **Methodology for simulating Wisconsin northern hardwood stand dynamics**. PhD dissertation (Univ. Microfilm No. 7602496), University. of Wisconsin, Madison, WI. 1975. 156 p.

PARRESOL, B.R. Baldcypress height-diameter equations and their prediction confidence intervals. **Canadian Journal Forest Research**, v.22, n.1429-1434, 1992.

VAN LAAR, A.; AKÇA, A. **Forest Mensuration**. Göttingen, Germany: Cuvillier Verlag, 1997. 418 p.

ZHANG, S.; BURKHART, H. E.; AMATEIS, R. L. The influence of thinning on tree height and diameter relationships in loblolly pine plantations. **Southern Journal of Applied Forestry**, v.21, p.199-205, 1997.

CAPÍTULO 2

REDUÇÃO DOS CUSTOS EM INVENTÁRIO DE POVOAMENTOS EQUIÂNEOS COM O EMPREGO DE REDES NEURAIIS

1. INTRODUÇÃO

A exatidão das estimativas do volume de povoamentos equiâneos depende, entre outros, da qualidade das estimativas obtidas em nível de árvores. As medições de diâmetros e alturas de árvores individuais são comumente aplicadas para obtenção da estimativa do estoque de crescimento e a realização de análises de viabilidade econômica. Diâmetros podem ser medidos facilmente com custos relativamente baixos, contudo as medições de altura são demoradas, muitas vezes imprecisas e relativamente de difícil aferição em plantios densos. Assim, as alturas são derivadas indiretamente dos diâmetros, utilizando uma relação conhecida ou estimada entre os diâmetros e alturas (VAN LAAR e AKÇA 1997; CAMPOS e LEITE, 2009).

A relação entre os diâmetros e as alturas pode ser descrita através de uma relação hipsométrica ou uma distribuição de altura-diâmetro bivariada. Na prática, são medidos os diâmetros de todas as árvores das parcelas e algumas alturas, que são utilizadas para gerar equações hipsométricas, visando a estimação da altura das demais árvores (ARABATIZIS e BURKHART, 1992; CAMPOS e LEITE, 2009). Estudos sobre modelos, procedimentos e equipamentos para a mensuração da altura de árvores são importantes, visto que, a mediação da altura das árvores é considerada um componente significativo no custo do inventário florestal (LEITE e ANDRADE, 2004).

Uma alternativa para reduzir tempo e custo de condução de inventários pode ser o emprego de redes neurais artificiais (RNA) para estimação da altura. Assim, foi idealizado este estudo com o objetivo de propor, construir e validar um modelo de estimação da altura de árvores em povoamentos equiâneos empregando redes neurais artificiais e avaliar a eficiência desse modelo na redução do custo do inventário florestal.

2. MATERIAL E MÉTODOS

2.1 Dados

Os dados utilizados neste estudo foram provenientes de medições de 91 parcelas permanentes instaladas em povoamentos de um clone de eucalipto com cerca de 2000 ha, distribuídos em cinco tipos de solos. Essas parcelas foram medidas em sete ocasiões. A altura total das árvores (H) foi estimada em função das seguintes variáveis: diâmetro com casca a 1,3 m de altura (dap) em cm, altura dominante da parcela (Hd) em m e idade (I) em meses.

2.2 Treinamento e avaliação das RNA

Avaliou-se a quantidade de parcelas nas quais se deixaria de medir a altura das árvores. A simulação consistiu no sorteio das seguintes proporções: 90, 80, 70, 60, 50, 40, 30, 20 e 10% das parcelas para treinamento das redes neurais artificiais e o restante dos dados (10, 20, 30, 40, 50, 60, 70, 80 e 90%, respectivamente) foram utilizados para a generalização, ou seja, para avaliar a capacidade de estimação das redes para dados desconhecidos. Dessa maneira, nas parcelas do conjunto de treinamento ocorreram as medições normalmente de todos os dap , das alturas dominantes e da altura total de algumas árvores, enquanto que nas parcelas do conjunto de generalização, omitiram-se as medições das alturas totais.

As redes neurais artificiais (RNA) treinadas foram do tipo *multilayer perceptron* (MLP) ou *perceptron* de múltiplas camadas, com uma camada de entrada que recebe as variáveis de entrada (dap , Hd e I) e transfere-as ponderadas por pesos sinápticos à camada intermediária ou oculta que aplica transformações matemáticas e transfere para a camada de saída que fornece a resposta da rede (estimativa da H).

Foram treinadas 50 redes para cada proporção de parcelas e a melhor rede (para cada proporção) em termos dos índices de treinamento e da correlação foi aplicada aos dados de generalização. As estimativas das alturas totais das árvores foram avaliadas com base na correlação, nos gráficos de erros percentuais em função dos valores observados e das frequências percentuais dos erros percentuais e na raiz quadrada do erro médio (RMSE%), sendo:

$$RMSE(\%) = \frac{100}{\bar{H}} \sqrt{\frac{\sum_{i=1}^n (H_i - \hat{H}_i)^2}{n}}$$

em que, \bar{H} é a média das alturas observadas, H_i é a altura observada nos inventários na i -ésima observação, \hat{H}_i é a altura estimada pela RNA na i -ésima observação e n é o número total de observações.

2.3 Avaliação Econômica

A avaliação baseou-se na simulação de cenários sobre um estudo de caso para a mensuração de um povoamento florestal de 5000 ha, com uma intensidade amostral média de uma parcela para seis hectares, resultando em 982 parcelas. A contribuição econômica da metodologia proposta foi avaliada em função da redução dos custos com relação a um orçamento proposto comercialmente por uma empresa de consultoria em mensuração florestal. Em função da inexistência de coeficientes técnicos para a real avaliação econômica dos benefícios da metodologia proposta, simularam-se diversas porcentagens na diminuição do tempo de medição das parcelas no campo.

Na Tabela 1 é apresentado o orçamento comercial proposto para o estudo de caso avaliado. Para melhor avaliação da metodologia proposta foram ajustados modelos lineares que descrevem o custo médio da parcela em função da redução do tempo de medição e do valor normal de medição da parcela (Tabela 2).

Tabela 1 – Orçamento proposto para a realização do inventário florestal para o estudo de caso proposto.

Serviço	Quantidade	Unidade	Preço/ Unidade	Total
Planejamento, treinamento e acompanhamento	1	Parcela	9.731,69	9.731,69
Coleta de dados - Parcelas permanentes	982	Parcela	56,27	55.261,19
Hospedagem - Inventário	982	Parcela	23,15	22.730,02
Alimentação - Inventário	982	Parcela	11,57	11.365,01
Transporte - Inventário	982	Parcela	16,91	16.607,13
Coleta de dados - Cubagem	802	Árvore	23,28	18.673,67
Hospedagem - Cubagem	802	Árvore	9,05	7.258,32
Alimentação - Cubagem	802	Árvore	4,53	3.629,16
Transporte - Cubagem	802	Árvore	6,60	5.289,88
Total				150.546,07

Tabela 2 – Equações do custo médio da parcela em função da redução do tempo de medição das árvores.

Porcentagem de Redução do Tempo de Medição	Equação
10%	$Custo = valorParcela - 0,00063(valorParcela \times Parc\%)$
20%	$Custo = valorParcela - 0,00184(valorParcela \times Parc\%)$
30%	$Custo = valorParcela - 0,00272(valorParcela \times Parc\%)$

Custo é o custo médio resultante da aplicação da metodologia (em reais), *valorParcela* é o valor atual da parcela (em reais), e *Parc%* representa a quantidade de parcelas sem medição de alturas.

3. RESULTADOS

As características das redes selecionadas com base nos índice de treinamento e na correlação entre as variáveis são apresentadas na Tabela 3.

Tabela 3 – Características e estatísticas das redes neurais artificiais selecionadas para estimar a altura total das árvores.

RNA	Arquitetura *	Índice do treino **	Índice da seleção **	Índice da avaliação **	%	Treinamento		Generalização	
						$r_{H\hat{H}}$	RMSE%	$r_{H\hat{H}}$	RMSE %
1	3-8-1	0,09	0,08	0,08	10	0,9964	2,1	0,9946	2,5
2	3-9-1	0,09	0,09	0,09	20	0,9961	2,2	0,9949	2,4
3	3-9-1	0,10	0,10	0,10	30	0,9947	2,5	0,9953	2,3
4	3-9-1	0,10	0,10	0,10	40	0,9950	2,4	0,9949	2,5
5	3-7-1	0,10	0,10	0,10	50	0,9951	2,3	0,9954	2,4
6	3-9-1	0,10	0,10	0,10	60	0,9949	2,5	0,9959	2,1
7	3-9-1	0,10	0,10	0,10	70	0,9950	2,3	0,9958	2,2
8	3-9-1	0,10	0,10	0,10	80	0,9952	2,4	0,9949	2,5
9	3-9-1	0,10	0,11	0,10	90	0,9950	2,4	0,9971	1,9

* Número de neurônios em cada camada. ** Índices do treino (obtenção da rede), seleção (parada do treino) e avaliação (qualidade da rede treinada) que devem ser estáveis, ou seja, pouca variação entre eles. % = Proporção de parcelas nas quais a altura total de algumas árvores é medida. $r_{H\hat{H}}$ = correlação entre os valores observados e estimados das alturas totais.

As redes selecionadas foram aplicadas aos dados das demais parcelas permanentes para estimação da altura, com base no *dap*, altura dominante e idade. A dispersão e frequência relativa dos erros percentuais mostram a alta precisão das estimativas (Figura 1 e 2).

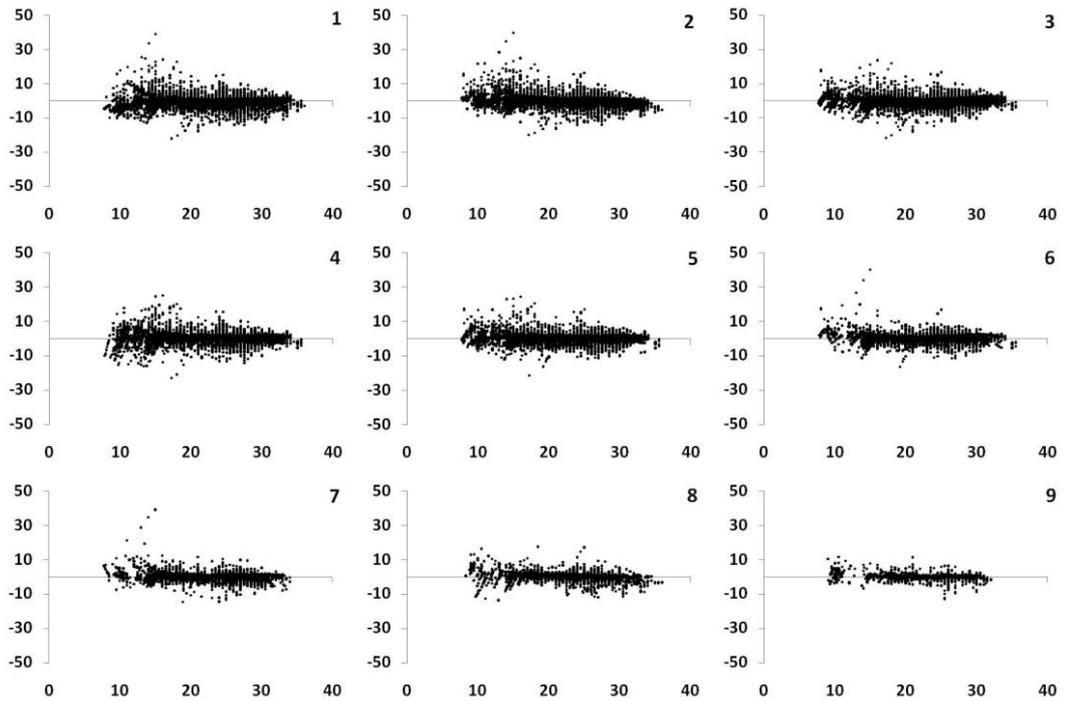


Figura 1 – Dispersão dos erros percentuais das alturas totais estimadas pelas redes neurais artificiais (eixo y) em função dos valores observados (eixo x).

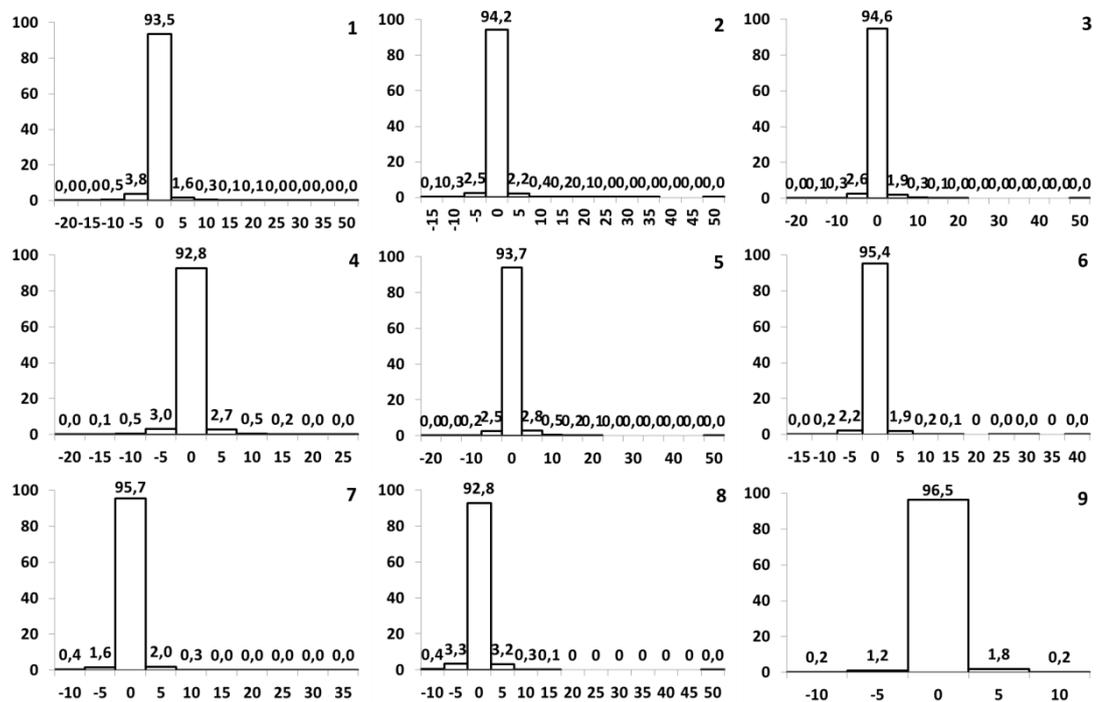


Figura 2 – Frequência percentual (y) dos erros percentuais (x) das alturas totais estimadas pelas redes neurais artificiais.

O custo total previsto foi de R\$150.546,07, considerando que o inventário fosse feito da forma convencional (Tabela 1). Na Figura 3, são apresentadas as simulações de redução do custo em função da diminuição do tempo de medição das parcelas.

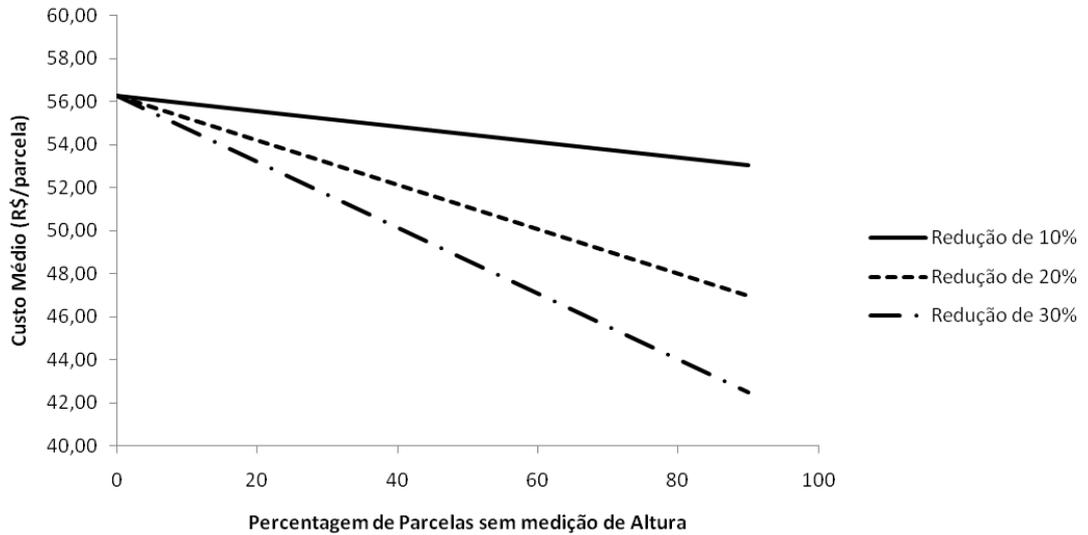


Figura 3 – Simulação de redução dos custos para o estudo de caso proposto em função da melhoria do tempo de medição.

4. DISCUSSÃO

A utilização da relação hipsométrica é responsável por uma redução significativa no tempo e no custo de medição da parcela. Diversos modelos para a predição da altura de povoamentos são encontrados na literatura (CURTIS, 1967; MONSERUD, 1975; EK et al., 1984; LARSEN e HANN, 1987; FLEWELLING e DE JONG, 1994; INOUE e YOSHIDA, 2004; CAMPOS e LEITE, 2009).

O presente estudo mostra o efeito de redução do custo de realização de inventário florestal pela redução das medições de altura com o aumento da exatidão obtida por métodos tradicionais. Simulou-se a realização de redução de percentuais de medição de parcelas e avaliou-se a aplicação de reduções destas em campo. Em função dos coeficientes operacionais para a medição da altura ser fortemente influenciado por condições climáticas (ex. ventos), e densidade populacional, optou-se por avaliar somente reduções no tempo de medição individual das parcelas.

A avaliação dos benefícios econômicos gerados pela aplicação da metodologia pode ser avaliada pela utilização das equações geradas (Figura 3). Para o estudo de caso apresentado, a medição de somente 10% das alturas das parcelas pode ser realizada. Considerando uma redução de 20% no tempo de medição da parcela, utilizando a equação proposta resulta em um custo de $56,27 - 0,00184 * 56,27 * 90 = R\$ 46,95$, ou seja, uma redução no custo de R\$9,32 (16%). Como neste estudo de caso considerou-se um total de 982 parcelas, isso representa uma economia de R\$ 9.152,00 no custo do inventário (R\$ 150.546,07 para R\$ 141.394,07).

A comprovação da precisão das estimativas da metodologia proposta é feita pela análise gráfica dos resíduos apresentada na Figura 1. Os histogramas apresentados na Figura 2 demonstram uma pequena dispersão dos erros percentuais

em torno de zero. A análise de ambas as figuras permite inferir sobre a inexistência de tendenciosidade nas estimativas proporcionadas pela metodologia proposta. O menor percentual de casos com erro de $\pm 2,5\%$ foi de 92,8%, sendo este um resultado extraordinário em se tratando de altura de árvores. De acordo com os resultados pode-se mensurar a altura em somente 10 % das parcelas inventariadas, sem perdas de precisão nas estimativas de alturas, isso reflete numa economia de tempo e custo como discutido acima. Como geralmente têm-se vários inventários sucessivos ao longo dos anos essa economia na verdade será bem maior.

A aplicação da metodologia permite reduzir consideravelmente o custo para a realização do inventário florestal. Além disso, a realização do próximo inventário seria feita somente com a mensuração do *dap* e das alturas das árvores dominantes do povoamento, por serem necessárias também para a classificação da capacidade produtiva. Opcionalmente a altura dominante pode ser definida com base no diâmetro, ou seja, altura dominante da parcela igual à média das alturas das 100 árvores normais de maior *dap* por hectare. Nesse caso a altura das maiores árvores seria determinada também pela RNA. Recomenda-se a mensuração da altura em somente algumas parcelas para o acompanhamento das estimativas.

5. CONCLUSÃO

O uso das redes neurais artificiais para estimar altura de povoamentos foi eficiente e permitiu reduzir o número de medições em altura sem perda de precisão e exatidão.

O tempo e o custo de condução de inventário florestal podem ser reduzidos de modo significativo com o emprego de redes neurais artificiais.

É possível estimar a altura de árvores em parcelas de inventário com erro de $\pm 2,5\%$ em mais de 90% das vezes, utilizando redes neurais artificiais treinadas com dados de altura obtidos em cerca de 10% das parcelas permanentes.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ARABATZIS, A. A.; BURKHART, H. E. An evaluation of sampling methods and model forms estimating height-diameter relationships in loblolly pine plantations. **Forest Science**, v. 38, n.1, p.192-198. 1992.

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

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

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3 ed. Viçosa: UFV, 2009. 548 p.

COLBERT, K. C.; LARSEN, D. R.; LOOTENS, J. R. Height-diameter equations for thirteen midwestern bottomland hardwood species. **Northern Journal of Applied Forestry**. v.19, n.4, p.171-176. 2002

CURTIS, R.O. Height-diameter and height-diameter-age equations for second growth Douglas-fir. **Forest Science**. v.13, p.365-375. 1967.

EK, A. R.; BIRDSALL, E. T.; SPEARS, R. J. **A simple model for estimating total and merchantable tree heights**. USDA Forest Service Research Note NC-309. 5 p. 1984.

FLEWELLING, J. R.; JONG, R. DE. Considerations in simultaneous curve fitting for repeated height-diameter measurements. **Canadian Journal Forest Research**. v.24, p.1408-1414. 1994.

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

INOUE, A.; YOSHIDA, S. Allometric model of the height-diameter curve for even-aged pure stands of Japanese cedar (*Cryptomeria japonica*). **Journal Forest Research**. v.9, p.325-331. 2004.

LARSEN, D. R.; HANN, D. W. **Height-diameter equations for seventeen tree species in southwest Oregon**. Res. Pap. 49, Forest Research Laboratory, Oregon State Univ., Corvallis, OR. 1987. 16 p.

LEITE, H.G.; ANDRADE, V. C. L. Uso do método da altura relativa em inventário florestal de um povoamento de Pinus. **Revista Árvore**, v.28, n.6, p. 865-873,2004.

MONSERUD, R. **Methodology for simulating Wisconsin northern hardwood stand dynamics**. PhD dissertation, University of Wisconsin, Madison, WI. 1975. 156 p.

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. **Revista Árvore**, v.33, n.6, p.1133-1139, 2009.

VAN LAAR, A.; AKÇA, A. **Forest Mensuration**. Göttingen, Germany: Cuvillier Verlag, 1997. 418 p.

CAPÍTULO 3

EMPREGO DE REDES NEURAS ARTIFICIAIS PARA A ESTIMAÇÃO DO VOLUME DE ÁRVORES

1. INTRODUÇÃO

O manejo florestal sustentável requer estimativas precisas de estoque de crescimento, para auxiliar os gestores florestais na análise econômica dos projetos florestais, elaboração e execução de planos de manejo. Esse estoque é expresso em m³ ou t de madeira, sendo que o volume pode ser facilmente estimado a partir de dimensões das árvores (diâmetro e altura) mensuradas no campo. O procedimento mais comum é a utilização de equações de volume com base nas relações entre volume e variáveis como o diâmetro e a altura (AKINDELE e LEMAY, 2006).

Para a obtenção de estimativas confiáveis deve-se utilizar equações compatíveis com a variação da população a ser inventariada. Nos plantios de eucalipto, o ideal é a utilização de equações específicas por clone ou espécie, espaçamento, classe de idade e regime de corte. Para a obtenção das informações para o ajuste dos modelos aconselha-se a cubagem de 100 a 150 árvores por estrato (GUIMARÃES e LEITE, 1996), podendo resultar em custo relativamente alto (OLIVEIRA et al., 2009; CAMPOS e LEITE, 2009).

Várias equações e métodos para estimação volumétrica e descrição da forma das árvores já foram propostas (BEHRE, 1923; KOZAK et al., 1969; DEMAERSCHALK, 1972; ORMEROD, 1973; MAX e BURKHART, 1976; DEMAERSCHALK e KOZAK, 1977; CAO et al., 1980; HILT, 1980; CLARK ET AL., 1991; CAMPOS e LEITE, 2009; OLIVEIRA et al., 2009). Contudo,

alguns modelos e situações podem apresentar tendenciosidades e níveis de erro inaceitáveis para os objetivos do manejo florestal.

O desenvolvimento de metodologias, abordagens ou modelos para estimar o estoque de madeira é justificável em função dos custos que esta etapa representa para o empreendimento florestal, bem como a importância do ganho em exatidão das estimativas. Novas ferramentas têm sido introduzidas na área florestal com o intuito de aumentar a exatidão das estimativas de volume e massa e reduzir os custos, dentre elas destaca-se os modelos de Redes Neurais Artificiais – RNA (GÖRGENS et al., 2009; LEITE et al., 2010).

Os modelos de RNA não necessitam de suposições sobre as relações entre as variáveis utilizadas no ajuste. Em vez disso, a rede é treinada e se encarrega de modelar as relações lineares e não lineares existentes. Diversos estudos comprovam a eficiência da aplicação de RNA na área florestal (GUAN e GERTNER, 1991; SCHMOLDT et al., 1997; BLACKARD e DEAN, 1999; ZHANG et al., 2000; LEDUC et al., 2001; LIU et al., 2003; DIAMANTOPOULOU, 2005; AVRAMIDIS et al., 2006; CORNE et al., 2004, SILVA, 2008; GÖRGENS et al., 2009; LEITE et al., 2010).

Em função do elevado custo de cubagem de árvores nas empresas florestais objetivou-se propor neste estudo um método com a utilização de RNA, para a redução do número de árvores a serem cubadas durante o processo de geração de modelos volumétricos.

2. MATERIAL E MÉTODOS

2.1 Dados

Os dados utilizados neste estudo foram provenientes de cubagens de 2700 árvores de povoamentos de clone de eucalipto. As variáveis utilizadas para estimar o volume comercial com e sem casca das árvores foram: diâmetro com casca a 1,30 m de altura (*dap*), altura total e os diâmetros com e sem casca ao longo do fuste, nas posições de 0,0, 0,5, 1,0, 1,5, 2,0 e 4,0 m.

2.2 Treinamento e Avaliação das RNA

O treinamento das RNA foi feito visando a obtenção de redes para a estimação dos volumes comerciais com e sem casca (*vcc* e *vsc*). Como variáveis de entrada utilizou-se o *dap*, a altura total (HT) e os diâmetros com e sem casca nas posições 0,0, 0,5, 1,0, 1,5, 2,0 e 4,0 m de altura (*dcc_i* e *dsc_i*) e os volumes com e sem casca obtidos até 2 e 4 m (*vcc₂*, *vsc₂*, *vcc₄* e *vsc₄*). O clone foi utilizado como variável categórica, representada pelo código do clone (C1 a C15). Testou-se duas combinações de variáveis de entrada que diferiam pela utilização do volume e diâmetro à 4m de altura (*vcc₄*, *vsc₄*, *dcc_{4,0}*, *dsc_{4,0}*).

Os dados foram divididos aleatoriamente em duas partes: 80% para treinamento das redes e 20% (540 árvores) para generalização, ou seja, aplicar as redes treinadas aos dados não utilizados no treinamento para validação. Para obter as redes neurais artificiais, utilizou-se o software Statistica 7 (STATSOFT, INC, 2007). As redes treinadas foram do tipo *perceptrons* de múltiplas camadas, comumente conhecidas como MLP (*Multilayer Perceptron*).

As melhores RNA foram selecionadas com base na correlação entre o volume observado e o volume estimado pelas redes ($r_{V\hat{V}}$) e na estabilidade dos índices de treinamento das redes fornecidos pelo software nas fases de treinamento (criação da RNA), de seleção (parada do treinamento) e de avaliação (aplicação de um conjunto de dados não utilizados no treinamento).

As redes selecionadas foram aplicadas aos dados separados para a generalização. Os valores estimados e observados foram comparados pelo procedimento estatístico proposto por Leite e Oliveira (2002). Avaliou-se também a dispersão dos erros percentuais, os histogramas de frequência dos erros percentuais e a raiz quadrada do erro médio (RMSE) (MEHTÄTALO et al., 2006):

$$RMSE(\%) = \frac{100}{\bar{V}} \sqrt{\frac{\sum_{i=1}^n (V_i - \bar{V}_i)^2}{n}}$$

em que, \bar{V} é a média dos volumes comerciais observados e n é o número total de observações.

3. RESULTADOS

A descrição das RNA obtidas e as respectivas estatísticas do treinamento e da generalização são apresentadas na Tabela 1. Na Figura 1 é apresentada a análise gráfica dos erros percentuais (dispersão e frequência) e a correlação entre valores observados e estimados. As estatísticas do procedimento estatístico L&O são apresentados na Tabela 2.

Tabela 1 – Características das redes neurais artificiais e estatísticas das estimativas do volume comercial de árvores.

Características das redes neurais artificiais				Treinamento				Generalização	
RNA	Arquitetura	Variáveis de entrada	Variável de saída	$r_{\hat{y}\hat{y}}$	Índice do treino	Índice da seleção	Índice da avaliação	$r_{\hat{y}\hat{y}}$	RMSE%
1	25-2-1	Grupo 1, dcc _{4,0} , vcc ₄	VCC	0,9982	0,06	0,06	0,06	0,9979	4,18
2	23-3-1	Grupo 1	VCC	0,9973	0,07	0,08	0,07	0,9973	4,75
3	25-2-1	Grupo 2, dsc _{4,0} , vsc ₄	VSC	0,9979	0,06	0,07	0,07	0,9976	4,41
4	23-2-1	Grupo 2	VSC	0,9968	0,08	0,08	0,08	0,9967	5,15

Grupo 1: Clone, dcc_{0,0}, dcc_{0,5}, dcc_{1,0}, dcc_{1,5}, dcc_{2,0}, dap, HT, vcc₂; Grupo 2: Clone, dsc_{0,0}, dsc_{0,5}, dsc_{1,0}, dsc_{1,5}, dsc_{2,0}, dap, HT, vsc₂.

Tabela 2 – Resultados obtidos quando aplicado o procedimento estatístico proposto pro Leite e Oliveira (2002), para as dez melhores RNA ($\alpha= 5\%$), com o resultado da avaliação da hipótese H_0 : A distribuição de diamétrica estimada é igual à projeção estimada.

Estatística	RNA 1	RNA 2	RNA 3	RNA 4
F(H0)	2,1798 ^{ns}	1,1561 ^{ns}	2,1486 ^{ns}	2,9619 ^{ns}
t(e)	1,4501 ^{ns}	1,2007 ^{ns}	1,9663 ^{ns}	0,9311 ^{ns}
$r_{y_j y_1} \geq 1 - \bar{e} $	sim	sim	sim	não
Erro médio (e)	0,0036	-0,0037	0,0053	-0,0020
Resultado	=	=	=	≠

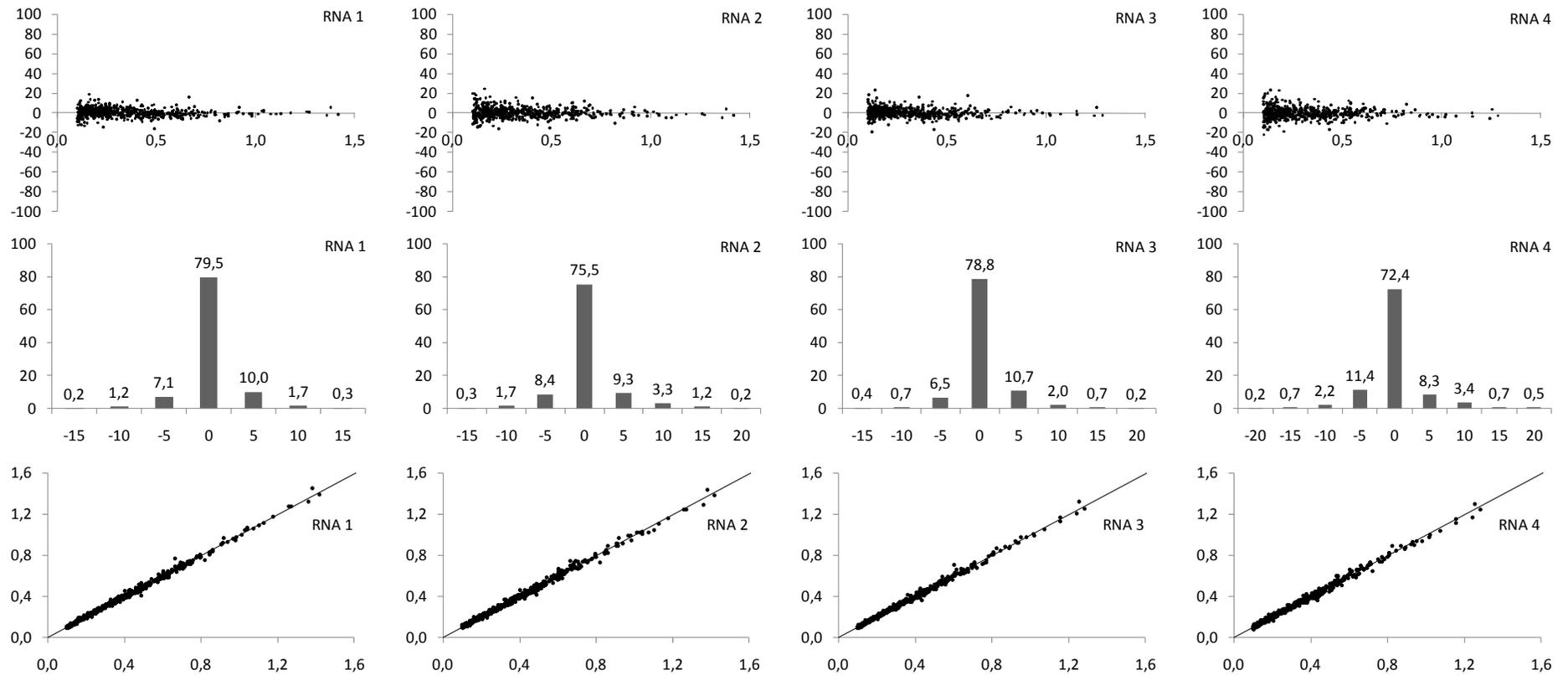


Figura 1 – Dispersão dos erros percentuais (y) em função dos volumes totais observados (x). Frequência percentual dos erros percentuais (y) em função das classes de erros percentuais (x). Relação entre valores observados (x) e estimados (y).

4. DISCUSSÃO

O presente trabalho teve como objetivo propor uma metodologia com foco principal na redução do número de medições de diâmetro ao longo do fuste de árvores-amostra para fins de cubagem de árvores em pé, ao invés de derrubá-las, mantendo os níveis de exatidão para o propósito de quantificação de estoques de crescimento em povoamentos florestais.

A amostragem para geração de equações volumétricas em florestas equiâneas é normalmente feita por espécie, espaçamento, classe de idade, tipos de solos e regime de corte (CAMPOS e LEITE, 2009). Na prática, quanto maior o número de estratos melhor a exatidão, porém maior será o custo de amostragem. Para alguns casos o número ideal de árvores-amostra para ajuste de um modelo volumétrico pode chegar a 100 ou 150 árvores (GUIMARÃES et al., 1996). Em situações práticas esse número de árvores-amostra pode ser inviável pela grande quantidade de estratos que podem ser formados (OLIVEIRA et al., 2009).

A utilização de métodos de amostragem que exigem menor número de medições de diâmetro ao longo do fuste pode reduzir o custo, mas isso só é aceitável se a confiabilidade das estimativas não sofrer interferência (WIANT et al., 1996). O modelo proposto exige somente a medição de posições da árvore que são facilmente mensuráveis em campo, sem a derrubada das árvores. Necessita-se apenas a cubagem de algumas árvores quando há a introdução de novos clones ou em situações não amostradas.

Outro método com enfoque similar ao proposto é o de similaridade de perfis, proposto por Oliveira et al. (2009). O método pode ser resumido da seguinte forma:

1) Com os dados de cubagem existentes na empresa, ajustar um modelo de taper para

todas as árvores; 2) abater eubar três árvores de clones novos, sendo uma árvore pequena, uma média e uma grande, em termos de *dap* e altura total, e ajustar o modelo de taper, para os dados dessas três árvores; 3) calcular a distância euclidiana entre os parâmetros do modelo de taper do clone novo com os parâmetros de cada árvore dos clones comerciais e determinar a árvore mais similar; e 4) determinar o estrato ao qual a árvore similar pertence e identificar a equação de volume a ser aplicada no novo clone. Este método apesar de eficiente não apresenta uma equação própria para o novo material genético introduzido, mas a equação mais similar. Caso a equação que possuir menor distância euclidiana para o novo material genético não for compatível com este, ocorrerá a adição de um viés, que pode influenciar significativamente a quantificação dos estoques volumétricos existentes.

O método proposto no presente estudo mostrou-se eficiente para a estimação do volume de árvores, sendo indicada para a obtenção dos volumes com e sem casca de povoamentos de eucalipto, possibilitando a construção de equações volumétrica com um custo relativamente baixo.

5. CONCLUSÕES

Os resultados deste estudo permitem concluir que:

- O emprego de RNA permite reduzir significativamente o número de árvores-amostra, sem perda de exatidão nas estimativas de volume por árvore.
- É possível estimar o volume de árvores de novos clones com o emprego de modelos de RNA, com exatidão.

6. REFERÊNCIAS BIBLIOGRÁFICAS

AKINDELE, S. O.; LEMAY, V. M. Development of tree volume equations for common timber species in the tropical rain forest area of Nigeria. **Forest Ecology and Management**, v.226, p.41–48, 2006.

AVRAMIDIS, S.; ILIADIS, L.; MANSFIELD, S.D. Wood dielectric loss factor prediction with artificial neural networks. **Wood Science Technol**, v.40, p.563-574, 2006.

BEHRE, C. E. Preliminary notes on studies of tree form. **Journal Forest**, v.21, p.507-511, 1923.

BLACKARD, J.; DEAN, D. Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables. **Computers and Electronics in Agriculture**, v.24, p.131-151, 1999.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3. ed. rev. e ampl. Viçosa: UFV, 2009. 548 p.

CAO, Q. V.; BURKHART, H. E.; MAX, T. A. Evaluation of two methods for cubic-volume prediction of loblolly pine to any merchantable limit. **Forest Science**, v.26, p.71-80, 1980.

CLARK, A.; SOUTER, R. A.; SCHLAEGEL, B. E. **Stem profile equations for southern tree species**. USDA For. Serv. Res. Pap.SE-282. 1991.

CORNE, S. A.; CARVER, S. J.; KUNIN, W. E.; LENNON, J. J.; VAN HEES, W. W. S. Predicting forest attributes in southeast Alaska using artificial neural networks. **Forest Science**, v.50, n.2, p.259-276, 2004.

DEMAERSCHALK, J. P. Converting volume equations to compatible taper equations. **Forest Science**, v.18, n.3, p.241-245, 1972.

DEMAERSCHALK, J.P.; KOZAK, A. The whole-bole system: a conditioned dual-equation system for precise prediction of tree profiles. **Canadian Journal of Forest Research**, v.7, p.488-497, 1977.

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

GÖRGENS, E. B.; LEITE, H. G.; SANTOS, H. N.; GLERIANI, J. M. Estimação do volume de árvores utilizando redes neurais artificiais. **Revista Árvore**, v.33, n.6, p.1141-1147, 2009.

GUAN, B.T.; GERTNER, G. Modeling red pine tree survival with an artificial neural network. **Forest Science**, v.37, n.5, p.1429-1440, 1991.

GUIMARÃES, D. P.; LEITE, H. G. Influência do número de árvores na determinação de equação volumétrica para *Eucalyptus grandis*. **Scientia Forestalis**, n.50, p. 37-42, 1996.

HILT, D.E. **Taper-based system for estimating stem volume of upland oaks**. USDA For. Serv. Res. Pap. NE-458, p. 12. 1980.

KOZAK, A.; MUNRO, D. O.; SMITH, J. H. G. Taper functions and their application in forest inventory. **The Forestry Chronicle**, v.45, p.278-283, 1969.

LEDUC, D.J.; MATNEY, T.G.; BELLI, K.L.; BALDWIN, V.C. **Predicting diameter distributions of longleaf pine plantations: a comparison between artificial neural networks and other accepted methodologies**. USDA Forest Service, Southern Research Station, Research Paper, SRS-025, Asheville, NC. 2001.

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 *Tectona grandis* Linn. trees using artificial neural networks. **European Journal of Forest Research**, v.130, n.2, p.263-269, 2011.

LEITE, H. G.; OLIVEIRA, F. H. T. Statistical procedure to test the identity of analytical methods. **Communications in soil science and plant analysis**, v.33, p.1105-1118, 2002.

LIU, C.; ZHANG, L.; DAVIS, C.J.; SOLOMON, D.S.; BRANN, T.B.; CALDWELL, L.E. Comparison of neural network and statistical methods in classification of ecological habitats using FIA data. **Forest Science**, v.49, n.4, p.619-631, 2003.

MAX, T. A.; BURKHART, H. E. Segmented polynomial regression applied to taper equations. **Forest Science**, v.22, p.283-289, 1976.

MEHTÄTALO, L.; MALTAMO, M.; KANGAS, A. The use of quantile trees in the prediction of the diameter distribution of a stand. **Silva Fennica**, v.40, n.3, p.501-516, 2006.

OLIVEIRA, M. L. R.; LEITE, H. G.; GARCIA, S. L. R.; CAMPOS, J. C. C.; SOARES, C. P. B.; SANTANA, R. C. Estimação do Volume de Árvores de Clones

de Eucalipto pelo Método da Similaridade de Perfis. **Revista Árvore**, v.33, p.133-141, 2009.

ORMEROD, D. W. A simple bole model. **The Forestry Chronicle**, v.49, p.136-138, 1973.

SCHMOLDT, D. L.; LI, P.; ABBOTT, A. L. Machine vision using artificial neural networks with local 3D neighborhoods. **Computers and Electronics in Agriculture**, v.16, p.255-271, 1997.

SILVA, M. L. M. Estimação **do diâmetro sem casca e diâmetro do cerne para árvores de *Tectona grandis* Linn. utilizando redes neurais artificiais**. 2008. 47 f. Monografia (Exigência para conclusão do curso de Engenharia Florestal) – Universidade Federal de Viçosa, Viçosa, MG, 2008.

Wiant, H. V. Jr.; Wood, G. B.; Williams, M. Comparison of three modern methods for estimating volume of sample trees using one or two diameter measurements. **Forest Ecology and Management**, v. 83, p.13-16, 1996.

ZHANG, Q. B.; HEBDA, R. I.; ZHANG, Q. J.; ALFARO, R. I. Tree-ring growth responses to climatic variables using artificial neural networks. **Forest Science**, v.46, n.2, p.229-239, 2000.

CAPÍTULO 4

PROJEÇÃO DA DISTRIBUIÇÃO DIAMÉTRICA DE POVOAMENTO EQUIÂNEOS UTILIZANDO REDES NEURAS ARTIFICIAIS

1. INTRODUÇÃO

O conhecimento da estrutura diamétrica de um povoamento florestal é de grande importância para o manejo de florestas equiâneas. Este permite inferir sobre o volume total do povoamento, sortimentos e a fase de desenvolvimento do povoamento, sendo fundamental para as análises econômicas e a simulação de regimes de manejo (CLUTTER et al., 1983; CAMPOS e LEITE, 2009).

Diferentes distribuições teóricas, tais como Beta, Weibull, Johnson SB e a função Hyper têm sido utilizadas para descrever a estrutura diamétrica de povoamentos florestais (BAILEY e DELL 1973; LOETSCH et al. 1973; HAFLEY e SCHREUDER, 1977; LEITE et al., 2010). A distribuição pode ser ajustada em termos da frequência de árvores ou da área basal por hectare (GOVE e PATIL, 1998, SIIPILEHTO, 1999). Destes, a distribuição de diâmetro com base nas frequências de árvores é a abordagem mais utilizada no Brasil.

Para projetar a distribuição, dois métodos principais podem ser aplicados: o método de predição ou projeção dos parâmetros e o método de recuperação de parâmetros (HYINK, 1980, HYINK e MOSER, 1983, GUIMARÃES, 1994). Na predição de parâmetros, os parâmetros de uma função densidade de probabilidade, por exemplo, a distribuição Weibull, são obtidos por modelos de regressão em função de características dos povoamentos (CAMPOS e LEITE, 2009). No método de recuperação de parâmetros, os parâmetros da função de distribuição são obtidos

empregando um sistema de equações fundamentados em medidas de posição. Dentre as formas de recuperação de parâmetros da distribuição destacam-se o método dos percentis da distribuição de diâmetro (BAILEY et al., 1981; CAO e BURKHART, 1984; GUIMARÃES, 1994; CAO, 2004) ou dos momentos da distribuição de diâmetro (NEWBY 1980; BURK e NEWBERRY, 1984; CAO 2004).

A computação neurobiológica tem sido empregada em alguns estudos na ciência florestal (DIAMONTOPOULOU, 2005; SILVA, 2008; GÖRGENS et al., 2009), destacando-se as redes neurais artificiais (RNA), que são processadores paralelamente distribuídos compostos por unidades de processamento simples (neurônios artificiais), que armazenam conhecimento experimental (aprendizagem) tornando-o disponível para uso (generalização) (BRAGA et al., 2000; HAYKIN, 2001; BINOTI, 2010). As RNA se aplicam principalmente aos problemas de aproximação de função, classificação de padrões, agrupamento de dados, predição (séries temporais), otimização, recuperação de conteúdo e controle (JAIN et al., 1996; HAYKIN, 2001, BINOTI, 2010).

A projeção de parâmetros de uma função densidade de probabilidade que descreva a estrutura diamétrica de um povoamento florestal se enquadra na tarefa de aproximação de função, que consiste em projetar uma rede neural que aproxime a função desconhecida $f(x)$, que descreve o mapeamento dos pares de entrada-saída $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ de um conjunto de n padrões de treinamento.

O objetivo deste estudo foi desenvolver e avaliar a aplicação de redes neurais artificiais para a projeção de parâmetros da distribuição Weibull, bem como comparar o método proposto com modelos de distribuição diamétrica comumente utilizados na área florestal.

2. MATERIAL E MÉTODOS

2.1 Dados

Os dados utilizados nesse estudo foram obtidos de aproximadamente 1250 parcelas permanentes retangulares de 340 m², instaladas e medidas em povoamentos de híbridos de *Eucalyptus grandis* x *Eucalyptus urophylla*, na região Centro Oeste do Estado de Minas Gerais. Esses povoamentos foram estabelecidos sob arranjo espacial de 3 x 3 m, sendo a madeira utilizada para produção de polpa de celulose de fibra curta, com idade de corte em torno dos 7 anos. Nesta idade a produtividade média desses povoamentos variou de 25 a 50 m³.ha⁻¹.ano⁻¹. As medições de diâmetros foram efetuadas nas árvores com *dap* acima de 5 cm, nas idades médias de 28, 40, 52, 64, 76, 88, 100, e 112 meses. O banco de dados foi separado em dois conjuntos de dados, sendo um para construção do modelo e o outro para sua validação. Nessa separação foi mantida uma frequência homogênea de dados por classe de capacidade produtiva, previamente definida empregando curvas de índices de local.

2.2 Estrutura Diamétrica

Para a construção dos modelos de distribuição de diâmetros, ajustou-se a função Weibull para todas as parcelas e ocasiões:

$$f(x) = \frac{\gamma}{\beta} \left(\frac{x}{\beta}\right)^{\gamma-1} e^{-\left(\frac{x}{\beta}\right)^{\gamma}}$$

em que, β parâmetro de escala ($\beta > 0$), γ é o parâmetro de forma ($\gamma > 0$), e x é o centro de classe de diâmetro ($x > 0$).

Para realização dos ajustes, os dados de distribuição diamétrica observados em cada parcela para todas as medições, foram agrupados em classes com amplitude

do 1,0 cm. Procedeu-se o ajuste das funções pelo Método da Máxima Verossimilhança (BAILEY e DELL, 1973).

Testou-se a aderência das funções aos dados pelo teste de Kolmogorov-smirnov (GIBBONS e SUBHABRATA, 1992; SOKAL e ROHLF, 1981). Este teste compara a frequência acumulada estimada, com a observada sendo a classe de maior divergência a estatística do teste (dn). O teste foi aplicado a cada ajuste com nível de significância 1%. Neste teste quanto maior o valor de dn , maior é a divergência entre os valores reais e os valores estimados pela função.

2.3 Treinamento e Generalização das redes neurais artificiais

O treinamento de uma RNA é o processo de ajuste de seus pesos através de um algoritmo de aprendizagem, que extrai características dos dados fornecidos e tem por objetivo gerar uma rede que desempenhe a tarefa de interesse (BINOTI, 2010).

Foram treinadas redes para projetar o parâmetro de forma, parâmetro de escala, diâmetro mínimo e diâmetro máximo. Os parâmetros e características do povoamento foram projetados em função das seguintes características: idade atual e futura, altura e diâmetro mínimo, máximo, médio, diâmetro quadrático e altura dominante, bem como os parâmetros da função na idade atual. Nesta etapa, foram utilizados dados de aproximadamente 5000 parcelas de inventário.

Foram treinadas 100 redes na estrutura de *Perceptrons de Múltiplas Camadas* (MLP) para projeção dos parâmetros de interesse. Nessa etapa do estudo utilizou-se a ferramenta *Intelligent Problem Solver* (IPS) do *software* Statística 7 (STATSOFT, INC, 2007). Diante do grande número de possíveis combinações das variáveis de entrada disponíveis foi solicitada a otimização desta seleção por algoritmos do IPS. Contudo para fins de operacionalização da técnica proposta, selecionou-se somente 4 redes, sendo uma com as variáveis mínimas necessárias (variável na idade atual e idades), duas com as variáveis selecionadas pelo otimizador e uma com a inclusão de todas as variáveis disponíveis.

A generalização é a capacidade de uma rede neural produzir saídas adequadas para entradas que não estavam presentes durante o treinamento (aprendizagem). Nesta etapa, foram utilizados dados de aproximadamente 5000 parcelas de inventário não utilizados durante o treinamento.

2.4 Modelo de Distribuição de Diâmetros

O modelo de distribuição diamétrica utilizados neste trabalho foi proposto por Nogueira et al., (2005) e é composto pelos seguintes modelos:

$$\beta_2 = \beta_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) dmax_2 + e$$

$$Ln\gamma_2 = Ln\gamma_1 e^{-(\alpha_0)(I_2^{\alpha_1} - I_1^{\alpha_1})} + e$$

$$dmax_2 = dmax_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) \beta_1 + e$$

$$dmin_2 = dmin_1 e^{-(\alpha_0)(I_2^{\alpha_1} - I_1^{\alpha_1})} + e$$

em que, I_1 e I_2 são as idades atual e futura, respectivamente, em meses; γ_1 e γ_2 são os parâmetros de forma da função Weibull nas idades atual e futura; β_1 e β_2 são os parâmetros de escala da função Weibull nas idades atual e futura; $dmax_1$ e $dmax_2$ são os diâmetros máximos nas idades atual e futura, em cm; $dmin_1$ e $dmin_2$ são os diâmetros mínimos nas idades atual e futura, em cm; e Ln indica o logaritmo neperiano e $e \sim NID(\Theta, \sigma^2)$.

Após o ajuste da função Weibull para cada parcela em cada medição procedeu-se o ajuste das funções que compõe o modelo de distribuição diamétrica (MDD) pelo método iterativo Quasi-Newton, empregando o software Statistica 9 (STATSOFT, INC, 2010). A distribuição diamétrica foi projetada a partir de uma distribuição em uma idade atual para uma idade futura. A primeira projeção foi feita a partir da média das frequências observadas em cada classe de diâmetro. As demais projeções foram feitas sobre as projeções imediatamente anteriores. Em função da baixa mortalidade apresentada nos dados do estudo não realizou-se a projeção do número de árvores, sendo este considerado constante.

2.5 Avaliação das projeções das Distribuições Diamétricas

A redistribuição teórica dos diâmetros foi avaliada pela capacidade das redes neurais e do sistema de equações em projetar o diâmetro máximo e mínimo e os parâmetros da função Weibull para qualquer idade. O coeficiente de correlação entre os valores observados e estimados e a análise gráfica dos erros percentuais foram utilizados para avaliar a precisão das estimativas de frequência por classe de diâmetro. Foi avaliada a consistência, a capacidade de projeção e o realismo biológico do sistema e das RNA, projetando diferentes distribuições iniciais

observadas, para idades futuras, e comparando com os respectivos valores observados de frequência por classe de diâmetro.

A avaliação das estimativas geradas pelas RNA, nas três etapas, treinamento, generalização e prognose, e pelo modelo de distribuição diamétrica foi feita empregando estatísticas e análise gráfica dos erros percentuais em relação aos valores observados, sendo:

$$erro_{\%} = \frac{(\hat{Y} - Y)}{Y} \times 100$$

em que, Y é a saída observada nos inventários, \hat{Y} é a saída estimada.

As estatísticas empregadas foram a correlação entre os valores estimados e observados dos parâmetros projetados e a raiz quadrada do erro médio (RMSE%).

$$r_{Y\hat{Y}} = \frac{cov(Y, \hat{Y})}{\sqrt{s^2(Y)s^2(\hat{Y})}}$$

em que, s^2 é a variância e cov é a covariância.

$$RMSE(\%) = 100 \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}} / \bar{Y}$$

em que, \bar{Y} é a média da saída observada e n é o número total de dados.

3. RESULTADOS

Ajustou-se a função Weibull para todas as parcelas e em cada situação, sendo cada estimação comparada com a distribuição observada. Todos os ajustes resultaram em aderência aos dados pelo teste Kolmogorov-Smirnov ($P > 0,01$).

As características das 16 redes treinadas, bem como as estatísticas utilizadas na avaliação são apresentadas na Tabela 1. Dentre as redes treinadas escolheu-se uma rede para a projeção de cada variável de interesse e a análise gráfica de erros percentuais: β_2 (RNA 4), $Ln(y_2)$ (RNA 8), $dmin_2$ (RNA 12) e $dmax_2$ (RNA 16). Os gráficos de dispersão e histogramas de erros (%) são apresentados na Figura 1.

Tabela 1 – Características e estatísticas das redes neurais artificiais treinadas para projetar os parâmetros da função densidade de probabilidade Weibull e as variáveis do povoamento.

RNA	Arquitetura *	r	Índice do treino **	Índice da seleção **	Índice da validação **	Variáveis de entrada	Variável de saída	RMSE%
1	3-9-1	0,9650	0,27	0,26	0,26	β_1, I_1, I_2	β_2	3,30
2	9-5-1	0,9724	0,23	0,23	0,24	$\beta_1, q_1, I_1, Hmin_1, Hd_1, Hmed_1, Hmax_1, dapmax_1, N_1, I_2$	β_2	1,71
3	11-5-1	0,9720	0,23	0,23	0,25	$\beta_1, \gamma_1, q_1, I_1, Hmin_1, Hmed_1, Hmax_1, dapmin_1, dapmax_1, N_1, I_2$	β_2	0,97
4	13-13-1	0,9728	0,23	0,24	0,23	$\beta_1, \gamma_1, q_1, I_1, Hmin_1, Hd_1, Hmed_1, Hmax_1, dapmin_1, dapmed_1, dapmax_1, N_1, I_2$	β_2	1,21
5	3-9-1	0,9541	0,29	0,31	0,30	$Ln(\gamma_1), I_1, I_2$	$Ln(\gamma_2)$	7,24
6	9-7-1	0,9549	0,30	0,30	0,30	$Ln(\gamma_1), q_1, I_1, Hd_1, Hmed_1, Hmax_1, dapmax_1, N_1, I_2$	$Ln(\gamma_2)$	7,17
7	10-6-1	0,9566	0,29	0,29	0,29	$Ln(\gamma_1), q_1, Hmin_1, Hd_1, Hmed_1, Hmax_1, dapmin_1, dapmax_1, N_1, I_2$	$Ln(\gamma_2)$	7,04
8	13-13-1	0,9581	0,28	0,29	0,29	$\beta_1, Ln(\gamma_1), q_1, I_1, Hmin_1, Hd_1, Hmed_1, Hmax_1, dapmin_1, dapmed_1, dapmax_1, N_1, I_2$	$Ln(\gamma_2)$	6,92
9	3-6-1	0,9414	0,33	0,32	0,33	$I_1, dapmin_1, I_2$	$dapmin_2$	8,71
10	7-6-1	0,9492	0,31	0,32	0,33	$q_1, I_1, Hd_1, dapmin_1, dapmed_1, N_1, I_2$	$dapmin_2$	8,13
11	8-3-1	0,9498	0,30	0,32	0,32	$I_1, Hd_1, Hmed_1, Hmax_1, dapmin_1, dapmax_1, N_1, I_2$	$dapmin_2$	8,08
12	11-12-1	0,9512	0,31	0,32	0,29	$q_1, I_1, Hmin_1, Hd_1, Hmed_1, Hmax_1, dapmin_1, dapmed_1, dapmax_1, N_1, I_2$	$dapmin_2$	7,97
13	3-9-1	0,9772	0,21	0,21	0,22	$I_1, dapmax_1, I_2$	$dapmax_2$	3,46
14	6-7-1	0,9804	0,20	0,19	0,20	$q_1, I_1, Hd_1, Hmax_1, dapmax_1, I_2$	$dapmax_2$	3,20
15	7-8-1	0,9806	0,20	0,19	0,20	$I_1, Hmed_1, Hmax_1, dapmin_1, dapmax_1, N_1, I_2$	$dapmax_2$	3,19
16	11-12-1	0,9811	0,19	0,20	0,20	$q_1, I_1, Hmin_1, Hd_1, Hmed_1, Hmax_1, dapmin_1, dapmed_1, dapmax_1, N_1, I_2$	$dapmax_2$	3,15

* Número de neurônios em cada camada. ** Índices do treino (obtenção da rede), seleção (parada do treino) e validação (qualidade da rede treinada) que devem ser estáveis, ou seja, pouca variação entre eles. β_1 = parâmetro β da função Weibull em I_1 , γ_1 = parâmetro γ da função Weibull em I_1 , q_1 = diâmetro médio em I_1 , I_1 = idade inicial, $Hmin_1$ = altura mínima em I_1 , Hd_1 = altura dominante em I_1 , $Hmed_1$ = altura média em I_1 , $Hmax_1$ = altura máxima em I_1 , $dapmin_1$ = dap mínimo em I_1 , $dapmed_1$ = média do dap em I_1 , $dapmax_1$ = dap máximo em I_1 , N_1 = número de árvores/ha em I_1 , I_2 = idade futura, β_2 = parâmetro β da função Weibull em I_2 , γ_2 = parâmetro γ da função Weibull em I_2 , r = correlação entre os valores da variável estimada e observada.

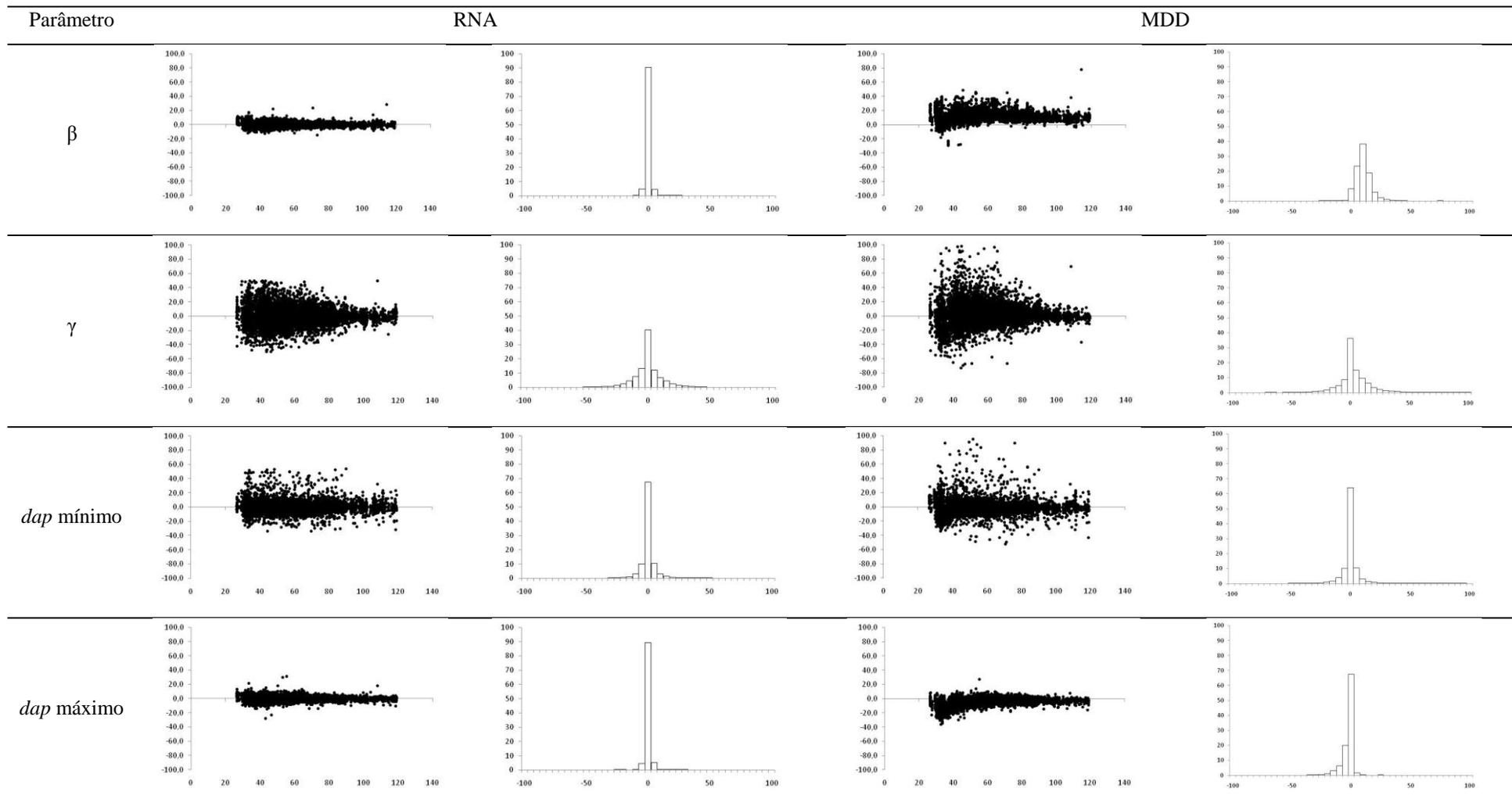


Figura 1 – Dispersão dos erros percentuais das estimativas obtidas pelas redes neurais artificiais (y) em função da idade (x).

O teste t foi significativo para todos os parâmetros em todas as equações que compõe o modelo de distribuição diamétrica. Todas as equações apresentaram valores relativamente altos de coeficientes de correlação, indicando que as variáveis independentes utilizadas nas equações têm grande influência nas variáveis dependentes. A análise gráfica dos erros percentuais permitiu inferir a inexistência de tendenciosidade nos ajustes (Figura 1). Os modelos ajustados e seus respectivos coeficientes de correlação e RMSE(%) foram:

$$\beta_2 = \beta_1 \left(\frac{I_1}{I_2} \right) + 1,2567 \left(1 - \frac{I_1}{I_2} \right) dmax_2 \quad r_{Y\hat{Y}} = 0,9640 \quad RMSE(\%) = 15,02$$

$$Ln\gamma_2 = Ln\gamma_1 e^{-(2456,3467)(I_2^{0,0002} - I_1^{0,0002})} \quad r_{Y\hat{Y}} = 0,8982 \quad RMSE(\%) = 29,21$$

$$dmax_2 = dmax_1 \left(\frac{I_1}{I_2} \right) + 4,457 \left(1 - \frac{I_1}{I_2} \right) \beta_1 \quad r_{Y\hat{Y}} = 0,9438 \quad RMSE(\%) = 4,95$$

$$dmin_2 = dmin_1 e^{(1354,7867)(I_2^{0,0001} - I_1^{0,0001})} \quad r_{Y\hat{Y}} = 0,8841 \quad RMSE(\%) = 8,78$$

4. DISCUSSÃO

Apesar da existência de diversas funções estatísticas, de diferentes métodos para ajuste dessas funções e de alternativas para estimação de distribuições diamétricas futuras, o enfoque de modelagem tem sido o mesmo, com predomínio de uma das formas da função Weibull. Dentre outras características, a boa correlação apresentada por seus parâmetros com características do povoamento tem sido fundamental para a utilização desta função (GUIMARÃES, 1994).

O processo de modelagem das distribuições diamétricas envolve basicamente a análise das alterações nos coeficientes, ou medidas de posição da função de distribuição, as quais são determinadas pelo crescimento do povoamento florestal. De modo geral, os modelos usuais de distribuição de diâmetros têm apresentado resultados satisfatórios, apenas no que concerne à estimativa do parâmetro de escala. Estudos na área de modelagem da distribuição diamétrica demonstram que o parâmetro de forma apresenta baixa correlação com características do povoamento (GUIMARÃES, 1994).

No presente estudo foi proposto um método alternativo para modelagem de distribuições diamétricas que não envolve o uso de modelos de regressão para a projeção da distribuição diamétrica em idades futuras. A proposta consiste na projeção dos parâmetros da distribuição diamétricas para idades futuras em função de características do povoamento empregando redes neurais artificiais.

A principal vantagem do método proposto se baseia na eficiência da modelagem de relações não lineares que podem existir entre as características do povoamento e os parâmetros da função densidade de probabilidade (fdp). A Figura 1 demonstra a superioridade da modelagem com RNA sobre as equações que compõe

o modelo de distribuição diamétrica. Conforme resultado desse estudo, o enfoque de modelagem resulta em estimativas livres de bias.

Na Tabela 2 são resumidos os percentuais de casos onde o erro de projeção ficou entre os limites de $\pm 5\%$ e $\pm 10\%$. Comprovando a maior eficiência e exatidão ao empregar RNA.

As estatísticas de precisão obtidas neste estudo são superiores àquelas obtidas em muitos modelos clássicos de distribuição de diâmetros e demonstram a eficiência do método proposto. Como proposta de trabalhos futuros sugere-se a incorporação de características edafoclimáticas que apesar das relações aparentemente desconhecidas em relação à variação dos parâmetros da função Weibull podem ser modeladas pelas RNA.

5. CONCLUSÃO

Redes neurais artificiais são mais eficientes do que os modelos de distribuição de diâmetros usuais.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- BAILEY, R. L.; ABERNATHY, N. C.; JONES, E. P. **Diameter distributions models for repeatedly thinned slash pine plantations**. USDA Forest Service Gen. Tech. Rep. SO-34. 1981.
- BAILEY, R. L.; DELL, T. R. Quantifying diameter distributions with the Weibull function. **Forest Science**, v.19, n.2, p.97-104, 1973.
- BINOTI, M. L. M. DA S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. 2010. 54 f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2010.
- BRAGA, A. P.; LUDERMIR, T. B.; CARVALHO, A. P. L. F. **Redes Neurais Artificiais: Teoria e Aplicações**. 2 ed. Rio de Janeiro: LTC, 2007.
- BURK, T.; NEWBERRY, J. A simple algorithm for moment-based recovery of Weibull distribution parameters. **Forest Science**, v.30, n.2, p.329-332, 1984.
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: Perguntas e respostas**. 3 ed. rev. e ampl. Viçosa: UFV, 2009.
- CAO, Q. V. Predicting parameters of a Weibull function for modeling diameter distribution. **Forest Science**, v.50, n.4, p.682-685, 2004.
- CAO, Q.; BURKHART, H. A segmented distribution approach for modeling diameter frequency data. **Forest Science**, v.30, n.1. p.129-137, 1984.
- CLUTTER, J. L.; FORTSON, J. C.; PIENAAR, L. V.; BRISTER, R. G. H.; BAILEY, R. L. **Timber management: a quantitative approach**. New York: John Willey e Sons, 1983. 333 p.
- DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. **Computers and electronics in agriculture**, v.10, p.235-244, 2005.

GIBBONS, J. D.; SUBHABRATA, C. **Nonparametric statistical inference**. 3th. ed. New York: Marcel Dekker, 544p. (Statistics: textbook and monograph, v.31). 1992.

GÖRGENS, E. B.; LEITE, H. G.; SANTOS, H. N.; GLERIANI, J. M. Estimação do volume de árvores utilizando redes neurais artificiais. **Revista Árvore**, v.33, n.6, p. 1141-1147, 2009.

GOVE, J. H.; PATIL, G. P. Modelling basal area size distribution of forest stands: a compatible approach. **Forest Science**, v.44, n.2, p.285-297, 1998.

GUIMARÃES, D. P. **Desenvolvimento de um modelo de distribuição diamétrica de passo invariante para prognose e projeção da estrutura de povoamentos de eucalipto**. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa. 1994.

HAFLEY, W. L.; SCHREUDER, H. T., Statistical distributions for fitting diameter and height data in even-aged stands. **Canadian Journal of Forest Research**, v.7, p.184-487, 1977.

HAYKIN, S. **Redes Neurais: princípios e prática**. Porto Alegre: Bookman, 2001.

HYINK, D. M. Diameter distribution approaches to growth and yield modeling. In: BROWN, K. M., CLARKE, F.R. (Eds.), **Forecasting Forest Stand Dynamics**. School of Forestry. Lakehead University, p.138-163, 1980.

HYINK, D. M.; MOSER, J. W. A generalized framework for projecting forest yield and stand structure using diameter distributions. **Forest Science**, v.29, p.85-95, 1983.

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. M. Artificial Neural Networks: A Tutorial. **Computer**, v.29, n.3, p.31-44, 1996.

LEITE, H. G.; BINOTI, D. H. B.; GUIMARÃES, D. P.; SILVA, M. L. M.; GARCIA, S. L. R. Avaliação do ajuste das funções Weibull e hiperbólica a dados de povoamentos de eucalipto submetidos a desbaste. **Revista Árvore**, v.34, n.2, p.305-311, 2010.

LOETSCH, F.; ZOHRER, F.; HALLER, R. E. **Forest inventory**. München, BLV Verlagsgesellschaft, 1973. v.2. 472p.

NEWBY, M. The properties of moment estimators for the Weibull distribution based on the sample coefficient of variation. **Technometrics**, v.22, p.187-194, 1980.

NOGUEIRA, G. S.; LEITE, H. G.; CAMPOS, J. C. C.; CARVALHO, A. F.; SOUZA, A. L. Modelo de distribuição diamétrica para povoamentos de *Eucalyptus* sp. submetidos a desbaste. **Revista Árvore**, v.29, n.4, p.579-589, 2005.

SIIPILEHTO, J. Improving the accuracy of predicted basal-area diameter distribution in advanced stands by determining stem number. **Silva Fennica**, v.33, n.4, p.281-301, 1999.

SILVA, M. L. M. **Estimação do diâmetro sem casca e diâmetro do cerne para árvores de *Tectona grandis* linn. utilizando redes neurais artificiais.** Monografia. Universidade Federal de Viçosa. 2008.

SOKAL, R. R.; ROHLF, F. J. **Biometry.** San Francisco: Freeman.1981.

CAPÍTULO 5

MODELAGEM DA DISTRIBUIÇÃO DE DIÂMETROS UTILIZANDO AUTÔMATOS CELULARES E REDES NEURAI ARTIFICIAIS

1. INTRODUÇÃO

A projeção da distribuição diamétrica é essencial para o manejo de florestas quando a madeira é destinada a multiprodutos. Essa projeção pode ser feita empregando modelos que estimam a distribuição de diâmetros a partir de uma distribuição atual ou de características do povoamento, permitindo simular desbastes e quantificar multiprodutos da madeira (KNOBELL, 1986; CLUTTER et al., 1983; NOGUEIRA, 2005; SIIPILEHTO, 2007). A característica comum nesse tipo de modelagem é a presença de uma função densidade de probabilidade (*fdp*).

Diferentes tipos de distribuições estatísticas já foram empregadas para descrever a estrutura diamétrica de povoamentos florestais, como: Gama (NELSON, 1964), log-normal (BLISS e REINKER, 1965), Beta (CLUTTER e BENNETT, 1965; PALAHÍ et al., 2007), Johnson's SB (HAFLEY e SCHUREUDER, 1977 e PALAHÍ et al., 2007) e a distribuição Weibull (WEIBULL, 1951; BAILEY e DELL, 1973; PALAHÍ et al., 2007). Desde 1973, a função Weibull tem sido amplamente difundida e utilizada na área florestal (CLUTTER e ALLISON, 1974; HAFLEY e SCHUREUDER 1977; MATNEY e SULLIVAN, 1982; KNOWE et al., 1997; CAO, 2004; NOGUEIRA et al., 2005; PALAHÍ et al., 2006; PALAHÍ et al., 2007).

Essas funções podem ser ajustadas pelos métodos dos momentos (FRAZIER, 1981), dos percentis (BAILEY et al., 1989), da máxima verossimilhança (FISHER, 1922), por combinação dos métodos dos momentos e dos percentis (BALDWIN e

FEDUCCIA, 1987), empregando a meta heurística *simulated annealing* (ABBASI et al., 2006) ou empregando redes neurais artificiais (ABBASI et al., 2008).

A estimação da distribuição diamétrica futura pode ser feita empregando modelos de predição ou modelos de projeção. No primeiro caso, os parâmetros da distribuição estatística são considerados como variáveis dependentes, em modelos de regressão onde as variáveis preditoras são características do povoamento (por exemplo, idade, diâmetros médio e máximo e índice de local). No segundo, os parâmetros observados numa idade atual também são incluídos como variáveis preditoras. Em 1994, Guimarães propôs um modelo de passo invariante, obtendo resultados compatíveis para a projeção da distribuição diamétrica a partir de quaisquer inventários florestais iniciais. A eficiência do modelo proposto por esse autor foi comprovada por Soares et al. (2007), com dados de povoamentos de eucalipto. Cao (2004) propôs os métodos da regressão de máxima verossimilhança e da regressão da função de distribuição acumulativa para estimar os parâmetros da distribuição estatística. Nogueira (2005) propôs um modelo de projeção da distribuição diamétrica através de um sistema de equações, onde os parâmetros da função Weibull em uma idade futura foram correlacionados com esses mesmos parâmetros em uma idade atual, resultando em um sistema de projeção compatível. Adaptações e aplicações dos diferentes métodos podem ser encontrados em Burkhart (1971), Päivinen (1980), Cao et al. (1982), Magnussen (1986), Gove e Patil (1998), Kangas e Maltamo (2000), Newton et al. (2005), Mehtätalo et al. (2006) e Nord-Larsen e Cao (2006).

Apesar da existência de diversas funções estatísticas, de diferentes métodos para ajuste dessas funções e de alternativas para estimação de distribuições diamétricas futuras, o enfoque de modelagem tem sido o mesmo, com predomínio de uma das formas da função Weibull. Na maioria das vezes os parâmetros desta função têm sido correlacionados com atributos dos povoamentos por meio de regressão. O presente estudo apresenta uma abordagem diferente daquelas usuais, envolvendo emprego de métodos de inteligência e de vida artificial ao invés da análise de regressão.

Técnicas da inteligência artificial, como redes neurais artificiais (DIAMONTOPOULOU, 2005; SILVA, 2008; GÖRGENS et al., 2009), e de vida artificial, como autômatos celulares (BERGER e HILDENBRANDT, 2000; BONE

et al., 2005; MARTHEY et al., 2007), tem sido utilizadas em algumas aplicações na área florestal.

Redes neurais artificiais (RNA) são processadores paralelamente distribuídos compostos por unidades de processamento simples (neurônios artificiais), que armazenam conhecimento experimental (aprendizagem) tornando-o disponível para uso (generalização) (BRAGA et al., 2007; HAYKIN, 2001). Autômatos celulares (AC) são sistemas de interações locais, discretos no tempo e no espaço, compostos de células, onde o estado de cada célula é resultado de seu estado atual, da sua regra de transição e da sua interação com células vizinhas (WOLFRAM, 1983; ILACHINSKI, 2001). As células são componentes simples que interagem localmente e a complexidade em âmbito global advém de regras entre essas unidades mais simples. Diversos métodos podem ser utilizados como regras de transição, com destaque para RNA (GARZON, 1995).

A modelagem empregando AC tem sido feita em vários tipos de estudos, envolvendo fenômenos químicos (d'HUMIERES e LALLEMAND, 1986; GERHARDT e SCHUSTER, 1989), crescimento vegetal (LINDEMAYER e ROZENBERG, 1976), crescimento de cristais (KESSLER, 1990), ecologia (PHIPS, 1992), propagação de doenças infecciosas (SEGEL, 1999), dinâmica social (EPSTEIN e AXTELL, 1996) e incêndios florestais (BAK et al., 1990).

O objetivo deste estudo foi construir, testar e avaliar um modelo de autômatos celulares, com sua regra de evolução baseada em redes neurais artificiais, para a projeção da distribuição de diâmetros em povoamentos de eucalipto.

2. MATERIAIS E MÉTODOS

2.1 Dados

Os dados utilizados nesse estudo foram obtidos de 400 parcelas permanentes retangulares de 340 m², instaladas em povoamentos de híbridos de *Eucalyptus grandis* x *Eucalyptus urophylla*, na região Centro Oeste do Estado de Minas Gerais. Esses povoamentos foram estabelecidos sob arranjo espacial de 3 x 3 m, com rotação regulatória de 7 anos, e produtividade média variando de 25 nos piores locais a 50 m³ha⁻¹ano⁻¹ nos melhores na idade de corte. As medições de diâmetros foram efetuadas nas árvores com *dap* acima de 5 cm, nas idades médias de 28, 40, 52, 64 e 76 meses. O banco de dados foi separado em dois conjuntos, sendo um para construção do modelo e o outro para sua validação, sendo mantida uma frequência homogênea de dados por classe de capacidade produtiva nos dois casos.

2.2 Modelo de Autômatos Celulares

Para um modelo unidimensional de AC, o valor da *i*^a célula no tempo *t*, denominado por *c_i(t)*, evolui de acordo com a regra F que é função do *c_i(t)* e das outras células que se encontram dentro da extensão *r* (para ambas as direções) de *c_i(t)* (ILACHINSKI, 2001), isto é:

$$C_i(t) = F(C_{i-r}(t-1), \dots, C_{i+r}(T+1))$$

O AC proposto neste estudo é unidimensional sendo que cada célula possui 4 vizinhos (*r* = 2). As células representam uma classe de diâmetro de 2 cm de amplitude, sendo que o estado de cada célula corresponde à probabilidade de ocorrência de árvores nesta classe.

Para a construção da regra de evolução do AC foram treinadas redes neurais artificiais, utilizando as seguintes variáveis de entrada: estado atual da célula, estado das células contidas na extensão da sua vizinhança e as idades atual e futura. Como variável de saída foi considerado o estado futuro da célula (Figura 1). Após a normalização dos dados foram testadas diferentes arquiteturas e tipos de redes, sendo selecionadas as que melhor representaram os dados. Foram testadas redes de função de base radial (RBF - *Radial Basis Function*), perceptrons de múltiplas camadas (MLP - *Multilayer perceptron*) e perceptrons de camada única (*Perceptron*), com diferentes arquiteturas, conforme Haykin (2001). A projeção da densidade (número de árvores por hectare) foi feita empregando RNA, tendo como variáveis de entrada a densidade atual, a idade atual e a idade futura. O *software* Statistica 7 (STATSOFT, Inc, 2007) foi utilizado nessa etapa do estudo.

2.3 Aplicação e validação do modelo AC

O modelo de AC foi empregado para projetar a distribuição de diâmetros para todas as idades de medição. Gráficos da correlação entre valores observados e estimados foram feitos para todas essas projeções e o procedimento estatístico proposto por Leite e Oliveira (2002) foi utilizado para avaliar a exatidão das projeções, a 5% de significância.

3. RESULTADOS E DISCUSSÃO

3.1 Regra de evolução

Cerca de 50 redes diferentes foram ajustadas, escolhendo-se as dez que melhor representaram o realismo biológico envolvido na evolução da distribuição diamétrica. Essas dez redes foram utilizadas para o desenvolvimento do estudo, sendo os melhores resultados foram obtidos com as redes MLP (Tabela 1).

Tabela 1 – Principais características das dez melhores redes neurais artificiais treinadas para avaliar a evolução da distribuição diamétrica de povoamentos de eucalipto.

Identificação	Arquitetura	Número de neurônios			Função de ativação	
		Camada de entrada	Camada oculta	Camada de saída	Camada oculta	Camada de saída
Rede-1	MLP	7	7	1	Exponencial	Logística
Rede-2	MLP	7	9	1	Logística	Seno
Rede-3	MLP	7	10	1	Logística	Identidade
Rede-4	MLP	7	12	1	Logística	Logística
Rede-5	MLP	7	10	1	Logística	Exponencial
Rede-6	MLP	7	7	1	Exponencial	Identidade
Rede-7	MLP	7	7	1	Exponencial	Identidade
Rede-8	MLP	7	8	1	Exponencial	Hiperbólica
Rede-9	MLP	7	6	1	Logística	Hiperbólica
Rede-10	MLP	7	12	1	Hiperbólica	Exponencial

3.2 Recuperação da distribuição diamétrica

Com as dez melhores redes obtidas foi feita a projeção da distribuição diamétrica nos dados destinados à validação do modelo. Resultados da aplicação do procedimento proposto por Leite e Oliveira (2002) são apresentados na Tabela 2. Esses resultados demonstram que as regras de evolução representaram a dinâmica do fenômeno estudado com eficiência e exatidão.

Tabela 2 – Resultados obtidos pela aplicação procedimento estatístico proposto por Leite e Oliveira (2002), para as dez melhores RNA ($\alpha= 5\%$), com o resultado da avaliação da hipótese H_0 : A distribuição de diamétrica observada é igual à projeção.

	Rede - 1	Rede - 2	Rede - 3	Rede - 4	Rede - 5	Rede - 6	Rede - 7	Rede - 8	Rede - 9	Rede - 10
F(H_0)	1,1780 ^{ns}	1,3505 ^{ns}	0,9479 ^{ns}	1,0291 ^{ns}	0,7727 ^{ns}	0,4837 ^{ns}	0,5619 ^{ns}	1,5080 ^{ns}	0,7352 ^{ns}	0,7248 ^{ns}
t(e)	0,6743 ^{ns}	1,3772 ^{ns}	0,3978 ^{ns}	1,1972 ^{ns}	0,7198 ^{ns}	0,7234 ^{ns}	0,0159 ^{ns}	0,7682 ^{ns}	0,6847 ^{ns}	0,6207 ^{ns}
$r_{y,y_1} \geq 1 - \bar{e} $	sim	sim	sim	sim	sim	sim	não	não	não	sim
r ²	0,8775	0,9470	0,9050	0,9224	0,8927	0,9437	0,9034	0,8644	0,8826	0,9107
RMS	0,0033	0,0015	0,0026	0,0023	0,0031	0,0015	0,0025	0,0032	0,0031	0,0020
Erro médio (e)	0,0837	0,1357	0,0490	-0,1408	0,0911	-0,0613	0,0009	0,0636	0,0501	0,0479
Resultado	=	=	=	=	=	=	≠	≠	≠	=

Escolheu-se uma parcela ao acaso para simulação detalhada da dinâmica de evolução da distribuição de diâmetros (Tabela 3). Como exemplo, seja a célula correspondente a classe de *dap* 10 cm no tempo $t = 0$ (27 meses). Esta célula apresentava o valor de 0,54 (probabilidade), sendo que suas vizinhas apresentavam os valores de 0,02; 0,11; 0,32 e 0,01, representando respectivamente os valores das classes 6, 8, 12 e 14. No instante $t = 1$ (39 meses), o valor dessa célula (célula correspondente à classe de *dap* 10 cm) é de 0,11. Esse resultado advém da aplicação da regra de transição sobre o estado atual do sistema no instante $t = 0$. A aplicação da regra ocorre discretamente no tempo, o que pode ser observado para as outras classes da distribuição.

Tabela 3 – Probabilidade de ocorrência de árvores (estado) por classes de *dap* em cinco medições de uma parcela em um povoamento de eucalipto.

Classe	4	6	8	10	12	14	16	18	20	22	24	26
Probabilidade observada												
t = 0	0,00	0,02	0,11	0,54	0,32	0,01	0,00	0,00	0,00	0,00	0,00	0,00
t = 1	0,00	0,01	0,03	0,11	0,36	0,41	0,13	0,01	0,00	0,00	0,00	0,00
t = 2	0,00	0,01	0,01	0,05	0,15	0,35	0,35	0,08	0,01	0,00	0,00	0,00
t = 3	0,00	0,01	0,01	0,03	0,10	0,23	0,36	0,18	0,04	0,01	0,00	0,00
t = 4	0,00	0,01	0,01	0,03	0,09	0,21	0,32	0,25	0,07	0,02	0,01	0,00
Probabilidade estimada												
t = 0	-	-	-	-	-	-	-	-	-	-	-	-
t = 1	0,00	0,02	0,01	0,11	0,33	0,43	0,10	0,01	0,00	0,00	0,00	0,00
t = 2	0,00	0,01	0,01	0,06	0,15	0,38	0,29	0,08	0,01	0,00	0,00	0,00
t = 3	0,00	0,01	0,02	0,04	0,12	0,27	0,34	0,16	0,03	0,00	0,00	0,00
t = 4	0,00	0,01	0,03	0,03	0,09	0,18	0,31	0,23	0,09	0,03	0,00	0,00

A RNA 2 foi escolhida como a mais representativa da evolução da distribuição de diâmetros. A partir dessa rede, foram projetadas 10 parcelas escolhidas ao acaso, dentre aquelas separadas para validação do modelo. A situação inicial da parcela 1 condiz com uma situação típica dos povoamentos equiâneos com

a distribuição tendendo a normal (Figura 1). A dinâmica de evolução tende a tornar a distribuição assimétrica à direita, em razão das árvores de maiores dimensões apresentarem maiores taxas de crescimento radial em relação às árvores menores (GUIMARÃES, 1994).

Parcela

28 meses

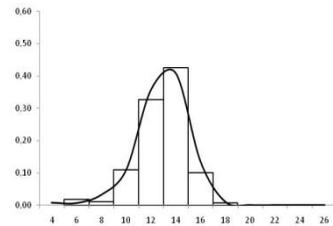
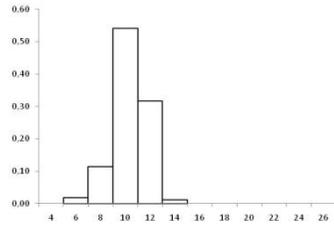
40 meses

52 meses

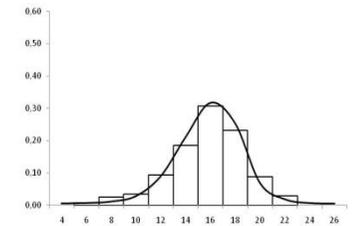
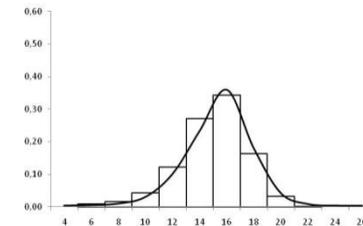
64 meses

76 meses

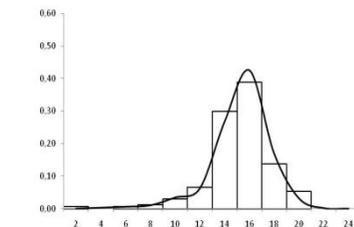
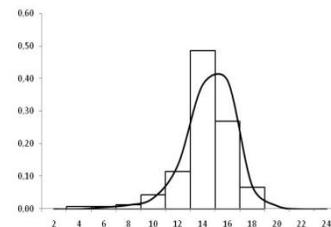
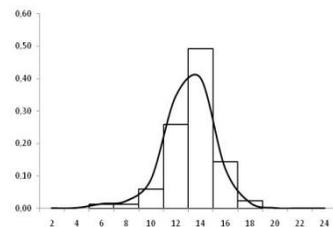
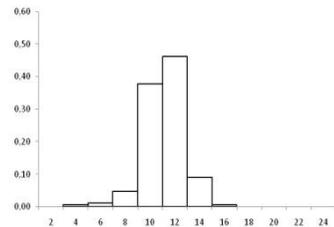
1



Sem Medição

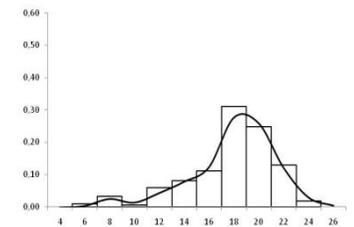
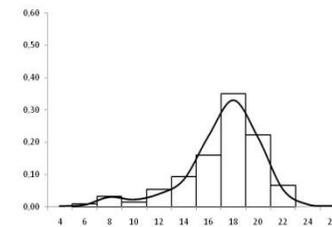
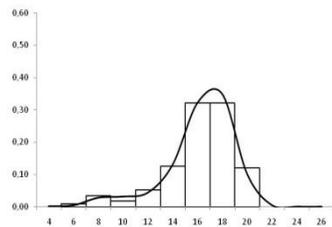
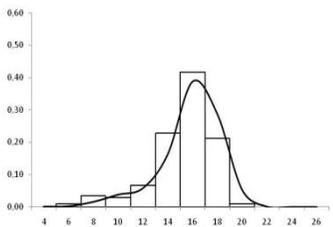
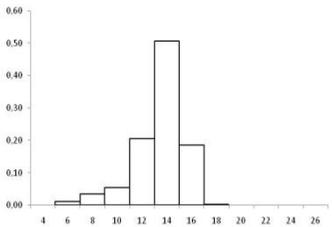


2

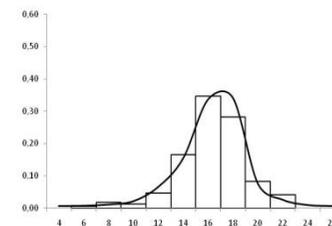
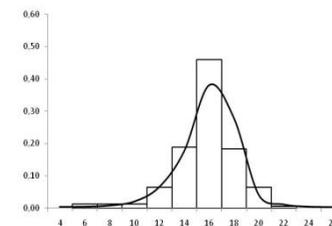
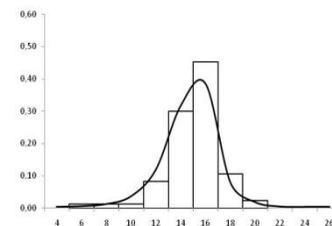
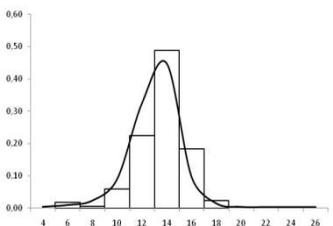
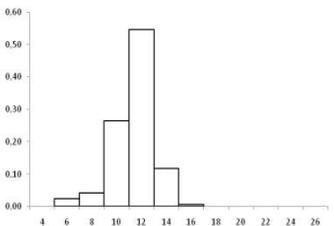


Sem Medição

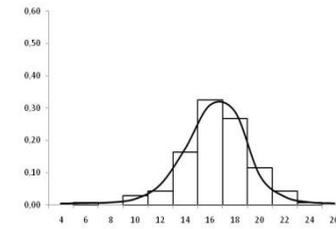
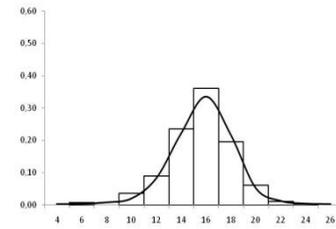
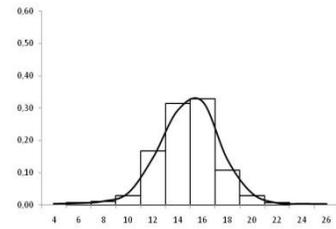
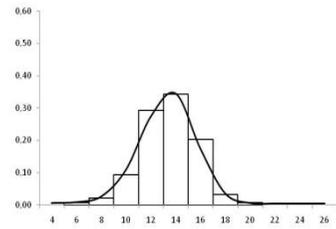
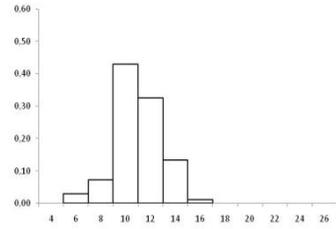
3



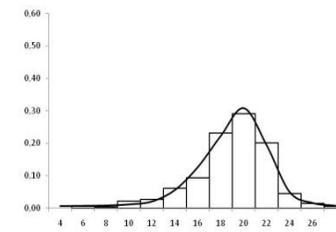
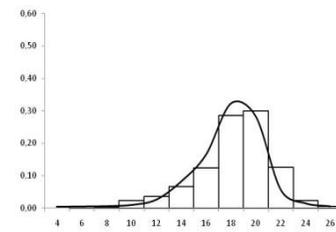
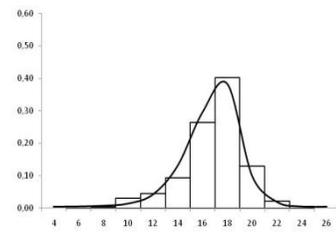
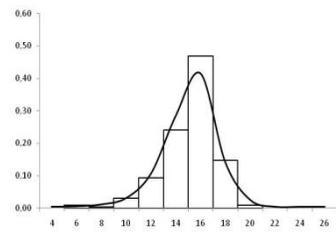
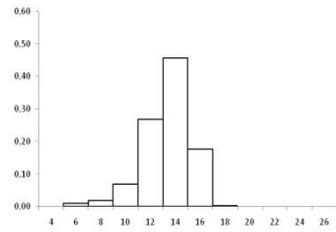
4



5



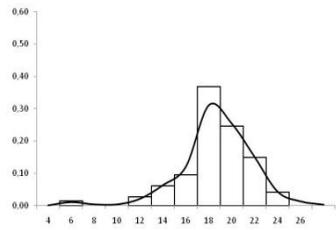
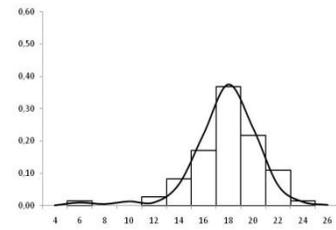
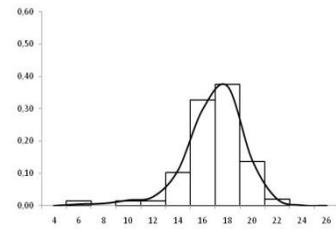
6



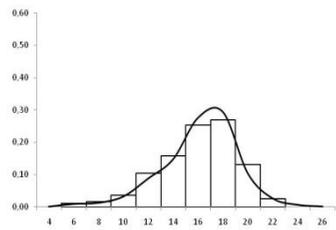
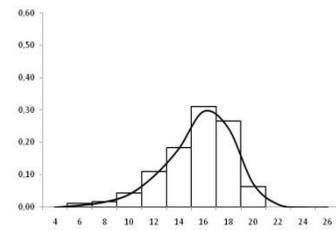
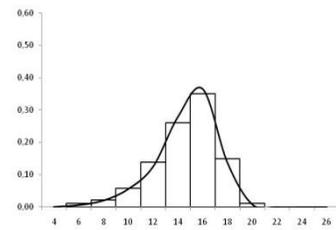
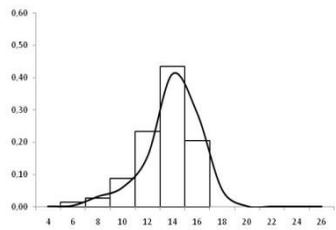
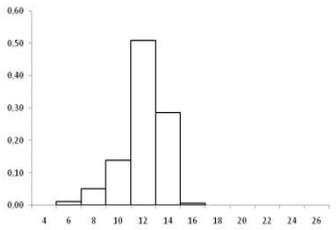
7

Sem Medição

Sem Medição



8



Continua...

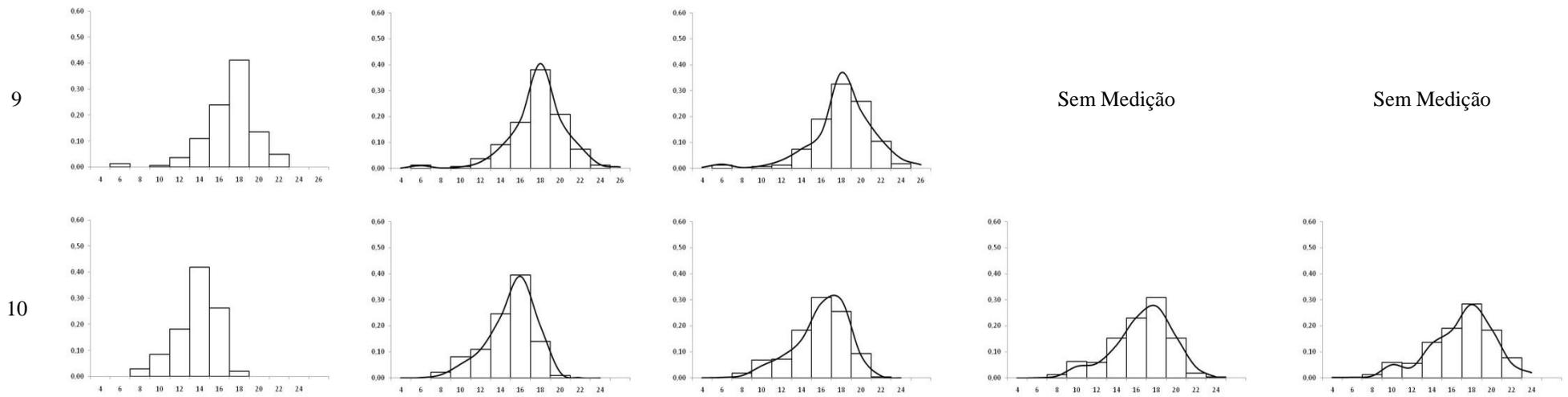


Figura 1 – Parcelas escolhidas demonstrando diversas dinâmicas de crescimento de povoamentos de eucalipto localizadas no estado de Minas Gerais, Brasil, e valores observados de probabilidade estimados pelo AC.

A distribuição inicial da parcela 2, com frequência concentrada em três classes de *dap*, indica que a competição deve ter iniciado em idade mais avançada. As parcelas 3, 7 e 9 apresentam um típico caso de evolução da distribuição diamétrica de povoamentos equiâneos onde, devido à competição, o crescimento radial nas árvores suprimidas estagnou. As parcelas 4, 5, 6, 8 e 10 demonstram que diferentes relações entre as classes de *dap* podem ser descritas pela regra de evolução. O modelo construído foi capaz de descrever diferentes tendências de evolução na distribuição de diâmetros.

A modelagem de distribuições de diâmetros em povoamentos de eucalipto tem sido feita com base em predição de parâmetros de uma distribuição estatística, empregando modelos de regressão. A função Weibull de dois parâmetros tem sido utilizada na maioria dos estudos por ser uma função flexível e pelo fato de seus parâmetros serem facilmente correlacionados com características dos povoamentos. Embora existam diferentes formas da função Weibull (MURTHY et al., 2004), na maioria das vezes esta função tem sido ajustada sem truncamentos à direita e à esquerda, pelo método da máxima verossimilhança. Os modelos de distribuição de diâmetros têm sido construídos como um conjunto de equações que estimam os parâmetros da distribuição estatística utilizada. Cada uma das equações desses sistemas tem um erro associado, sendo a estimativa dos parâmetros da função estatística dependente da exatidão das equações que compõem o sistema.

A eficiência das estimativas geradas pelos modelos de distribuição de diâmetros usuais é dependente da qualidade dos dados utilizados e da qualidade do ajuste das equações que compõem o modelo. Em alguns casos a frequência total estimada pelo modelo difere muito da frequência total observada ou estimada em nível de povoamento. Nesses casos, Clutter et al. (1983) recomendam que a diferença entre o somatório das frequências estimadas por classe de diâmetro e a frequência total prevista seja acrescida na última classe de diâmetro. Em povoamentos de eucalipto esta alternativa pode resultar em superestimação do volume por hectare. Outra alternativa, consiste em redistribuir a diferença proporcionalmente às frequências estimadas por classe, resultando conforme Campos e Leite (2009), em estimativas de volume de maior exatidão.

No presente estudo foi proposto um método alternativo para modelagem de distribuições diamétricas que não envolve o uso de modelos de regressão para estimar a proporção de árvores por classe de diâmetro. A proposta consiste em

associar um modelo de autômatos celulares e redes neurais artificiais. É uma abordagem simples que exige apenas distribuições de diâmetros observadas. Conforme resultado desse estudo, o enfoque de modelagem resulta em estimativas precisas e livres de *bias*. Especialmente no caso de povoamentos submetidos a desbaste, ou povoamentos com distribuições erráticas, como aquelas indicadas na idade de 76 meses para as parcelas 3 e 10 (Figura 1), a metodologia proposta é sugestiva. Outra proposta eficiente para modelagem de distribuições irregulares (erráticas) empregando a função Weibull foi apresentada por Zhang e Liu (2006). Embora não tenha sido apresentado por esses autores, a projeção de distribuições diamétricas a partir da função Weibull modificada também pode requerer o emprego modelos de regressão incluindo as variáveis independentes usuais (diâmetros médio e máximo, idade e índice de local).

As estatísticas de precisão obtidas neste estudo são semelhantes às aquelas obtidas em muitos modelos clássicos de distribuição de diâmetros e demonstram a eficiência da metodologia proposta. Embora não tenha sido apresentado, a modelagem apresentada neste estudo pode ser empregada para povoamentos submetidos a qualquer tipo de desbaste e também para povoamentos inequívocos. Conforme observado na Figura 1 (parcelas 3 e 10), a metodologia permite obter projeções descontínuas que podem ocorrer dependendo da mortalidade e do tipo de desbaste realizado. Isso normalmente não ocorre nos enfoques usuais, que geram distribuições contínuas. O modelo proposto descreve o fenômeno estudado com realismo biológico, podendo ser utilizado para a projeção da distribuição diamétrica de povoamentos equívocos.

4. CONCLUSÃO

Distribuições de diâmetro em povoamentos de eucalipto podem ser projetados com eficiência e exatidão empregando modelos de autômatos celulares e redes neurais artificiais.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ABBASI, B.; JAHROMIA, A. H. E.; ARKAT, J.; HOSSEINKOUCHACK, M. Estimating the parameters of Weibull distribution using simulated annealing algorithm. **Applied Mathematics and Computation**, v. 183, n.1, p.85-93, 2006.

ABBASI, B.; RABELO, L.; HOSSEINKOUCHACK, M. Estimating parameters of the three-parameter Weibull distribution using a neural network. **European Journal of Industrial Engineering**, v. 2, n.4, p.428-445, 2008.

BAILEY, R. L.; DELL, T. R. Quantifying diameter distributions with the Weibull function. **Forest Science**, v.19, n. 2, p.97-104, 1973.

BAILEY, R. L.; BURGAN, T. M.; JOKELA, E. J. Fertilized midrotation-aged slash pine plantations – Stand structure and yield prediction models. **Southern Journal of Applied Forestry**, v.13, n.2, p.76-80, 1989.

BAK, P.; CHEN, K.; TANG, C. A forest-fire model and some thoughts on turbulence. **Physics Letters A**, v.147, n.5-6, p. 297-300, 1990.

BALDWIN, V. C., Jr.; FEDUCCIA, D. P. **Loblolly pine growth and yield prediction of managed West Gulf plantations**. Res. Pap. SO-236. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 1987. 27 p.

BLISS, C. I.; REINKER, K. A. A lognormal approach to diameter distributions in even-aged stands. **Forest Science**, v.10, n.3, p.350-360, 1964.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. 2.ed. Rio de Janeiro: LTC, 2007. 226 p.

BULLINARIA, J. A. **Introduction to neural computation**. Lecture Notes. Disponível em <http://www.cs.bham.ac.uk/~jxb/inc.html>. 2010.

BURKHART, H. E. Note: Slash pine plantations yields estimates based on diameter distribution: an evaluation. **Forest Science**, v.17, n.4, p.452-453, 1971.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal**: perguntas e respostas. 3. ed. rev. e ampl. Viçosa: UFV, 2009. 548 p.

CAO, Q. V. Predicting parameters of a Weibull function for modeling diameter distribution. **Forest Science**, v. 50, n. 5, p.682-685, 2004.

CAO, Q. V.; BURKHART, H. E.; LEMIN, R.C., Jr. **Diameter distributions and yields of thinned loblolly pine plantations**. School of Forestry and Wildlife Resources. Publication No. FWS-1-82. Blacksburg: Virginia Polytechnic Institute and State University, 1982. 62 p.

BONE, C.; DRAGICEVIC, S.; ROBERTS, A. A fuzzy-constrained cellular automata model of forest insect infestations. **Ecological Modelling**, v.192, n.1-2, p.107-125, 2006.

CLUTTER, J. L.; ALLISON, B. J. A Growth and Yield Model for Pinus Radiata in New Zealand. In: Proc. **Growth Models for Tree and Stand Simulation** (J. FRIES, Ed.). Department of Forest Research, Research Notes 30, Royal College of Forestry, Stockholm, Sweden. 1974.

CLUTTER, J. L.; BENNETT, F. A. **Diameter distributions in old-field slash pine plantations**. Georgia Forest Research Council Report, 1965. 9 p. (Report, 13).

CLUTTER, J. L.; FORTSON, J. C.; PIENAAR, L. V.; BRISTER, R. G. H.; BAILEY, R. L. **Timber management**: a quantitative approach. New York: John Wiley e Sons, 1983. 333 p.

COHEN, C. A.; WHITTEN, B. Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution. **Communications in Statistics -Theory and Methods**, v.11, n.23, p. 2631-2656, 1982.

d'HUMIERES, D.; LALLEMAND, P. Lattice gas cellular automata, a new experimental tool for hydrodynamics. **Physica**, 140A, 1986.

DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. **Computers and Electronics in Agriculture**, v. 48, n.3, p.235-244, 2005.

EPSTEIN, J. M.; AXTELL, R. **Growing Artificial Societies**: Social Science from the Rottorn Up. Cambridge: MIT Press, 1996. 224p.

FRAZIER, J. R. **Compatible whole-stand and diameter distribution models for loblolly pine plantations**. 1981. 125 p. Ph.D. Dissertation, Virginia Polytech. Inst. and State Univ., Blackburg, VA.

FISHER, R. A. On the interpretation of χ^2 from contingency tables, and the calculation of P. **Journal of the Royal Statistical Society**, v.85, n.1, p.87-94, 1922.

GERHARDT, M.; SCHUSTER, H. A cellular automaton describing the formation of spatially ordered structures in chemical systems. **Physica D: Nonlinear Phenomena**, v.36, n.3, p.209-221, 1989.

GÖRGENS, E. B.; LEITE, H. G.; SANTOS, H. N.; GLERIANI, J. M. Estimação do volume de árvores utilizando redes neurais artificiais. **Revista Árvore**, v.33, n.6, p. 1141-1147, 2009.

GOVE, J. H.; FAIRWEATHER, S. E. Maximum-likelihood estimation of Weibull function parameters using a general interactive optimizer and grouped data. **Forest ecology and Management**, v.28, n.1, p. 61-69, 1989.

GOVE, J. H.; PATIL, G. P. Modelling the basal area-size distribution of forest stands: a compatible approach. **Forest Science**, v.44, n.2, p. 285-297, 1998.

GUIMARÃES, D. P. **Desenvolvimento de um modelo de distribuição diamétrica de passo invariante para prognose e projeção da estrutura de povoamentos de eucalipto**. 1994. 178 f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 1994.

HAFLEY, W. L.; SCHREUDER, H. T. Statistical distributions for fitting diameter and height data in even-aged stands. **Canadian Journal of Forest Research**, v.7, n.3, p.481-487, 1977.

HAYKIN, S. **Redes Neurais: princípios e prática**. 2 ed. Porto Alegre: Bookman, 2001. 900 p.

ILACHINSKI, A. **Cellular Automata: A Discrete Universe**. World Scientific Publishing Company, 2001. 808p.

KANGAS, A.; MALTAMO, M. Calibrating predicted diameter distribution with additional information. **Forest Science**, v.46, n.3, p.390-396, 2000.

KESSLER, D. A.; LEVINE, H.; REYNOLDS W. N. Coupled-map lattice model for crystal growth. **Physical Review A**, v. 42, n.10, p. 6125-6128, 1990.

KNOEBELL, B. R.; BURKHART, H. E.; BECK, D. E. A growth and yield model for thinned stands of yellow-poplar. *Forest Science*, v. 32, n. 2, p.62, 1986.

LEITE, H. G.; OLIVEIRA, F. H. T. Statistical procedure to test the identity of analytical methods. **Communications in soil science and plant analysis**. v. 33, p.1105-1118. 2002.

LINDEMAYER, A.; ROZENBERG, G. **Automata, Languages, Development**. North- Holland, Amsterdam, 1976, 529p.

GARZON, M. **Models of Massive Parallelism: Analysis of Cellular Automata and Neural Networks**. Springer-Verlag, Berlin, 1995. 272p.

MAGNUSSEN, S. Diameter distribution in *Picea abies* described by the weibull model. **Scandinavian Journal Forest Research**, n.1, p.493-502, 1986.

PALAHÍ, M.; PUKKALA, T.; BLASCO, E.; TRASOBARES, A. Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (north-east of Spain). **European Journal of Forest Research**, v.126, n.4, p.563-571. 2007.

- MATHEY, A. H.; KRCCMAR, E.; TAIT, D.; VERTINSKY, I.; INNES, J. Forest planning using co-evolutionary cellular automata. **Forest Ecology Management**. v.239, n.3, p.45-56. 2007.
- MATNEY, T. G.; SULLIVAN, A. D. Variable top volume and height predictions for slash pine trees. **Forest science**, v.28, n.2, p. 274-82, 1982.
- MEHTÄTALO, L.; MALTAMO, M.; KANGAS, A. The use of quantile trees in the prediction of the diameter distribution of a stand. **Silva Fennica**, v.40, n.3, p. 501-516. 2006.
- MURTHY, D. N. P.; XIE, M.; JIANG, R. Weibull Models. A John Wiley e Sons, Inc., Publication. 2004. 396 p.
- NELSON, T. C. Diameter distribution and growth of loblolly pine. **Forest Science**. v.10, n. 1, p.105-114. 1964.
- NEWTON, P. F.; LEI, Y.; ZHANG, S.Y. Stand-level diameter distribution yield model for black spruce plantations. **Forest Ecology and Management** v. 209 p.181-192, 2005.
- NOGUEIRA, G. S.; LEITE, H. G.; CAMPOS, J. C. C.; CARVALHO, A. F.; SOUZA, A. L. de. Modelo de distribuição diamétrica para povoamentos de Eucalyptus sp. submetidos a desbaste. **Revista Árvore**, v.29, n.4, p.579-589, 2005.
- NORD-LARSEN, T.; CAO, Q. V. A diameter distribution model for even-aged beech in Denmark. **Forest Ecology and Management** v.231, p.218-225, 2006.
- PALAHÍ, M.; PUKKALA, T.; TRASOBARES, A. Modelling the diameter distribution of Pinus sylvestris, Pinus nigra and Pinus halepensis forest stands in Catalonia using the truncated Weibull function. **Forestry**, v.79, n.5, p. 553-562. 2006.
- PÄIVINEN, R. On the estimation of stem diameter distribution and stand characteristics. **Folia Forestalia**. v.442. 28 p. 1980.
- PHIPPS, M. J. From local to global: the lesson of cellular automata, pages 165-187 In DEANGELIS, D. L.; GROSS, L. J. **Individual-Based Models and Approaches in Ecology: Populations, Communities and Ecosystems**, Chapman and Hall, 1992.
- SEGEL, L. A. Immunology viewed as the study of an autonomous decentralized system, pages 65-88 In DASGUPTA, D. **Artificial Immune Systems and their Applications**, Springer-Verlag, 1999.
- SIIPILEHTO, J.; SARKKOLA, S.; MEHTÄTALO, L. Comparing regression estimation techniques when predicting diameter distributions of Scots pine on drained peatlands. **Silva Fennica** v. 41, n.2, p.333-349. 2007.
- SILVA, M. L. M. **Estimação do diâmetro sem casca e diâmetro do cerne para árvores de *Tectona grandis* Linn. utilizando redes neurais artificiais**. 2008. 47 f. Monografia (Graduação Engenharia Florestal) – Universidade Federal de Viçosa, Viçosa, MG, 2008.

SOARES, T. S.; LEITE, H. G.; VALE, A. B.; SOARES, C. P. B. e SILVA, G.F. avaliação de um modelo de passo invariante na predição da estrutura de um povoamento de *Eucalyptus* sp. **Revista Árvore**, v.31, n.2, p.275-283, 2007.

BERGER, U.; HILDENBRANDT, H. A new approach to spatially explicit modelling of forest dynamics: spacing, ageing and neighborhood competition of mangrove trees. **Ecological Modelling**, v.132 p.287-302. 2000.

WOLFRAM, S. **Cellular automata**, Los Alamos: Science, 1983.

ZHANG, L.; LIU, C. Fitting irregular diameter distributions of forest stands by Weibull, modified Weibull, and mixture Weibull models. **Journal Forest Research**, v.11, p.369-372. 2006.

WEIBULL, W. A Statistical distribution function of wide Applicability. **Journal of Applied Mechanics**, v.18, p.293-297. 1951.

CAPÍTULO 6

MODELAGEM DA DISTRIBUIÇÃO DIAMÉTRICA DE POVOAMENTOS EUCALIPTO SUBMETIDOS A DESBASTE UTILIZANDO AUTÔMATOS CELULARES

1. INTRODUÇÃO

A estimação da distribuição diamétrica de povoamentos equiâneos é essencial para a avaliação econômica de multiprodutos da madeira. Essa estimação é comumente realizada empregando modelos de distribuição diamétrica. O procedimento utilizado nesses modelos se baseia no ajuste de uma função densidade de probabilidade (*f_{dp}*) aos dados de parcelas permanentes e no emprego de modelos de regressão relacionando os parâmetros dessa *f_{dp}* com características dos povoamentos (CLUTTER et al., 1983; CAMPOS e LEITE, 2009).

Diferentes funções (NELSON, 1964; BLISS e REINKER, 1965; CLUTTER e BENNETT, 1965; HAFLEY e SCHUREUDER, 1977; PALAHÍ et al., 2007; BAILEY e DELL, 1973; NOGUEIRA et al., 2005) e formas de obtenção dos seus parâmetros em idades futuras são utilizadas para a caracterização da distribuição diamétrica de povoamentos equiâneos (CAO, 2004). Assim como a função Weibull, a função hiperbólica destaca-se dentre as *f_{dp}* devido principalmente à sua flexibilidade, facilidade de ajuste e correlação significativa dos seus parâmetros com características do povoamento (GUIMARÃES, 2002; CAMPOS e LEITE, 2009; LEITE et al., 2010).

Técnicas de inteligência computacional, como redes neurais artificiais (DIAMONTOPOULOU, 2005; SILVA, 2008; GÖRGENS et al., 2009) e de vida

artificial, como autômatos celulares (MARTHEY et al., 2007), têm sido utilizadas em algumas aplicações na área florestal. Redes neurais artificiais (RNA) são processadores paralelamente distribuídos compostos por unidades de processamento simples (neurônios artificiais), que armazenam conhecimento experimental (aprendizagem) tornando-o disponível para uso (generalização) (BRAGA et al., 2007; HAYKIN, 2001). Autômatos celulares (AC) são sistemas de interações locais, discretos no tempo e no espaço, compostos de células, onde o estado futuro de cada célula é resultado de seu estado atual, da sua regra de transição e da sua interação com células vizinhas (WOLFRAM, 1983; ILACHINSKI, 2001). As células são componentes simples que interagem localmente e a complexidade em âmbito global advém de regras entre essas unidades mais simples. Diversos métodos podem ser utilizados como regras de transição, com destaque para RNA (GARZON, 1995).

A modelagem empregando AC tem sido feita para a descrição de sistemas complexos como, por exemplo: fenômenos químicos (d'HUMIERES e LALLEMAND, 1986; GERHARDT e SCHUSTER, 1989), crescimento vegetal (LINDEMAYER e ROZENBERG, 1976), crescimento de cristais (KESSLER, 1990), ecologia (PHIPS, 1992), propagação de doenças infecciosas (SEGEL, 1999), dinâmica social (AXTELL e EPSTEIN, 1996) e incêndios florestais (BAK et al. 1990).

O objetivo deste estudo foi construir, testar e avaliar um modelo de autômatos celulares, com sua regra de evolução baseada em redes neurais artificiais, para a projeção da distribuição de diâmetros em povoamentos de eucalipto e com simulação de desbastes.

2. MATERIAIS E MÉTODOS

2.1 Dados

Os dados utilizados neste estudo foram obtidos de 400 parcelas permanentes retangulares com área de 340 m², instaladas em povoamentos de clones de eucalipto sob arranjo espacial de 3 x 3 m, com ciclo de corte médio de 7 anos, e produtividade média, na idade de corte, de 40 m³ ha⁻¹ ano⁻¹ nessa idade. As medições de diâmetros foram efetuadas nas árvores com *dap* acima de 5 cm, nas idades médias de 28, 40, 52, 64, 76 e 88 meses. O arquivo de dados foi separado em dois subconjuntos, sendo um para construção do modelo (50% dos dados) e o outro para sua validação. Nessa separação foi mantida uma frequência homogênea de dados por classe de capacidade produtiva.

2.2 Modelo de Autômatos Celulares

Para um modelo unidimensional de AC, o valor da *i*^a célula no tempo *t*, denominado por $c_i(t)$, evolui de acordo com a regra *F* que é função do $c_i(t)$ e das outras células que se encontram dentro da extensão *r* (para ambas as direções) de $c_i(t)$ (ILACHINSKI, 2001), isto é:

$$C_i(t) = F(C_{i-r}(t-1), \dots, C_{i+r}(T+1))$$

O AC proposto neste estudo é unidimensional sendo que cada célula possui 4 vizinhos ($r = 2$). As células representam uma classe de diâmetro de 1 cm de amplitude, sendo que o estado de cada célula corresponde à probabilidade de ocorrência de árvores nesta classe.

Para a construção da regra de evolução do AC foram utilizadas redes neurais artificiais, com as seguintes estruturas: (1) O arquivo de dados foi formatado com

simulações aleatórias de desbastes. As redes foram treinadas tendo como variáveis de entrada: estado atual da célula, estado das células contidas na extensão da vizinhança, intensidade de desbaste e as idades atual e futura; (2) variáveis de entrada: estado atual da célula, estado das células contidas na extensão da vizinhança e as idades atual e futura. Os treinamentos das redes e a projeção, nesse caso, foram realizados até a idade do desbaste e após o desbaste, com a mesma rede; (3) variáveis de entrada: estado atual da célula, estado das células contidas na extensão da vizinhança e as idades atual e futura. Os treinamentos das redes e a projeção, nesse caso, foram realizados até a idade do desbaste (3a) e a após o desbaste (3b) com outra rede. Para a simulação de desbastes ajustou-se um modelo $dap = f(I)$, para cada árvore de cada parcela, sendo posteriormente, simuladas aleatoriamente cinco intensidades de desbastes variando entre 20 e 50 % da área basal em idades variando de 30 a 150 meses, sendo o corte final projetado para os 180 meses. A estrutura de formatação dos dados para treinamento das redes é apresentada na Tabela 1. Em todos os métodos propostos projetou-se o estado futuro da classe de *dap* (célula).

Tabela 1 – Estrutura utilizada no treinamento de RNA para a simulação de desbastes com o modelo de autômatos celulares, demonstrando a probabilidade de ocorrência de árvores em cada classe de diâmetro.

Estrutura	Idade (meses)	Intensidade de desbaste (%)	Classe de dap																											
			5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
1	20	-	0	0,01	0,07	0,25	0,46	0,18	0,03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	30	20%	0	0	0	0,01	0	0,06	0,14	0,33	0,27	0,14	0,05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	40	-	0	0	0	0	0	0	0	0	0	0,23	0,35	0,24	0,12	0,05	0	0	0	0	0	0	0	0	0	0	0	0	0	
	50	-	0	0	0	0	0	0	0	0	0	0	0,09	0,24	0,34	0,18	0,09	0,05	0	0	0	0	0	0	0	0	0	0	0	
	60	-	0	0	0	0	0	0	0	0	0	0	0	0,08	0,23	0,24	0,24	0,09	0,09	0,01	0	0	0	0	0	0	0	0	0	
	70	-	0	0	0	0	0	0	0	0	0	0	0	0,01	0,12	0,23	0,19	0,26	0,08	0,08	0,03	0	0	0	0	0	0	0	0	
	80	-	0	0	0	0	0	0	0	0	0	0	0	0	0,04	0,16	0,2	0,2	0,22	0,07	0,08	0,03	0	0	0	0	0	0	0	
	90	30%	0	0	0	0	0	0	0	0	0	0	0	0	0,03	0,08	0,2	0,19	0,23	0,11	0,08	0,08	0	0	0	0	0	0	0	
	100	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,15	0,3	0,3	0,08	0,09	0,08	0	0	0	0	0	0	
	120	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,02	0,23	0,26	0,25	0,09	0,08	0,08	0	0	0	0	0	
	140	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,13	0,23	0,26	0,15	0,09	0,08	0,06	0	0	0	0	
	160	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,09	0,17	0,25	0,23	0,09	0,04	0,11	0,02	0	0	0	
180	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,17	0,3	0,17	0,13	0,09	0,09	0,04	0	0	0		
2	20	-	0	0,01	0,07	0,25	0,46	0,18	0,03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	40	-	0	0	0	0	0	0	0	0	0	0,23	0,35	0,24	0,12	0,05	0	0	0	0	0	0	0	0	0	0	0	0	0	
	50	-	0	0	0	0	0	0	0	0	0	0	0,09	0,24	0,34	0,18	0,09	0,05	0	0	0	0	0	0	0	0	0	0	0	
	60	-	0	0	0	0	0	0	0	0	0	0	0	0,08	0,23	0,24	0,24	0,09	0,09	0,01	0	0	0	0	0	0	0	0	0	
	70	-	0	0	0	0	0	0	0	0	0	0	0	0,01	0,12	0,23	0,19	0,26	0,08	0,08	0,03	0	0	0	0	0	0	0	0	
	80	-	0	0	0	0	0	0	0	0	0	0	0	0	0,04	0,16	0,2	0,2	0,22	0,07	0,08	0,03	0	0	0	0	0	0	0	
	100	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,15	0,3	0,3	0,08	0,09	0,08	0	0	0	0	0	0	
	120	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,02	0,23	0,26	0,25	0,09	0,08	0,08	0	0	0	0	0	
	140	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,13	0,23	0,26	0,15	0,09	0,08	0,06	0	0	0	0	
	160	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,09	0,17	0,25	0,23	0,09	0,04	0,11	0,02	0	0	0	
	180	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,17	0,3	0,17	0,13	0,09	0,09	0,04	0,04	0	0	
	3	20	Sem Desbaste	0	0,01	0,07	0,25	0,46	0,18	0,03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30		Sem Desbaste	0	0	0	0,01	0	0,06	0,14	0,33	0,27	0,14	0,05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
40		Sem Desbaste	0	0	0	0	0,01	0	0,02	0,06	0,08	0,26	0,26	0,18	0,09	0,04	0	0	0	0	0	0	0	0	0	0	0	0	0	
50		Sem Desbaste	0	0	0	0	0	0,01	0,01	0,03	0,02	0,04	0,16	0,22	0,27	0,13	0,07	0,04	0	0	0	0	0	0	0	0	0	0	0	
60		Sem Desbaste	0	0	0	0	0	0,01	0,01	0	0,05	0,01	0,02	0,15	0,22	0,19	0,19	0,07	0,07	0,01	0	0	0	0	0	0	0	0	0	
70		Sem Desbaste	0	0	0	0	0	0,01	0,01	0	0,02	0,03	0,01	0,07	0,15	0,2	0,15	0,2	0,07	0,06	0,02	0	0	0	0	0	0	0	0	
80		Sem Desbaste	0	0	0	0	0	0,01	0	0,01	0	0,05	0	0,02	0,1	0,16	0,19	0,15	0,17	0,06	0,06	0,02	0	0	0	0	0	0	0	
90		Sem Desbaste	0	0	0	0	0	0	0,01	0,01	0	0,05	0	0,01	0,07	0,12	0,19	0,15	0,17	0,09	0,07	0,06	0	0	0	0	0	0	0	
100		Sem Desbaste	0	0	0	0	0	0	0,01	0,01	0	0,02	0,03	0	0,03	0,09	0,16	0,17	0,17	0,17	0,14	0,06	0,04	0	0	0	0	0	0	
120		Sem Desbaste	0	0	0	0	0	0	0,01	0,01	0	0,02	0,03	0	0,01	0,07	0,12	0,16	0,14	0,15	0,13	0,06	0,05	0,04	0	0	0	0	0	
140		Sem Desbaste	0	0	0	0	0	0	0,01	0,01	0	0	0,04	0,01	0	0,03	0,08	0,16	0,17	0,13	0,14	0,09	0,05	0,05	0,03	0	0	0	0	
160		Sem Desbaste	0	0	0	0	0	0	0,01	0,01	0	0	0,04	0,01	0	0,02	0,08	0,14	0,16	0,11	0,14	0,12	0,06	0,02	0,07	0,01	0	0	0	
180		Sem Desbaste	0	0	0	0	0	0	0,01	0,01	0	0	0,04	0,01	0	0,02	0,05	0,09	0,15	0,15	0,17	0,09	0,07	0,06	0,06	0,02	0	0	0	
40		Após Desbaste	0	0	0	0	0	0	0	0	0	0,23	0,35	0,24	0,12	0,05	0	0	0	0	0	0	0	0	0	0	0	0	0	
50		Após Desbaste	0	0	0	0	0	0	0	0	0	0	0,09	0,24	0,34	0,18	0,09	0,05	0	0	0	0	0	0	0	0	0	0	0	
60		Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0,08	0,23	0,24	0,24	0,09	0,09	0,01	0	0	0	0	0	0	0	0	0	
70		Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0,01	0,12	0,23	0,19	0,26	0,08	0,08	0,03	0	0	0	0	0	0	0	0	
80		Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0,04	0,16	0,2	0,2	0,22	0,07	0,08	0,03	0	0	0	0	0	0	0	0	
90		Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0	0,03	0,08	0,2	0,19	0,23	0,11	0,08	0,08	0	0	0	0	0	0	0	
100		Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,15	0,3	0,3	0,08	0,09	0,08	0	0	0	0	0	0	
120	Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,02	0,23	0,26	0,25	0,09	0,08	0,08	0	0	0	0	0		
140	Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,13	0,23	0,26	0,15	0,09	0,08	0,06	0	0	0	0		
160	Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,09	0,17	0,25	0,23	0,09	0,04	0,11	0,02	0	0	0		
180	Após Desbaste	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,17	0,3	0,17	0,13	0,09	0,09	0,04	0,04	0	0		

Foram testadas redes de função de base radial (RBF - *Radial Basis Function*), perceptrons de múltiplas camadas (MLP - *Multilayer perceptron*) e perceptrons de camada única (*Perceptron*), com diferentes arquiteturas, sendo que para cada metodologia, foram treinadas 10 redes de cada tipo, sendo selecionadas, aquelas que melhor representaram os dados. O *software* Statistica 10 (STATSOFT, Inc, 2012) foi utilizado nessa etapa do estudo.

2.3 Modelo de distribuição diamétrica

Para avaliação das projeções realizadas pelo modelo de AC empregado nesse estudo, realizou-se a projeção da distribuição diamétrica por um modelo de distribuição diamétrica comumente utilizado. Utilizou-se como função densidade probabilidade a função hiperbólica (hyper) descrita por Guimarães (2002):

$$f(X) = \left\{ \omega \frac{\gamma}{\beta} \left(\frac{X - \alpha}{\beta} \right)^{(\gamma-1)} \left(1 - \tanh \left(\left(\frac{X - \alpha}{\beta} \right)^\gamma \right) \right)^2 \right\}$$

em que, ω é o valor assintótico, α o parâmetro de locação, β parâmetro de escala ($\beta > 0$), γ é o parâmetro de forma ($\gamma > 0$), e X é o centro de classe de diâmetro ($X > 0$). Sendo que a supressão do parâmetro de locação caracteriza a função hiperbólica de dois parâmetros.

Para a obtenção dos parâmetros da *fdp* em uma idade futura utilizou-se o sistema de equações proposto por Nogueira et al. (2005).

$$\beta_2 = \beta_1 \left(\frac{I_1}{I_2} \right) + \alpha_1 \left(1 - \frac{I_1}{I_2} \right)$$

$$Ln\gamma_2 = Ln\gamma_1 e^{(-\alpha_0)(I_2^{\alpha_1} - I_1^{\alpha_1})}$$

$$N_2 = N_1 e^{(-\alpha_0)(I_2^{\alpha_1} - I_1^{\alpha_1})}$$

em que: I_1 e I_2 são as idades atual e futura, respectivamente, em meses; γ_1 e γ_2 são os parâmetros de forma da função hiperbólica nas idades atual e futura; β_1 e β_2 são os parâmetros de escala da função hiperbólica nas idades atual e futura; N_1 e N_2 são as frequências de árvores por hectare nas idades atual e futura; Ln indica o logaritmo neperiano; e a_0 e a_1 são os parâmetros dos modelos.

O modelo de AC e o modelo de distribuição diamétrica foram empregados para projetar a distribuição de diâmetros para todas as idades. Para a avaliação da

exatidão das projeções foi utilizado o procedimento estatístico proposto por Leite e Oliveira (2002) a 5% de significância. Avaliou-se graficamente as projeções entre valores estimados e projetados pelo MDD e pelos modelos de AC.

3. RESULTADOS

Dentre todas as redes treinadas as redes do tipo *MLP* apresentaram melhores resultados, quando comparadas com as redes do tipo *RBF* e *Perceptron*, para todas as estruturas de dados, sendo que a arquitetura das duas melhores, em função do erro de treinamento, é apresentada na Tabela 2.

Tabela 2 – Principais características de duas redes neurais artificiais treinadas para avaliar a evolução da distribuição diamétrica de povoamentos de eucalipto em duas formas de entradas de dados.

Estrutura	Rede	Número de neurônios			Função de Ativação	
		Camada de entrada	Camada oculta	Camada de saída	Camada oculta	Camada de saída
1	Rede-1	8	11	1	Exponencial	Seno
	Rede-2	8	11	1	Exponencial	Logística
2	Rede-1	7	10	1	Exponencial	Logística
	Rede-2	7	12	1	Exponencial	Seno
3A	Rede-1	7	9	1	Tangente hiperbólica	Exponencial
	Rede-2	7	10	1	Logística	Tangente hiperbólica
3B	Rede-1	7	11	1	Exponencial	Logística
	Rede-2	7	12	1	Exponencial	Identidade

Para todas as simulações e idades ajustou-se a função hiperbólica de dois parâmetros, sendo que, todos os ajustes apresentaram aderência aos dados pelo teste de aderência Kolmogorov-Smirnov a 5% de significância. O teste t foi significativo para todos os parâmetros em todas as equações do sistema. Todas as equações apresentaram valores relativamente altos de coeficientes de correlação entre valores observados e estimados, indicando que as variáveis independentes utilizadas nas equações têm grande influência nas variáveis dependentes. A análise gráfica dos resíduos permitiu inferir a inexistência de tendenciosidade nos ajustes (Figura 1).

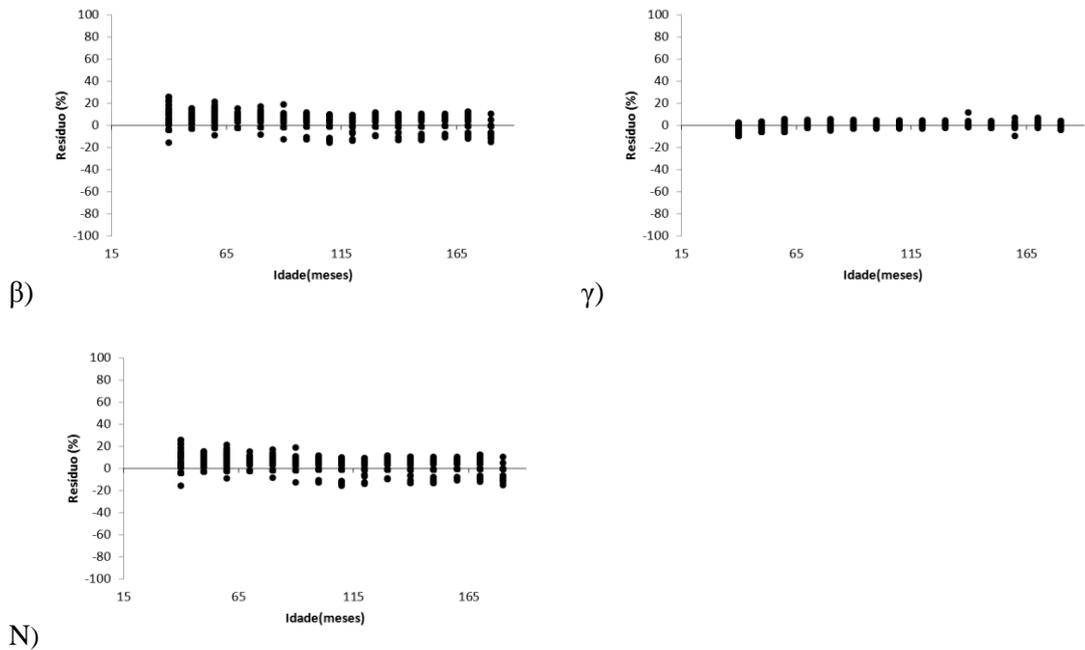


Figura 1 – Dispersão dos resíduos (%) em função da idade para as três equações do modelo de distribuição diamétrica.

Os modelos ajustados e seus respectivos coeficientes de correlação foram:

$$\beta_2 = \beta_1 \left(\frac{I_1}{I_2} \right) + 23,5565 \left(1 - \frac{I_1}{I_2} \right) \quad r_{yy} = 0,964$$

$$\text{Ln}\gamma_2 = \text{Ln}\gamma_1 e^{-(0,1687)(I_2^{0,2564} - I_1^{0,2564})} \quad r_{yy} = 0,954$$

$$N_2 = N_1 e^{-(14,2704)(I_2^{-0,0104} - I_1^{-0,0104})} \quad r_{yy} = 0,995$$

Com as duas melhores redes obtidas para cada estrutura e com o modelo de distribuição diamétrica, foram feitas as projeções das distribuições diamétricas nos dados destinados à validação do modelo. Resultados da aplicação do procedimento proposto por Leite e Oliveira (2002) são apresentados na Tabela 3. Esses resultados demonstram que as regras de evolução representaram a dinâmica do fenômeno estudado com eficiência e exatidão.

A superioridade do método proposto em relação ao modelo de distribuição de diâmetros é verificada na Tabela 3 e na Figura 2. Conforme essa tabela, enquanto o erro médio com o MDD foi acima de 1% o maior erro médio obtido com o AC foi de 0,2095.

Tabela 3 – Resumo dos resultados obtidos com o procedimento estatístico proposto por Leite e Oliveira (2002), para as melhores RNA e para o modelo de distribuição diamétrica ($\alpha= 5\%$), com o resultado da avaliação da hipótese H_0 : A distribuição de diamétrica estimada é igual à projeção estimada.

Estatística	MDD	Estrutura 1		Estrutura 2		Estrutura 3A		Estrutura 3B	
		Rede-1	Rede-2	Rede-1	Rede-2	Rede-1	Rede-2	Rede-1	Rede-2
F(Ho)	787,7473	96,9555	270,4260	0,3675	117,4866	1,3798 ^{ns}	1,9361 ^{ns}	0,1527 ^{ns}	0,8395 ^{ns}
t(e)	36,5778	4,0442	3,7029	1,1618 ^{ns}	9,0503	0,5752 ^{ns}	0,6030 ^{ns}	1,0000 ^{ns}	0,6255 ^{ns}
$r_{y, \hat{y}_i} \geq 1 - \hat{r} $	sim	não	não	sim	sim	sim	sim	sim	sim
r ²	0,3163	0,7113	0,4467	0,9288	0,9167	0,8129	0,8491	0,9426	0,9394
RMS	0,0060	0,0023	0,0034	0,0005	0,0008	0,0028	0,0021	0,0005	0,0004
Erro médio (e)	1,1111	0,1133	0,2095	-0,0918	-0,0461	0,2131	0,2498	-0,0625	0,0357
Resultado	#	#	#	=	#	=	=	=	=

A avaliação gráfica entre valores observados e estimados foi realizada para as melhores redes de cada forma de ajuste. Na Figura 2 é apresentado as análises gráficas de uma parcela escolhida aleatoriamente, com um desbaste de 30% da área basal aos 80 meses e outro de 50% da área basal aos 120 meses.

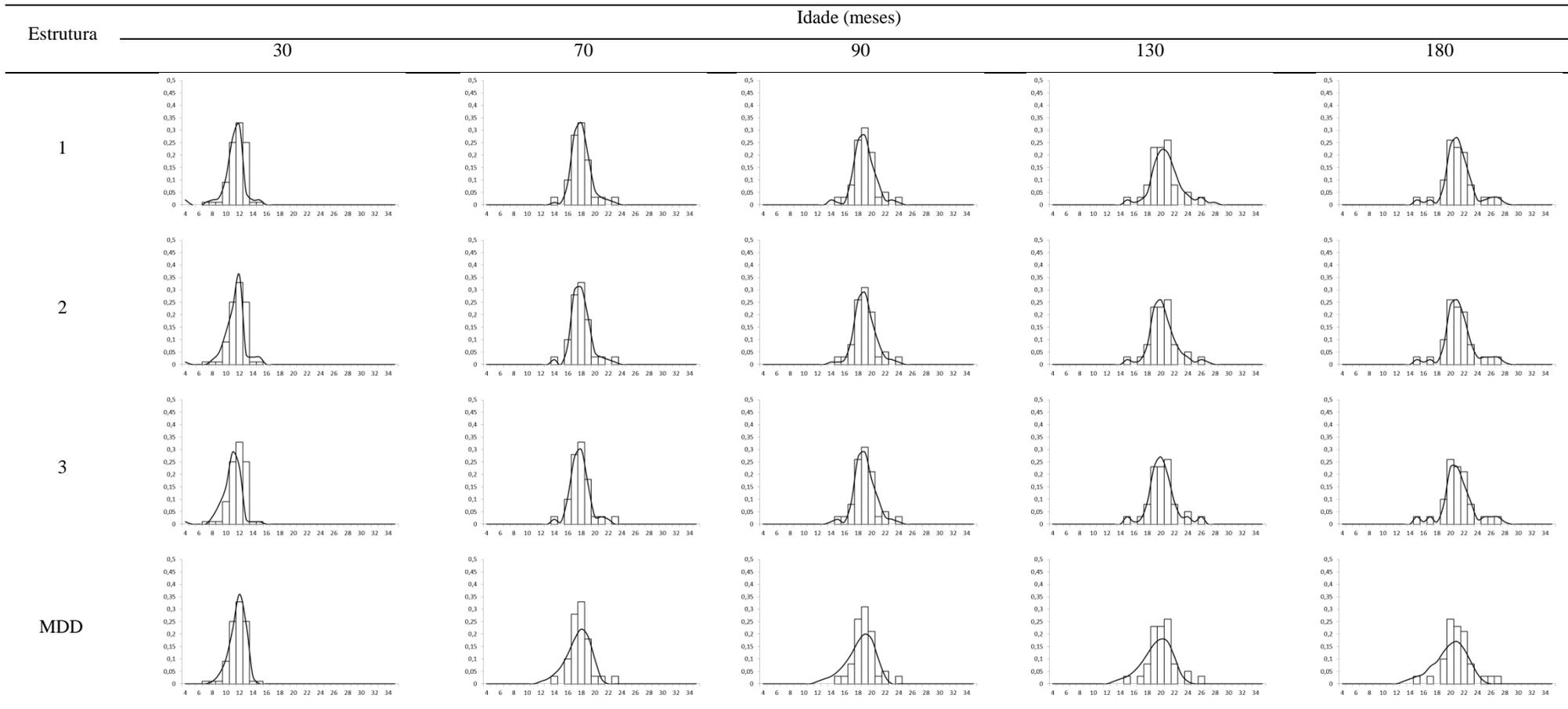


Figura 2 – Análises gráficas de uma parcela escolhida aleatoriamente, com um desbaste aos 80 meses de 30% área basal e outro desbastes na idade 120, de 50% da área basal, para as 3 estruturas propostas e modelo de distribuição de diâmetros (MDD).

4. DISCUSSÃO E CONCLUSÕES

A modelagem de distribuições de diâmetros em povoamentos de eucalipto tem sido feita com base em predição e projeção de parâmetros de uma distribuição estatística, empregando modelos de regressão (NOGUEIRA, et al., 2005). Diferentes tipos de distribuições estatísticas já foram empregadas para descrever a estrutura diamétrica de povoamentos florestais, como: Gama (NELSON, 1964), log-normal (BLISS e REINKER, 1965), Beta (CLUTTER e BENNETT, 1965; PALAHÍ et al., 2007), Johnson's SB (HAFLEY e SCHUREUDER, 1977; PALAHÍ et al., 2007) e a distribuição Weibull (WEIBULL, 1951; BAILEY e DELL, 1973; PALAHÍ et al., 2007). Desde 1973, a função Weibull tem sido amplamente difundida e utilizada na área florestal (CLUTTER e ALLISON, 1974; HAFLEY e SCHREUDER 1977; MATNEY e SULLIVAN, 1982; CAO, 2004; NOGUEIRA et al. 2005; PALAHÍ et al., 2006; PALAHÍ et al., 2007). Os trabalhos de Guimarães (2002) e Leite, et al. (2010) demonstraram a superioridade da função hiperbólica, para a descrição da distribuição diamétrica de povoamentos equiâneos.

A eficiência das estimativas geradas pelos modelos de distribuição de diâmetros usuais é dependente da qualidade dos dados utilizados e da qualidade do ajuste das equações que compõem o modelo. Cada uma das equações desses sistemas tem um erro associado, sendo a estimativa dos parâmetros da função estatística dependente da exatidão das equações que compõem o sistema.

No presente estudo foi proposto um método alternativo para modelagem de distribuições diamétricas e simulações de desbaste, que não envolvam o uso de modelos de regressão para estimar a proporção de árvores por classe de diâmetro. A proposta consiste em associar um modelo de autômatos celulares e redes neurais artificiais. É uma abordagem simples que exige apenas distribuições de diâmetros

observadas e se baseia no fato da probabilidade de cada classe de diâmetro em uma idade futura ser dependente do número atual de árvores nessa classe e nas classes vizinhas. Por esse fato, as árvores suprimidas apresentam uma melhor representação como pode ser observado na Figura 2.

Os três métodos propostos diferem na forma de ajuste e simulação de desbastes. O primeiro apresenta como principal característica a projeção e a simulação de desbaste completa com apenas uma estrutura, ou seja, uma única rede é responsável pela projeção e desbaste, sendo que este deve ser definido como variável de entrada para o modelo. A presença do desbaste como variável de entrada é determinada pela presença de oito neurônios na camada de entrada (Tabela 2). Com o segundo método a simulação dos desbastes deve ser feita pelo usuário, toda a projeção deve ser realizada por uma única rede que deve captar as diferentes tendências de crescimento expressas pelo povoamento antes e após a realização do desbaste. No terceiro método proposto, as diferentes tendências de crescimento, existentes antes e após o desbaste são expressas por diferentes redes, esse fato apresenta uma considerável melhoria na descrição da dinâmica de crescimento, o que pode ser observado nas estatísticas do teste L&O (Tabela 3). Portanto, recomenda-se a utilização da terceira metodologia para a projeção da distribuição diamétrica de povoamentos de eucalipto submetidos a desbaste.

Conforme resultado desse estudo, o enfoque de modelagem resulta em estimativas livres de *bias*, podendo ser utilizado para a projeção da distribuição diamétrica de povoamentos submetidos a desbaste.

5. REFERÊNCIAS BIBLIOGRÁFICAS

AXTELL, R.; EPSTEIN, J. M. **Growing Artificial Societies: Social Science from the Rottorn Up**, MIT Press, 1996.

BAILEY, R. L.; DELL, T. R. Quantifying diameter distributions with the Weibull function. **Forest Science**, v.19, n.2, p 97-104,1973.

BAK, P.; CHEN, K.; TANG, C. A forest-fire model and some thoughts on turbulence. **Physics Letters A**, v.147, n. 5 e 6, p. 297-300, 1990.

BLISS, C. L.; REINKER, K. A. A log normal approach to diameter distributions in even-aged stands. **Forest Science**, v. 10, n.3, p.350-360, 1964.

BRAGA, A. P.; LUDERMIR, T. B.; CARVALHO, A. P. de L. F. **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro, 2000. 262 p.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3 ed. Viçosa: UFV, 2009. 548 p.

CAO, Q. V. Predicting parameters of a Weibull function for modeling diameter distribution. **Forest Science**, v. 50, n. 4, p.682-685, 2004.

CLUTTER, J. L.; ALLISON, B. J. A growth and yield model for Pinus radiata in New Zealand for tree and stand simulation. **Royal College of Forestry**. Research Notes, n. 30, p. 136-160, 1974.

CLUTTER, J. L.; BENNETT, F. A. Diameter distributions in old-field slash pine plantations. **Georgia Forest Research Council Report**, n.13, p. 1-9, 1965.

CLUTTER, J. L.; FORTSON, J.C.; PIENAAR, L.V.; BRISTER, R.G.H.; BAILEY, R.L. **Timber management: a quantitative approach**. New York: John Willey e Sons, 1983. 333 p.

d'HUMIERES, D.; LALLEMAND, P. Lattice gas cellular automata, a new experimental tool for hydrodynamics, **Physical Review Letters**, v.56, n.14, p. 1505-1508, 1986.

DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. **Computers and electronics in agriculture**, v.48, n.3, p. 235-244, 2005.

GERHARDT, M.; SCHUSTER, H. A cellular automaton describing the formation of spatially ordered structures in chemical systems. **Physica D: Nonlinear Phenomena**, v.36, n.3, p. 209-221, 1989.

GÖRGENS, E. B.; LEITE, H. G.; SANTOS, H. N.; GLERIANI, J. M. Estimação do volume de árvores utilizando redes neurais artificiais. **Revista Árvore**, v.33, n.6, p. 1141-1147, 2009.

HAFLEY, W. L.; SCHREUDER, H. T., Statistical distributions for fitting diameter and height data in ever-aged stands. **Canadian Journal of Forest Research**, v.7, p.184-487, 1977.

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

ILACHINSKI, A. **Cellular Automata: A Discrete Universe**, 808p. World Scientific Publishing Company, 2001.

KESSLER, D. A.; LEVINE, H.; REYNOLDS W. N. Coupled-map lattice model for crystal growth. **Physical Review A**, v. 42, n.10, p. 6125-6128 ,1990.

LEITE, H. G.; OLIVEIRA, F. H. T. Statistical procedure to test identity between analytical methods. **Communications in soil science and plant analysis**, v.33, n. 7 e 8, p.1105-1118, 2002.

LEITE, H. G.; BINOTI, D. H. B.; GUIMARÃES, D. P.; SILVA, M. L. M.; GARCIA, S. L. R. Avaliação do ajuste das funções Weibull e hiperbólica a dados de povoamentos de eucalipto submetidos a desbaste. **Revista Árvore**, v.34, n.2, p. 305-311, 2010.

LINDEMAYER, A.; ROZENBERG, G. Review of "Automata, languages, development". **North-Holland Publishing Company**, v. 9, n.1, p. 16-18, 1976.

MATHEY, A. H.; KRUMHOLTZ, E.; TAIT, D.; VERTINSKY, I.; INNES, J. Forest planning using co-evolutionary cellular automata. **Forest Ecology and Management**, v. 239, n. 1-3, p. 45-56, 2007.

MATNEY, T. G.; SULLIVAN, A. D. Variable top volume and height predictions for slash pine trees. **Forest science**, v.28, n.2 p.74-82, 1982.

NELSON, T. C. Diameter distribution and growth of loblolly pine. **Forest Science**, v.10 n.1, p.105-114,1964.

NOGUEIRA, G. S.; LEITE, H. G.; CAMPOS, J. C. C.; CARVALHO, A. F.; SOUZA, A. L. de. Modelo de distribuição diamétrica para povoamentos de *Eucalyptus* sp. submetidos a desbaste. **Revista Árvore**, v.29, n.4, p.579-589, 2005.

PALAHÍ, M.; PUKKALA, T.; TRASOBARES, A. Modelling the diameter distribution of *Pinus sylvestris*, *Pinus nigra* and *Pinus halepensis* forest stands in Catalonia using the truncated Weibull function. **Forestry**, v.79, n.5, p.553-562, 2006.

PHIPPS, M. J. From local to global: the lesson of cellular automata. In: DEANGELIS, D. L.; GROSS, L. J. (Ed.) **Individual-Based Models and Approaches in Ecology: Concepts and Models**. Chapman and Hall, 1992.

SEGEL, L. A. Immunology viewed as the study of an autonomous decentralized system. In: DASGUPTA, D. **Artificial Immune Systems and their Applications**. Springer-Verlag, 1999.

SILVA, M. L. M. **Estimação do diâmetro sem casca e diâmetro do cerne para árvores de *Tectona grandis* Linn. utilizando redes neurais artificiais**. 2008. 47 f. Monografia (Graduação Engenharia Florestal) – Universidade Federal de Viçosa, Viçosa, MG, 2008.

WEIBULL, W. A Statistical distribution function of wide applicability. **Journal of Applied Mechanics**, v.18, p.293-297, 1951.

WOLFRAM, S. **Cellular automata**. Los Alamos: Science, 1983.1083p.

CAPÍTULO 7

PARAMETRIZAÇÃO DE REDES NEURAIS ARTIFICIAIS PARA ESTIMAÇÃO DO VOLUME DE ÁRVORES

1. INTRODUÇÃO

As variáveis que interferem no volume das árvores, como idade, espaçamento, índice de local, solo, apresentam na maioria das vezes relações complexas e tendências muitas vezes não lineares, dificultando o ajuste de modelos volumétricos com inclusão dessas variáveis. As redes neurais artificiais (RNA) apresentam-se como uma ferramenta apropriada para ser utilizada nessas situações, onde a inclusão de variáveis categóricas não é factível (BINOTI et al., 2009).

Uma rede neural artificial (ou simplesmente uma rede neural ou rede) é um modelo computacional biologicamente inspirado, constituído por elementos de processamento simples (neurônios artificiais) que aplicam uma determinada função matemática aos dados (função de ativação) gerando uma única resposta, são dispostos em camadas e ligados entre si, sendo estas conexões, geralmente, associadas a coeficientes denominados de pesos. O ajuste destes pesos é realizado por um processo chamado treinamento ou aprendizado, sendo responsável pela extração das características dos dados e armazenamento de conhecimento das redes. Enquanto que a aplicação de uma rede consiste no processo de generalização, ou seja, utilização de uma rede treinada em dar resposta para dados inéditos (BRAGA et al., 2007).

Atualmente muitas formas de RNA são propostas e usadas com variações principalmente, no número de neurônios por camada, tipo de função de ativação dos

neurônios, número de camadas ocultas e tipo de conexão entre os neurônios. Outro fator que interfere significativamente no tempo de treinamento e na eficiência de uma rede, são os algoritmos utilizados, bem como os parâmetros e coeficientes utilizados em seu treinamento.

O número de neurônios na camada oculta é definido principalmente de forma empírica pela experiência do usuário, embora existam algumas metodologias propostas na literatura, como em Hirose et al. (1991), Arai (1993) e Fujita (1998). Além disso, no processo de determinação do número de neurônios deve-se considerar que um número excessivo de neurônios pode acarretar a memorização dos dados de treinamento, processo conhecido como *overfitting*. Por outro lado, um pequeno número de neurônios na camada escondida pode não ser suficiente para a realização da tarefa desejada, fenômeno conhecido como *underfitting* (SILVA et al., 2010).

A definição do algoritmo de treinamento a ser utilizado, bem como de seus parâmetros influenciam principalmente na fuga de mínimos locais, no desempenho da tarefa desejada e no tempo de treinamento. Por exemplo, taxas de aprendizado muito baixas acarretam um aprendizado muito lento, enquanto que taxas de aprendizado muito altas geram oscilações no processo de treinamento, dificultando e até mesmo impedindo a convergência do algoritmo de treinamento *error backpropagation*, pois a busca no espaço de soluções “pula” o erro mínimo global (HAYKIN, 2001; BRAGA et al., 2007; BULLINARIA, 2010; SILVA et al., 2010).

Objetivou-se nesse trabalho avaliar diferentes configurações do treinamento de RNA para a estimação do volume de árvores. Como método comparativo ajustou-se o modelo de Schumacher e Hall (1933).

2. MATERIAL E MÉTODOS

2.1 Dados

Os dados utilizados neste estudo foram provenientes de cubagens de 2.307 árvores de povoamentos de eucalipto (*Eucalyptus grandis* x *Eucalyptus urophylla*) com produtividade média, aos 7 anos, de cerca de 40 m³ha⁻¹ano⁻¹.

As variáveis numéricas (quantitativas) consideradas foram: idade em meses, diâmetro com casca a 1,30 m de altura, em cm (*dap*), altura total, em m (*Ht*) e volume da árvore, em m³ (*V*) (Tabela 1). Enquanto que as variáveis categóricas (qualitativas) foram: projeto, espaçamento, e clone. O número de classes das variáveis categóricas foram 23, 7 e 15, respectivamente. As árvores-amostra foram cubadas em seções de 1 m, sendo o volume total com casca (*V*) obtido pela aplicação da fórmula de Smalian (CAMPOS e LEITE, 2009).

Tabela 1 – Análise descritiva das variáveis numéricas.

Variável	Mínimo	Máximo	Média	Desvio padrão
Idade (anos)	2,71	11,00	6,50	± 2,35
<i>dap</i> (cm)	10,82	32,47	18,27	± 4,37
Altura total (m)	17,94	43,80	28,22	± 5,10
Volume (m ³)	0,100	1,580	0,379	± 0,23

2.2 Modelo Volumétrico

Ao estratificar os dados foram definidos 53 estratos, com base nas variáveis: projeto, espaçamento, idade e clone. Para cada estrato ajustou-se o modelo de Schumacher e Hall (CAMPOS e LEITE, 2009):

$$V = \beta_0 dap^{\beta_1} Ht^{\beta_2} \varepsilon$$

em que, V = volume, m^3 ; dap = diâmetro a 1,3 m de altura, cm; Ht = altura total da árvore, m; β_i = parâmetros; e ε = erro aleatório, $\varepsilon \sim NID(0, \sigma^2)$.

2.3 Redes Neurais Artificiais

As variáveis numéricas foram normalizadas linearmente no intervalo de 0 a 1 e as variáveis categóricas foram normalizadas utilizando-se a metodologia 1-de-N (GOLDSCHMIDT e PASSOS, 2005; HEATON, 2010).

A camada de entrada foi constituída de 48 sendo um neurônio para cada variável numérica e classe de variável categórica, ou seja: dap (1), altura (1), idade (1), projeto (23), espaçamento (7), clone (15). Como saída utilizou-se o volume total da árvore. As redes foram constituídas de apenas uma camada oculta, sendo que o número de neurônios nessa camada variou de 1 a 48 (que corresponde ao número de neurônios na camada de entrada). Como funções de ativação da camada oculta e de saída utilizou-se as funções: linear, log, tangente hiperbólica, seno e sigmoide (HEATON, 2010).

Os algoritmos de treinamentos utilizados foram *error backpropagation*, *resilient propagation*, *Manhattan update rule*, *scaled conjugate gradient*, *levenberg marquardt*, *quick propagation*, *simulated annealing*, e algoritmos genéticos.

Para o algoritmo *backpropagation* utilizou-se variações no valor da taxa de aprendizado de 0,001 a 1,000 (a cada 0,001), bem como variações no parâmetro *momentum* de 0,0000 a 1,000 (a cada 0,0005), totalizando 1000 valores de taxas de aprendizado e 2001 valores de *momentum*. Os valores de taxa de aprendizado e taxas de *momentum* foram combinados totalizando 2.001.000 combinações. Além dessas combinações utilizou-se a estratégia de auto ajuste dos valores de taxa de aprendizado e *momentum*, processo que reduz os valores das taxas de aprendizado e *momentum* com a evolução do algoritmo resultando em 8.004.000 configurações do algoritmo *backpropagation*.

Para o algoritmo *resilient propagation*, considerou-se quatro variações básicas denominadas RPROP+, RPROP-, iRPROP+, iPROP (HEATON, 2010). No algoritmo *Manhattan update rule*, que por definição necessita de pequenas taxas de treinamento, restringiu-se a utilização de somente um valor de taxa de aprendizado: 0,00001. Os algoritmos *scaled conjugate gradient*, *levenberg marquardt* apresentam somente uma configuração não dependendo de valores de parâmetros de treinamento.

O algoritmo de treinamento conhecido como *quick propagation* é baseado no método de Newton utilizado para obtenção de raízes de uma função. Esse algoritmo é tolerante a altas taxas de aprendizado, por isso avaliou-se taxas variando de 0,1 a 10,0 (a cada 0,1) totalizando 100 configurações desse algoritmo.

A metaheurística *Simulated annealing* é baseado no processo de recozimento de metais utilizado em processos metalúrgicos, o processo consiste em aquecer um metal e verificar como esse metal se resfria e congela numa estrutura cristalina de energia mínima. Como principais parâmetros desse algoritmo tem-se o número de ciclos por interação e os valores de temperatura máxima e mínima. Os valores de números de ciclos avaliados foram variados de 100 a 1000 (a cada 100), com temperatura mínima de dois e máxima de dez, totalizando 10 configurações.

Outra metaheurística utilizada foi o Algoritmo Genético (AG). O AG baseia-se no processo de evolução natural das espécies e nos princípios de reprodução genética. O sistema possui como principais parâmetros configuráveis o tamanho da população, a probabilidade de mutação e a taxa de cruzamento. Testou-se valores de população variando de 50 a 1000 (a cada 50) totalizando 20 valores de população, valores de mutação variando de 0,1 a 0,5 (a cada 0,1) totalizando 5 valores de mutação, valores de cruzamento variando 0,05 a 0,70 (a cada 0,05) totalizando 14 valores. Foram testados no total 114 configurações de AG.

Foram avaliadas todas as possíveis combinações entre números de neurônios na camada oculta, tipo de funções de ativação na camada oculta e de saída, número de ciclos de treinamento e configurações dos algoritmos. Os pesos iniciais de todas as redes foram gerados aleatoriamente. Para cada combinação avaliada gerou-se 100 redes, sendo considerada uma simulação. A escolha da melhor configuração e o processo de treinamento das redes neurais artificiais foi feito avaliando o erro quadrático médio e o desvio padrão dos erros quadráticos médios de cada simulação, sendo escolhidas as configurações mais simples e consistentes. Como critério de parada dos algoritmos de treinamento, utilizou-se: o número total de ciclos (com os

valores variando de 500 a 3000 (a cada 100), totalizando 26 valores) ou o erro quadrático médio inferior a 1%. Portanto, finalizou-se o treinamento quando um dos critérios foi atingido.

O treinamento e a avaliação das redes foram feitos no sistema NeuroForest – Volumetric, desenvolvido por Helio G. Leite, Daniel H. B. Binoti e Mayra Luiza M. da S. Binoti, disponível gratuitamente no site <http://neuroforest.ucoz.com/>, a versão disponível no site foi alterada pelos próprios autores a fim de rodar scripts executando as configurações testadas nesse trabalho.

2.4 Avaliação dos Modelos

As estimativas das RNA e dos modelos de Schumacher e Hall foram avaliadas com base nos coeficientes de correlação com os valores observados e a raiz quadrada do erro médio (RMSE), bem como na análise gráfica dos resíduos (erros percentuais) e histograma de erros percentuais.

$$RMSE(\%) = 100 \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}} / \bar{Y}$$

em que, Y é a saída observada nos inventários, \hat{Y} é a saída estimada, \bar{Y} é a média da saída observada e n é o número total de dados.

3. RESULTADOS E DISCUSSÃO

Foram estimados 53 equações volumétricas a partir do modelo de Schumacher e Hall. O coeficiente de correlação entre os valores observados e estimados pelas equações foi de 0,993 e o valor de RMSE estimado foi de 7,28%. As estimativas apresentaram-se sem tendenciosidade e livre de bias. A dispersão dos resíduos em função dos valores observados e o histograma de resíduos são apresentados na Figura 1.

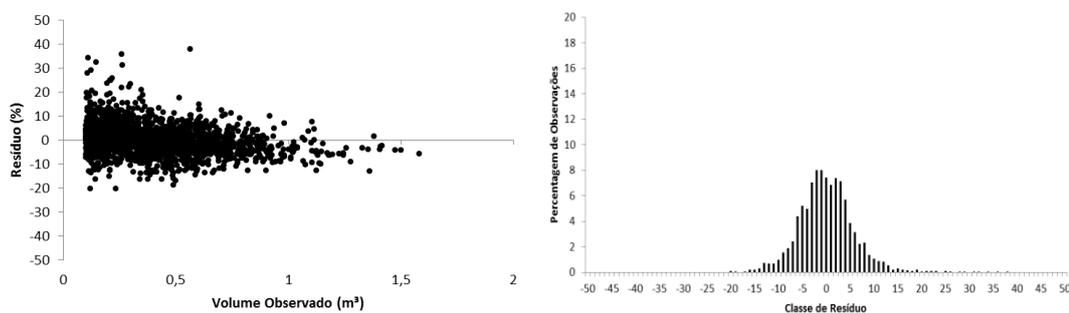


Figura 1 – Gráfico de resíduos e histograma de erros obtidos pelas estimativas realizadas pelo modelo de Schumacher e Hall.

Foram avaliadas 1200 configurações diferentes de RNA referentes ao número de neurônios na camada oculta e combinações de funções de ativação. Como critério de parada utilizou-se o número de ciclos totalizando 26 critérios. Essas configurações foram combinadas com as simulações de cada algoritmo *backpropagation* (8.004.000), *resilient propagation* (4), *Manhattan update rule* (1), *scaled conjugate gradiente* (1), *levenberg marquardt* (1), *quick propagation* (100), *simulated annealing* (10), e algoritmos genéticos (114). No total foram testadas $2,49 \cdot 10^{11}$ configurações sendo treinadas $2,49 \cdot 10^{13}$ RNA.

Em função da enorme quantidade de redes treinadas, foram avaliadas apenas as configurações que apresentaram valores satisfatórios de erro quadrático médio, ou seja, valores menores que 1%.

Obteve-se resultados satisfatórios com configurações de apenas um neurônio na camada oculta, demonstrando que redes de estruturas complexas não são necessárias para a estimação do volume de árvores. Provavelmente, isto se justifica pelo fato do número de neurônios na camada oculta ser responsável principalmente pela extração de características de não linearidade dos dados, neste caso, um único neurônio foi capaz de extrair a relação entre as variáveis consideradas e o volume.

Recomenda-se que nas aplicações de RNA, sejam escolhidas configurações mais simples, com o menor número de neurônios possível na camada oculta, a fim de evitar o *overfitting* e facilitar o processo de busca e otimização da configuração para uma determinada tarefa (BULLINARIA, 2010).

Não houve diferença entre as funções de ativação na camada oculta avaliadas. Para as funções de ativação na camada de saída a função linear, log seno e tangente hiperbólica apresentaram valores negativos para algumas estimativas volumétricas, devido à normalização das variáveis numéricas ter sido feita entre valores de 0 a 1, enquanto que estas funções resultam em valores entre -1 e 1.

Não houve redes satisfatórias com os algoritmos *scaled conjugate gradient*, *Manhattan update rule*, *levenberg marquardt*, *simulated annealing* e algoritmos genéticos, para os dados desse estudo em nenhuma configuração testada.

Para taxas de aprendizado com valores acima de 0,1, o algoritmo *backpropagation* mostrou-se instável, não convergindo. Valores baixos de taxa de aprendizagem (menores que 0,01) e do *momentum* (menores que 0,01) apresentaram resultados satisfatórios, não influenciando de forma perceptível, o tempo de treinamento do algoritmo, apesar de exigir um maior número de ciclos para a convergência.

Não foi objetivo avaliar o tempo de processamento, pelo fato do sistema NeuroForest - Volumetric ser extremamente eficiente no processo de treinamento (cada simulação gastou aproximadamente 10 segundos), o que torna o tempo fator irrelevante nesse caso.

Para o algoritmo *resilient propagation* somente as configurações com mais do que três neurônios na camada oculta apresentaram resultados satisfatórios. Não

houve diferença entre as quatro variações básicas RPROP+, RPROP-, iRPROP+, iPROP.

O algoritmo *quick propagation* apresentou resultados satisfatórios somente para configurações que possuíam mais do que três neurônios na camada oculta e taxa de aprendizado inferior a 1,0. Taxas superiores a 1,0 causaram instabilidade no algoritmo não apresentando convergência.

Selecionou-se duas configurações que apresentaram resultados satisfatórios para demonstração gráfica (Figura 2): A) um neurônio na camada oculta, com função de ativação tangente hiperbólica na camada oculta e sigmoide na camada de saída, treinamento utilizando o algoritmo *backpropagation* com taxa de aprendizado igual a 0,001 e *momentum* nulo; e B) 12 neurônios na camada oculta, com função de ativação tangente hiperbólica na camada oculta e sigmoide na camada de saída, treinamento utilizando o algoritmo *resilient propagation* (RPROP+).

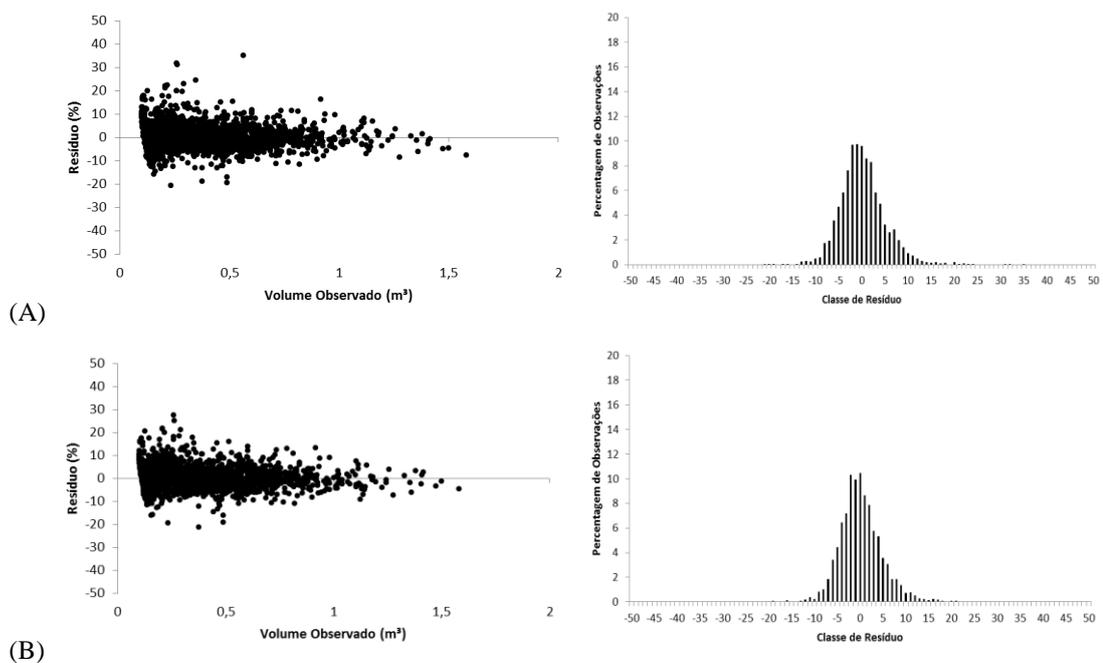


Figura 2 – Gráfico de resíduos e histograma de erros obtidos pelas estimativas realizadas pelas RNA treinadas. A) um neurônio na camada oculta, com função de ativação tangente hiperbólica na camada oculta e sigmoide na camada de saída, treinamento utilizando o algoritmo *backpropagation* com taxa de aprendizado igual a 0,001 e *momentum* nulo; e B) 12 neurônios na camada oculta, com função de ativação tangente hiperbólica na camada oculta e sigmoide na camada de saída, treinamento utilizando o algoritmo *resilient propagation* (RPROP+).

Observou-se pouca diferença nas estimativas com relação ao modelo de Schumacher e Hall. Deve-se salientar que as estimativas foram feitas utilizando 53 equações volumétricas (estratificação) contra somente uma única RNA para todos os dados. Fato que favorece a utilização das RNA bem como a aplicação operacional do método, por ser possível reduzir muito o tempo gasto com ajustes e avaliações de modelos volumétricos por estrato. Além disso, não houve diferenças perceptíveis entre a configuração A (pouca complexidade) e B (maior complexidade), fato que justifica a utilização de RNA com baixa complexidade, ou seja, poucos neurônios na camada oculta, o que pode levar a rede a memorizar os dados de treinamento (*overfitting*), ao invés de extrair as características gerais dos dados que permitirão a generalização.

4. CONCLUSÕES

A estimação do volume de árvores por meio de redes neurais artificiais pode ser feita com diferentes configurações, variando o número de neurônios na camada oculta, as funções de ativação, os algoritmos e os parâmetros de treinamento.

Redes neurais simples, com apenas um neurônio na camada oculta podem ser utilizadas com eficiência para estimação do volume de árvores.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ARAI, M. Bounds on the number of hidden units in binary-valued three-layer neural networks. **Neural Networks**, v.6, p. 855-860, 1993.

BINOTI, M. L. M. S.; 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. **Revista Árvore**, v. 33, n. 6, p. 1133-1139, 2009.

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

BULLINARIA, J. A. Introduction to Neural Computation. **Notas de aula**. 2008. Disponível em: <<http://www.cs.bham.ac.uk/~jxb/inc.html>>.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3 ed. Viçosa: UFV, 2009. 548 p.

FUJITA, O. Statistical estimation of the number of hidden units for feedforward neural networks. **Neural Networks**, v.11, p. 851-859, 1998.

GOLDSCHMIDT, R.; PASSOS, E. **Data mining: um guia prático**. Rio de Janeiro: Elsevier, 2005. 261 p.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001.

HEATON, J. **Programming Neural Networks with Encog3 in Java**. 2 nd ed. Heaton Research, Incorporated, 2011. 240 p.

HIROSE, Y.; YAMASHITA, K.; HIJIYA, S. Back-propagation algorithm which varies the number of hidden units. **Neural Networks**, v.4, p.61-66, 1991.

SCHUMACHER, F. X.; HALL, F. S. Logarithmic expression of timber-tree volume. **Journal of Agricultural Research**, v.47, n.9, p.719-734, 1933.

SILVA, I. N.; SPATTI, H. D.; FLAUZINO, R. A. **Redes Neurais Artificiais: para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010.

CAPÍTULO 8

ESTIMAÇÃO DO VOLUME DE ÁRVORES UTILIZANDO O SISTEMA NEUROFOREST - VOLUMETRIC

1. INTRODUÇÃO

Desde a introdução de técnicas de regressão no manejo florestal, estas têm sido utilizadas para a estimação do volume e altura de árvores, predição ou projeção do volume de povoamentos em idades futuras, dentre outras funções, gerando na maioria dos casos estimativas consistentes.

A estimação do volume de árvores é comumente feita pela correlação do volume total com variáveis facilmente mensuráveis como altura e diâmetro a 1,3 m de altura (*dap*). O modelo de Schumacher e Hall (1933) é predominante na área florestal devido a suas propriedades estatísticas, facilidade de ajuste e uso, além da geração de estimativas de boa qualidade (CAMPOS e LEITE, 2009).

A utilização de redes neurais artificiais (RNA) tem se mostrado uma alternativa promissora em relação às técnicas de regressão para no manejo dos recursos florestais. Diversos trabalhos têm sido desenvolvidos visando sua adaptação e parametrização para diversas situações como a estimação do volume de árvores (SILVA et al., 2009), crescimento e produção (BINOTI, 2010) e *taper* (LEITE et al., 2010), dentre outras aplicações.

Entretanto a utilização de técnicas de RNA depende de ferramentas computacionais especializadas, o que pode acarretar um custo expressivo para a sua utilização e entendimento. Ferramentas computacionais especializadas para aplicação

de RNA na área florestal ainda não haviam sido divulgadas e este fato, em muitos casos, dificulta ou impossibilitava a utilização de RNA na área florestal.

Recentemente foi desenvolvido o sistema NeuroForest – Volumetric, exatamente para viabilizar o uso eficiente das técnicas de RNA para resolver problemas de quantificação do volume de árvores em povoamentos florestais. O sistema é específico, de interface amigável ao usuário e de distribuição gratuita.

1.1 Objetivo

O objetivo deste estudo foi de demonstrar um caso de uso do NeuroForest – Volumetric para quantificação do volume de árvores e avaliar sua eficiência na aplicação de redes neurais artificiais. Também foi objetivo apresentar uma fundamentação teórica sucinta, porem suficiente sobre redes neurais artificiais.

2. MATERIAL E MÉTODOS

2.1 Desenvolvimento do Projeto

O NeuroForest - Volumetric foi desenvolvido utilizando a linguagem de programação Java. Como ambiente de desenvolvimento foram utilizados o IDE (Integrated Development Environment) Netbeans 7.1 e a JDK 7.3 (Java Development Kit). Os testes do sistema foram realizados em ambiente Windows. O sistema utiliza a biblioteca Encog 3 para a geração, treinamento e construção das RNA (HEATON, 2011).

2.2 Estudo de Caso

Inicialmente foi elaborado um texto com a fundamentação teórica sobre RNA, suficiente para o entendimento e uso do NeuroForest – Volumetric. Em seguida foi conduzido um caso de uso do software conforme descrito em

2.2 Redes Neurais Artificiais

Uma rede neural artificial (RNA) é uma rede de neurônios artificiais, e, portanto, constitui uma aproximação grosseira de partes do cérebro real (redes neurais biológicas). Consiste em um sistema computacional paralelo composto de elementos de processamento simples (neurônios artificiais) que calculam determinadas funções matemáticas e estão conectados entre si de uma maneira específica, a fim de executar uma determinada tarefa. Na maioria dos modelos essas conexões estão associadas a pesos que armazenam o conhecimento adquirido durante o processo de aprendizado ou treinamento (fase de ajuste dos pesos) e ponderaram as

entradas recebidas por cada neurônio da rede (HAYKIN, 2001; BRAGA et al., 2007; BULLINARIA, 2010).

As unidades de processamento de uma RNA, denominadas neurônios artificiais, são modelos bem simplificados dos neurônios biológicos encontrados nos sistemas nervosos biológicos e no cérebro humano. O neurônio biológico é dividido em três partes principais: dendritos, corpo celular e axônio (Figura 1). Os dendritos recebem informações na forma de impulsos nervosos oriundos de outros neurônios e os conduz até o corpo celular que processa os sinais recebidos e caso atinjam o limiar de excitação do neurônio, um impulso elétrico é produzido e propagado através do axônio para os neurônios seguintes. A região de contato entre a terminação axônica de um neurônio e os dendritos de outro consiste na sinapse que pondera a transmissão de impulsos elétricos, ressalta-se que o contato físico inexistente, mas sim substâncias neurotransmissoras unem funcionalmente os neurônios, formando as redes neurais biológicas (Figura 2). Este processo simples de comunicação entre neurônios é responsável pela maioria das funções realizadas pelo cérebro, porém a capacidade de realizar funções complexas surge da atividade em paralelo de vários neurônios do sistema nervoso, ou seja, das redes neurais biológicas (BRAGA et al., 2007; SILVA et al., 2010).

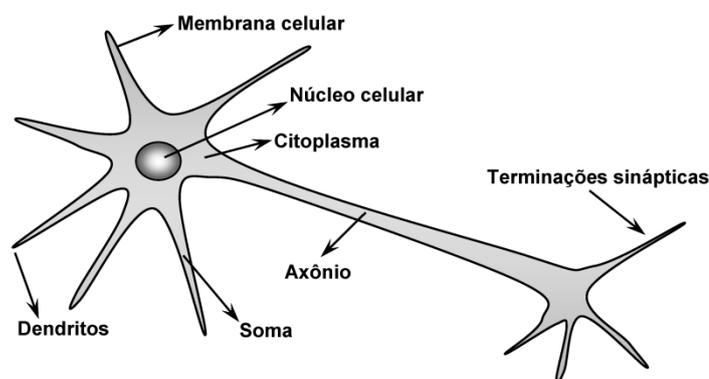


Figura 1 – Neurônio biológico (SILVA et al., 2010).

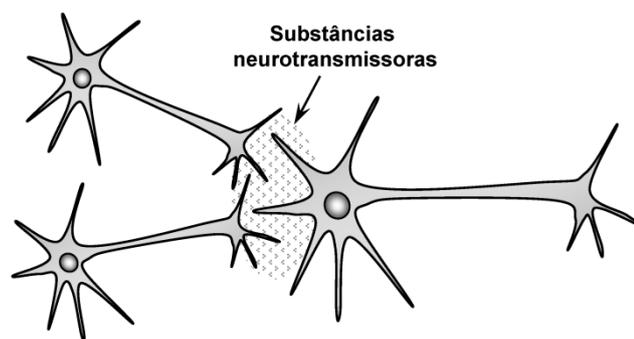


Figura 2 – Conexões sinápticas entre neurônios biológicos (SILVA et al., 2010).

O primeiro modelo de neurônio artificial foi proposto por McCulloch e Pitts em 1943. De maneira similar ao neurônio biológico, o neurônio artificial (Figura 3) possui n terminais de entrada (dendritos) que recebem cada um, um valor x_1, x_2, \dots, x_n (ativações dos neurônios anteriores). Os terminais de entrada possuem pesos associados w_1, w_2, \dots, w_n cujos valores podem ser positivos ou negativos e ponderam os valores das entradas, além de armazenar o conhecimento adquirido (sinapse). Um combinador linear (Σ) soma o produto de cada entrada do neurônio por seu respectivo peso, posteriormente é subtraído desta soma o limiar de excitação do neurônio (θ), resultando no potencial de ativação do neurônio (u), ao qual é aplicada uma função matemática (função de ativação, $g(u)$) cujo objetivo é limitar a saída do neurônio (y). Assim, o neurônio artificial pode ser descrito pela seguinte função matemática (BRAGA et al., 2007; SILVA et al., 2010):

$$y = g(u), \text{ sendo } u = \sum_{i=1}^n x_i w_i - \theta$$

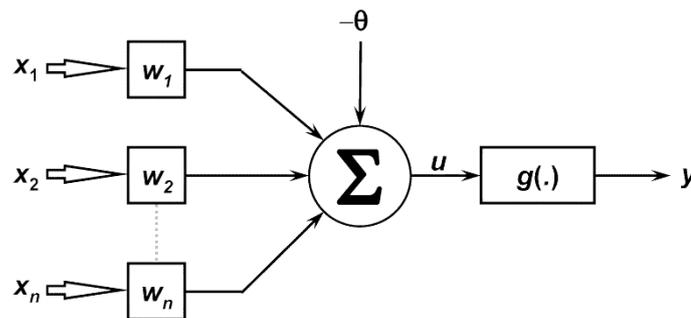


Figura 3 – Neurônio artificial (SILVA et al., 2010).

Em termos de implementação computacional, é conveniente tratar o limiar de excitação do neurônio (θ) como um peso sináptico (w_0) associado a uma entrada fixa ($x_0 = 1$), dessa maneira, o limiar torna-se um parâmetro a ser ajustado pelo algoritmo de treinamento da rede (HAYKIN, 2001; SILVA et al., 2010). Matematicamente, o modelo permanece equivalente:

$$y = g(u), \text{ sendo } u = \sum_{i=0}^n x_i w_i$$

Similar aos neurônios biológicos, os neurônios artificiais individuais possuem capacidade computacional limitada, entretanto, um conjunto de neurônios artificiais conectados entre si, constituindo uma rede, é capaz de resolver problemas de elevada complexidade. A disposição dos seus neurônios em uma rede, comumente definida como arquitetura, é caracterizada pelo número de camadas, pelo número de neurônios por camada e pelo tipo de conexão entre os neurônios. As camadas de uma

rede podem ser dos seguintes tipos: entrada (recebimento dos dados), oculta ou intermediária (extração de características) e saída (produção e apresentação do resultado final da rede). Ressalta-se que apenas os neurônios das camadas ocultas e de saída realizam o processamento descrito anteriormente, os neurônios da camada de entrada apenas recebem os dados e transmite-os para a próxima camada. Dessa maneira, uma rede é considerada de camada única, quando não possui camadas ocultas, possuindo apenas a camada de saída, uma vez que a camada de entrada não realiza processamento dos dados, enquanto que uma rede de múltiplas camadas pode ter uma ou mais camadas ocultas, além da camada de saída. As conexões entre os neurônios podem ser do tipo *feedforward* (alimentada para frente) ou recorrente (as saídas dependem das entradas e do seu valor atual) (BRAGA et al., 2007).

2.2.1 Normalização

A normalização dos dados é uma etapa de pré-processamento dos dados essencial para determinados métodos de Mineração de Dados, dentre estes as Redes Neurais Artificiais. Consiste na transformação dos valores de cada variável numérica de forma a compreenderem pequenos intervalos, como de -1 a 1 ou de 0 a 1. Esta etapa impede que valores maiores de algumas variáveis influenciem de forma tendenciosa nas estimativas, tratando todas as variáveis de forma equivalente (GOLDSCHMIDT e PASSOS, 2005). Além disso, as funções matemáticas calculadas pelos neurônios artificiais, comumente denominadas de funções de ativação, limitam a amplitude das saídas aos intervalos citados. Assim, uma normalização adequada das variáveis antes do processo de treinamento é importante para obter bons resultados e reduzir significativamente o tempo de cálculo (SOLA e SEVILLA, 1997; HAYKIN, 2001).

A normalização linear considera o valor mínimo e máximo de cada variável na transformação para valores em um intervalo fechado, por exemplo, de 0 até 1 ou de -1 a 1, mantendo as distâncias entre os valores normalizados proporcionais às distâncias entre os dados originais, cuja fórmula é (VALENÇA, 2010):

$$x' = \frac{(x - x_{min})(b - a)}{(x_{max} - x_{min})} + a$$

em que: x' é o valor normalizado; x é o valor original; x_{min} é o valor mínimo da variável, x_{max} é o valor máximo da variável; a é o limite inferior do intervalo de normalização (0 ou -1); e b é o limite superior do intervalo de normalização (1).

É necessário um pós-processamento dos dados ao final do processamento de uma rede, revertendo a normalização realizada anteriormente.

A normalização linear foi escolhida para ser implementada no pré-processamento dos valores das variáveis numéricas ou quantitativas, enquanto que as variáveis categóricas ou qualitativas são codificadas de forma a atender às necessidades de uma rede, ou seja, que os dados estejam em uma representação numérica. Dessa maneira, as variáveis categóricas cujos valores são nominais, são submetidas à codificação de categórica para numérica do tipo representação binária 1-de-N. Nesse tipo de representação, o código tem um comprimento igual ao número de categorias da variável a ser codificada, e cada elemento recebe o valor 0, exceto aquele que representa a categoria em questão (GOLDSCHMIDT e PASSOS, 2005). Por exemplo, considere que a variável “local” possui cinco categorias, a codificação é demonstrada na Tabela 1:

Tabela 1 – Codificação por representação binária 1-de-N.

Valores originais	Código
Fazenda Velha	10000
Fazenda Jacutinga	01000
Beira do Riacho	00100
Santa Cruz	00010
Providência	00001

2.2.2 Função de Ativação

A função de ativação ($g(\mathbf{u})$) define a saída do neurônio em termos do potencial de ativação (\mathbf{u}). Seu objetivo é limitar o intervalo de amplitude da saída a um valor entre 0 e 1 ou entre -1 e 1. Diversas funções de ativação podem ser utilizadas afetando o desempenho da RNA, por isso implementou-se as seguintes funções: função identidade, função sigmoidal, função tangente hiperbólica, função Log e função Seno.

A função identidade ou linear resulta no valor do próprio potencial de ativação (\mathbf{u}), lembrando que este é resultado da combinação linear do produto de cada entrada por seu respectivo peso, incluindo o limiar de excitação (\mathbf{w}_0 com $\mathbf{x}_0 = 1$). A função sigmoidal (logística) resulta em valores reais entre 0 e 1, nas

funções tangente hiperbólica, Log e Seno, este intervalo é entre -1 e 1 (BRAGA et al., 2007; VALENÇA, 2010; HEATON, 2011).

2.2.3 Treinamento

O treinamento ou aprendizado de uma RNA consiste no processo iterativo de ajuste dos seus parâmetros, ou seja, pesos e limiar de excitação (trado com um peso) a partir de apresentações sucessivas de um conjunto de exemplos a um algoritmo de treinamento, até que um critério de parada seja atingido, finalizando o processo (BISHOP, 1995; HAYKIN, 2001; BRAGA et al., 2007).

Os algoritmos propagativos de treinamento são utilizados para o treinamento de RNA que se enquadram no paradigma de aprendizado do tipo supervisionado, ou seja, o algoritmo de treinamento recebe um par de entradas e com suas respectivas saídas desejadas. O algoritmo de propagação vai passar por uma serie de iterações, visando minimizar o erro das estimativas. Em cada iteração todos os dados são apresentados à rede, a cada observação apresentada, a matriz de pesos é alterada pelas etapas conhecidas como *forward* (para frente) e *backward* (para trás). Na etapa *forward*, os valores de entrada são apresentados à rede e propagados até a camada de saída que gera uma resposta, as estimativas obtidas são comparadas com os valores observados, obtendo dessa forma o erro. Ressalta-se que nesta etapa, os pesos da rede são todos fixos. A etapa *backward* consiste na propagação do erro da camada de saída em direção à camada de entrada, ou seja, para trás através da rede, direcionando o ajuste dos pesos de acordo com uma regra de correção de erro, para que o valor da saída da rede seja o mais próxima do desejado (HAYKIN, 2001).

Diversos algoritmos podem ser utilizados para o treinamento de RNA. Buscou-se implementar maior gama de algoritmos disponíveis para facilitar na obtenção de RNA satisfatórias. Os algoritmos implementados no sistema foram: *Error BackPropagation*, *Manhattan Update Rule*, *Quick propagation* (QPROP), *Resilient Propagation* (RPROP), *Scaled Conjugate Gradient* (SCG), *Levenberg Marquardt algorithm* (LMA), Algoritmos Genéticos e *Simulated Annealing*.

Alguns algoritmos requerem a definição de parâmetros específicos como a taxa de aprendizado e o *momentum*. A taxa de aprendizado mede a rapidez com que os pesos são atualizados. O *momentum* é adicionado à equação de ajuste dos pesos para acelerar o processo de treinamento e evitar mínimos locais, reduzindo a instabilidade (BRAGA et al., 2007).

Os Algoritmos Genéticos são inspirados nos princípios da evolução natural e da reprodução genética, sendo utilizados para resolver problemas de busca e otimização, como o ajuste dos pesos de uma rede. O processo iterativo de busca por uma solução compreende as seguintes etapas: geração de uma população inicial aleatória formada por um conjunto aleatório de indivíduos (possíveis soluções); seleção de indivíduos mais adaptados; reprodução gerando novos indivíduos (cruzamentos (crossover), mutações ou recombinação genética); até a obtenção de uma solução satisfatória (critério de parada) (LINDEN, 2006).

Simulated annealing é um processo de busca e otimização que baseia-se na analogia com processos de recozimento sólidos (metalurgia). A partir de uma solução inicial e através de um mecanismo de movimento, uma nova solução é gerada e avaliada, definindo se o movimento será aceito com base em algum critério pré-estabelecido. Os parâmetros de temperatura limitam a probabilidade de aceitação de soluções “ruins”, a fim de explorar o espaço de soluções e escapar de mínimos locais (RODRIGUES et al., 2004).

O treinamento de uma rede é finalizado quando um determinado critério de parada é atingido, neste estudo, foram utilizados três critérios: erro médio, número de ciclos e convergência. Considerando a k -ésima amostra de treinamento, a função erro quadrático mede o desempenho dos valores produzidos pelos neurônios de saída (j), ou seja:

$$e(k) = \frac{1}{2} \sum_{j=1}^n \left(y_d^j(k) - y_j(k) \right)^2$$

em que y_d^j é o valor desejada da saída e y_j é o valor obtido pela rede. Considerando um conjunto de treinamento composto por p amostras, o desempenho global do algoritmo de treinamento pode ser medido pelo erro quadrático médio (e_M), ou simplesmente erro médio (SILVA et al., 2010):

$$e_M = \frac{1}{p} \sum_{k=1}^p e(k)$$

Um ciclo (ou época) corresponde à apresentação completa de todos os elementos do conjunto de treinamento acompanhada dos ajustes dos pesos da rede. Portanto, um critério de parada com base no número de ciclos, encerra o treinamento de uma rede após atingir uma determinada quantidade de ciclos. O critério

denominado convergência define o número de ciclos após o qual, se o erro médio não diminuir, o treinamento é finalizado (BRAGA et al., 2007).

3. RESULTADOS

O sistema computacional denominado por NeuroForest - Volumetric (Sistema para estimação do volume de árvores utilizando redes neurais artificiais) foi desenvolvido visando tornar-se o mais amigável possível ao usuário, admitindo certa flexibilidade na escolha de construção das RNA, normalização dos dados, escolha de algoritmos e uma consistente visualização dos resultados obtidos.

Definiu-se como entrada padrão para o treinamento das redes os parâmetros:

- Local: nome ou código do local (talhão, unidade de manejo, lote, projeto, etc.) no qual a árvore está plantada.
- Regime: regime de corte, sendo 1 para a primeira rotação (altofuste), 2 para a segunda rotação (primeira talhadia) e assim sucessivamente.
- Espaçamento: distância de plantio entre as árvores na fila e entre filas, por exemplo, 3x3.
- Clone: nome ou código do material genético ou espécie.
- Árvore: código da árvore no banco de dados de cubagem (apenas para efeito de controle do usuário).
- Idade: período de tempo do estabelecimento (plantio ou condução de brotação) até a medição, em meses ou anos.
- Dap: diâmetro à 1,30 m em centímetros.
- Altura: altura total ou comercial em metros.
- Volume: volume em metros cúbicos totalizado por alguma fórmula aproximativa a partir das medições da cubagem.

As variáveis qualitativas ou categóricas são: local, espaçamento e clone. Enquanto que as variáveis regime, idade, dap, altura e volume são quantitativas ou numéricas.

A tela inicial do sistema desenvolvido é apresentada na Figura 4.

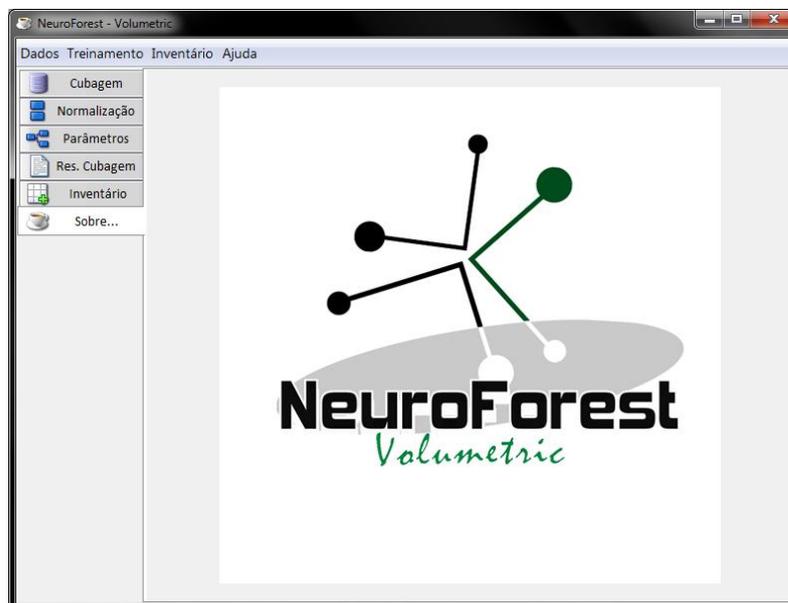


Figura 4 – Tela inicial do sistema.

No menu “Dados” são importados os dados para processamento. Após a importação os dados são apresentados na aba “Cubagem” (Figura 5). Os dados devem ser importados a partir de arquivos de texto (.txt) com as colunas seguindo a ordem do sistema. Realizada a importação dos dados e sua verificação no sistema, deve-se acionar o botão “Descrição dos Dados”, para realização do processo de descrição e avaliação dos dados (mínimos, máximos, número de classes das variáveis categóricas).

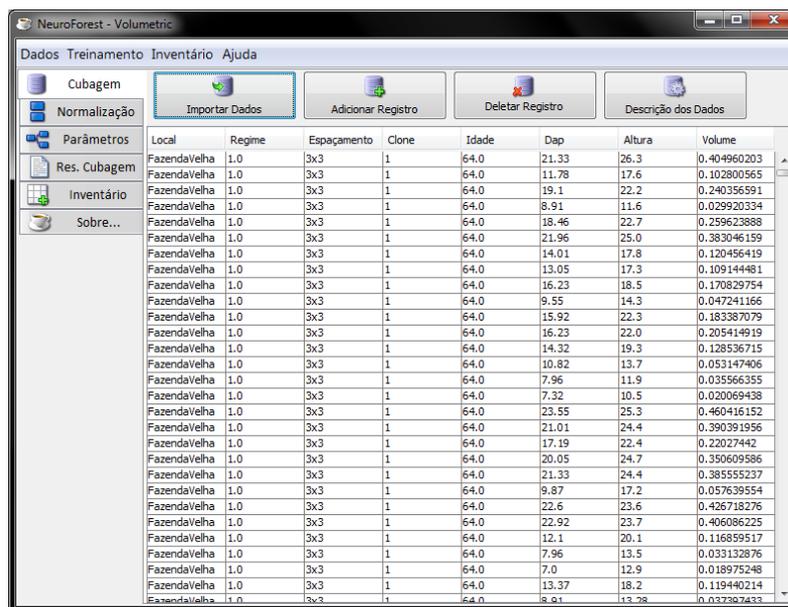


Figura 5 – Apresentação dos dados importados para o sistema.

Após a descrição dos dados o usuário é levado à aba “Normalização”, onde são específicos pelo próprio sistema, os valores máximo e mínimo de cada variável. O usuário pode especificar os valores para normalização das variáveis. Clicando no botão “Normalizar” os dados são normalizados e apresentados (Figura 6).

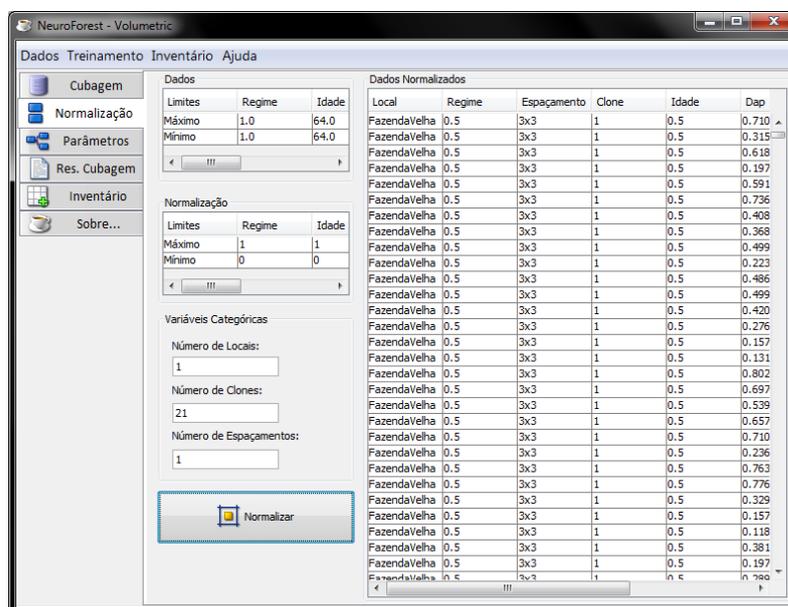


Figura 6 – Tela de normalização dos dados de treinamento.

Na aba “Parâmetros” são apresentados os principais parâmetros que podem ser ajustados para a obtenção das RNA: números de neurônios da camada oculta ou intermediária (a camada de entrada que depende exclusivamente das variáveis utilizadas e a camada de saída corresponde a apenas uma variável); função de

ativação de cada camada; número de redes treinadas (pode-se escolher treinar diversas redes ao mesmo tempo, fornecendo ao usuário a possibilidade de escolha da melhor RNA); tipo de treinamento (algoritmo de treinamento) (Figura 7).

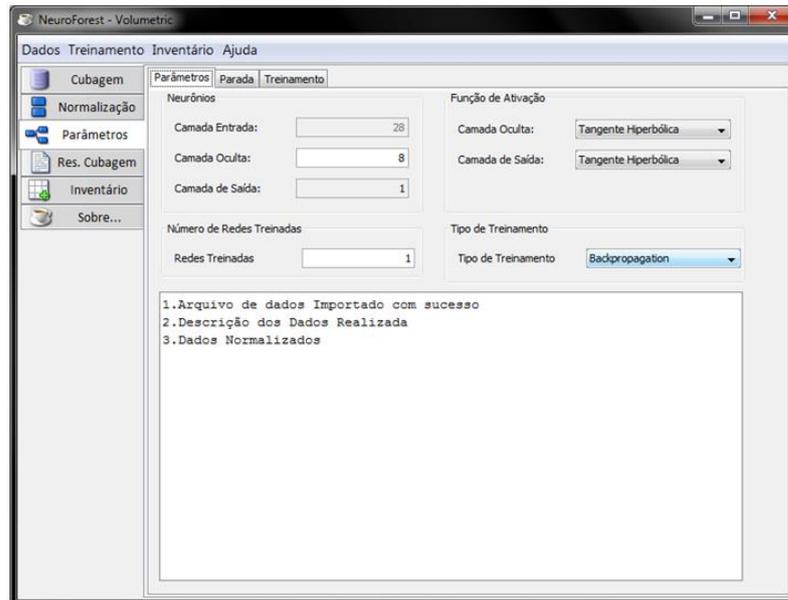


Figura 7 – Tela de parâmetros para configuração das redes neurais artificiais.

Os critérios de parada da RNA (Figura 8) e os principais parâmetros dos algoritmos de treinamento (Figura 9) podem ser especificados nessa aba.

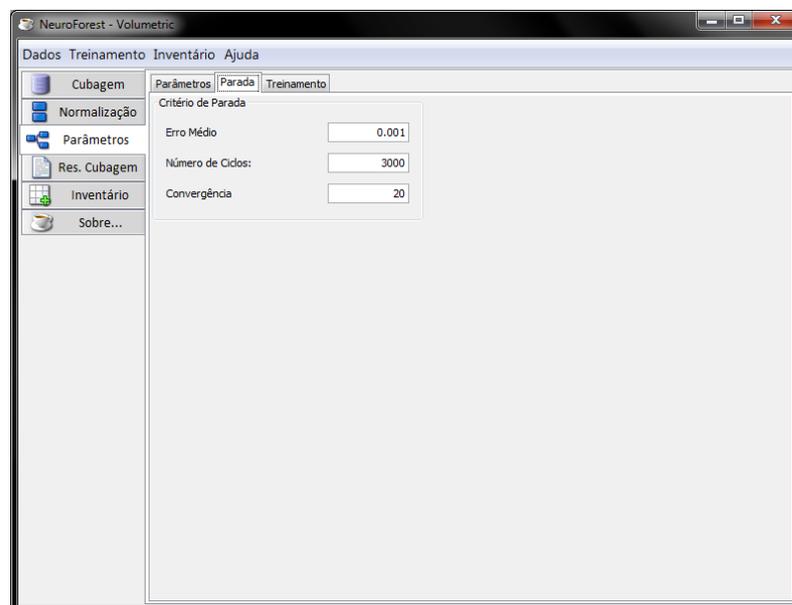


Figura 8 – Configuração dos critérios de parada para o treinamento das RNA.

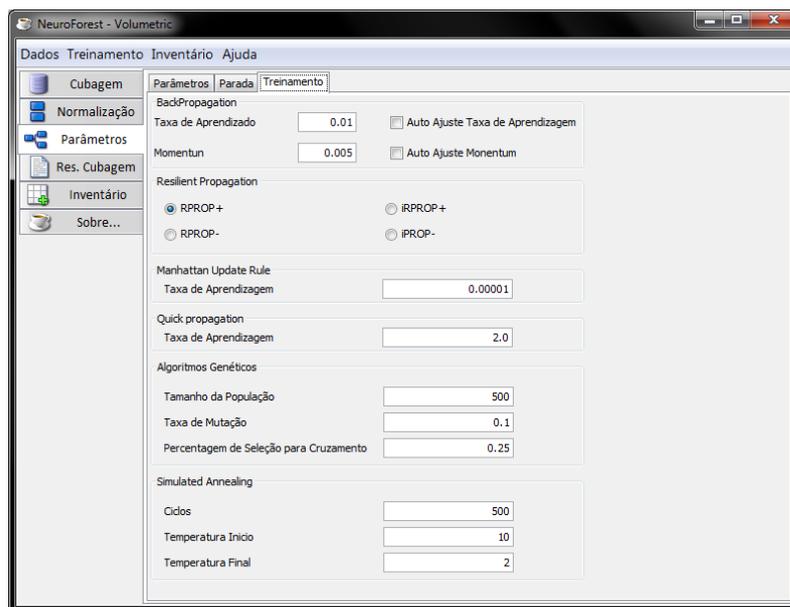


Figura 9 – Configuração dos algoritmos de treinamento das RNA.

Os resultados do treinamento das RNA são apresentados na aba “Res. Cubagem”. Caso o usuário tenha escolhido treinar mais de uma rede são apresentados os valores estimados para cada rede treinada (Figura 10). Nessa mesma aba são apresentados os gráficos de avaliação das estimativas obtidas pelas redes: correlação entre valores observados e estimados (Figura 11), resíduo ou erro percentual em função dos valores observados (Figura 12) e histograma da frequência dos erros percentuais, sendo o eixo y a proporção (ao multiplicar por 100 obtém-se a frequência percentual) e o eixo x a classe de erro percentual (Figura 13).

The screenshot shows the 'Res. Cubagem' (Res. Estimation) tab. A table displays the estimated values for 30 trained RNA networks. The columns are: Local, Regime, Espaçamento, Clone, Idade, Dap, Altura, VolumeObs..., and RNA. The data is as follows:

Local	Regime	Espaçamento	Clone	Idade	Dap	Altura	VolumeObs...	RNA
FazendaVelha	1.0	3x3	1	64.0	21.33	26.3	0.404960203	0.410
FazendaVelha	1.0	3x3	1	64.0	11.78	17.6	0.102800565	0.090
FazendaVelha	1.0	3x3	1	64.0	19.1	22.2	0.240356591	0.261
FazendaVelha	1.0	3x3	1	64.0	8.91	11.6	0.029920334	0.030
FazendaVelha	1.0	3x3	1	64.0	18.46	22.7	0.259623888	0.251
FazendaVelha	1.0	3x3	1	64.0	21.96	25.0	0.383046159	0.400
FazendaVelha	1.0	3x3	1	64.0	14.01	17.8	0.120456419	0.111
FazendaVelha	1.0	3x3	1	64.0	13.05	17.3	0.109144481	0.101
FazendaVelha	1.0	3x3	1	64.0	16.23	18.5	0.170829754	0.150
FazendaVelha	1.0	3x3	1	64.0	9.55	14.3	0.047241166	0.051
FazendaVelha	1.0	3x3	1	64.0	15.92	22.3	0.183387079	0.180
FazendaVelha	1.0	3x3	1	64.0	16.23	22.0	0.205414919	0.190
FazendaVelha	1.0	3x3	1	64.0	14.32	19.3	0.128536715	0.130
FazendaVelha	1.0	3x3	1	64.0	10.82	13.7	0.053147406	0.060
FazendaVelha	1.0	3x3	1	64.0	7.96	11.9	0.035566355	0.020
FazendaVelha	1.0	3x3	1	64.0	7.32	10.5	0.020069438	0.010
FazendaVelha	1.0	3x3	1	64.0	23.55	25.3	0.460416152	0.470
FazendaVelha	1.0	3x3	1	64.0	21.01	24.4	0.390391956	0.360
FazendaVelha	1.0	3x3	1	64.0	17.19	22.4	0.22027442	0.210
FazendaVelha	1.0	3x3	1	64.0	20.05	24.7	0.350609586	0.330
FazendaVelha	1.0	3x3	1	64.0	21.33	24.4	0.385555237	0.370
FazendaVelha	1.0	3x3	1	64.0	9.87	17.2	0.057639554	0.060
FazendaVelha	1.0	3x3	1	64.0	22.6	23.6	0.426718276	0.400
FazendaVelha	1.0	3x3	1	64.0	22.92	23.7	0.406086225	0.420
FazendaVelha	1.0	3x3	1	64.0	12.1	20.1	0.116859517	0.100
FazendaVelha	1.0	3x3	1	64.0	7.96	13.5	0.033132876	0.030
FazendaVelha	1.0	3x3	1	64.0	7.0	12.9	0.018975248	0.020
FazendaVelha	1.0	3x3	1	64.0	13.37	18.2	0.119440214	0.110
FazendaVelha	1.0	3x3	1	64.0	8.91	13.28	0.037397433	0.040

Figura 10 – Estimativas das RNA treinadas pelo sistema.

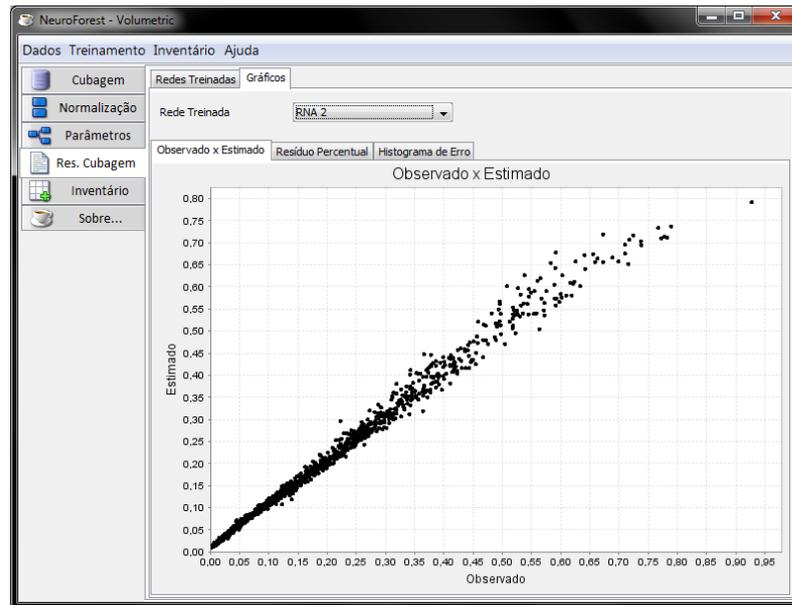


Figura 11 – Gráfico de correlação entre os valores observados e estimados.

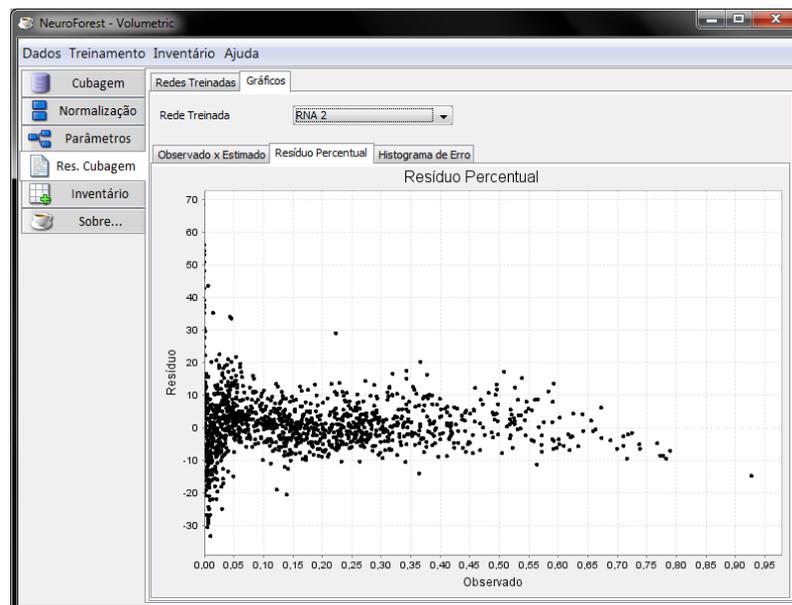


Figura 12 – Gráficos dos resíduos percentuais em função dos valores observados.

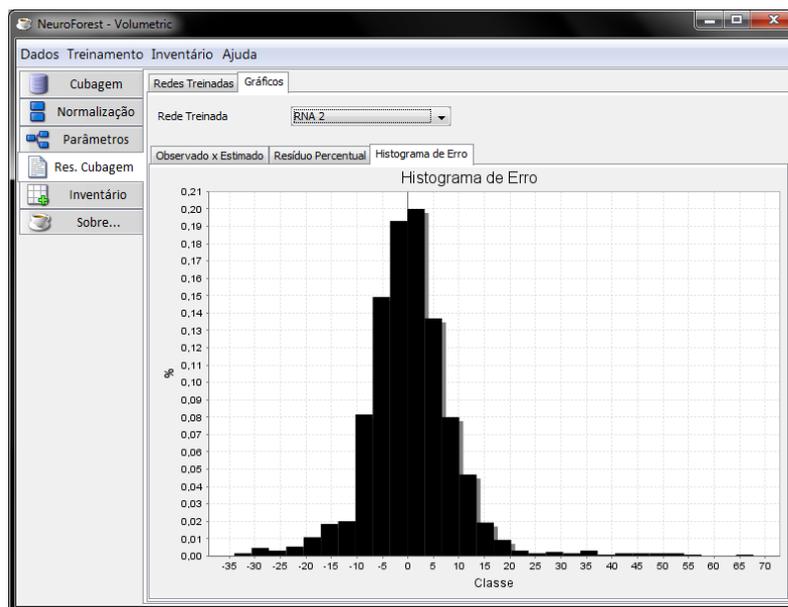


Figura 13 – Histograma de dispersão dos erros percentuais.

A aba “Inventário” permite a importação de dados de inventário, aplicação das RNA treinadas ou salvas anteriormente e a exportação das estimativas realizadas (Figura 14).

Local	Regime	Espaçamento	Clone	Idade	Dap	Altura	Volume Esti...
FazendaVelha	1	3x3	1	64.0	21.33	26.3	0.45815884...
FazendaVelha	1	3x3	1	64.0	11.78	17.6	0.09129451...
FazendaVelha	1	3x3	1	64.0	19.1	22.2	0.27972100...
FazendaVelha	1	3x3	1	64.0	8.91	11.6	0.04165842...
FazendaVelha	1	3x3	1	64.0	18.46	22.7	0.26996502...
FazendaVelha	1	3x3	1	64.0	21.96	25.0	0.45069760...
FazendaVelha	1	3x3	1	64.0	14.01	17.8	0.11610510...
FazendaVelha	1	3x3	1	64.0	13.05	17.3	0.10171008...
FazendaVelha	1	3x3	1	64.0	16.23	18.5	0.15458267...
FazendaVelha	1	3x3	1	64.0	9.55	14.3	0.05710058...
FazendaVelha	1	3x3	1	64.0	15.92	22.3	0.19718726...
FazendaVelha	1	3x3	1	64.0	16.23	22.0	0.19971185...
FazendaVelha	1	3x3	1	64.0	14.32	19.3	0.13322170...
FazendaVelha	1	3x3	1	64.0	10.82	13.7	0.06352621...
FazendaVelha	1	3x3	1	64.0	7.96	11.9	0.03577044...
FazendaVelha	1	3x3	1	64.0	7.32	10.5	0.02598007...
FazendaVelha	1	3x3	1	64.0	23.55	25.3	0.52429906...
FazendaVelha	1	3x3	1	64.0	21.01	24.4	0.39725642...
FazendaVelha	1	3x3	1	64.0	17.19	22.4	0.22908077...
FazendaVelha	1	3x3	1	64.0	20.05	24.7	0.36755863...
FazendaVelha	1	3x3	1	64.0	21.33	24.4	0.40992716...
FazendaVelha	1	3x3	1	64.0	9.87	17.2	0.07287323...
FazendaVelha	1	3x3	1	64.0	22.6	23.6	0.44103225...
FazendaVelha	1	3x3	1	64.0	22.92	23.7	0.45679147...
FazendaVelha	1	3x3	1	64.0	12.1	20.1	0.11171822...
FazendaVelha	1	3x3	1	64.0	7.96	13.5	0.04190349...
FazendaVelha	1	3x3	1	64.0	7.0	12.9	0.03182340...
FazendaVelha	1	3x3	1	64.0	13.37	18.2	0.11173236...
FazendaVelha	1	3x3	1	64.0	8.91	13.78	0.04877651...

Figura 14 – Aplicação das RNA a dados de inventário florestal contínuo.

4. DISCUSSÃO

O projeto NeuroForest conta com o desenvolvimento de outros sistemas para a aplicação de RNA, como: NeuroForest – hipsometric para a estimação da altura de árvores; NeuroForest – taper para a estimação do perfil de árvores; NeuroForest – *Prediction* para a predição da produção futura de plantios florestais; NeuroForest – Diameter Distribution utilizado para o ajuste de funções densidade de probabilidade; bem como a projeção dos parâmetros dessas funções; NeuroForest – AC para a projeção da distribuição diamétrica utilizando modelos de autômatos celulares, dentre outros.

Sistemas como o RPF, utilizado para geração de modelos de regulação florestal, OtimToras, utilizado para o corte ótimo de toras, FitFD, sistema utilizado para o ajuste de funções de densidade probabilidade e Select utilizado para a seleção de parcelas para ajuste de modelos de crescimento e produção foram desenvolvidos pelo mesmo grupo de autores do projeto Neuroforest.

Esses softwares possuem a filosofia de ser livre, ou seja, os softwares estão disponíveis para serem usados, copiados, e distribuídos, por qualquer pessoa, seja na sua forma original ou com modificações, seja gratuitamente ou com custo (HEXSEL, 2002). Este fato possibilita a utilização do sistema sem custo, visando popularizar a utilização de RNA e outras técnicas aplicadas na área florestal.

No presente estudo foi conduzido um estudo de caso de aplicação de um desses softwares, o NeuroForest – Volumetric. A flexibilidade e eficiência do software foi comprovada sendo verificado: simplicidade de uso, importação de dados e geração de resultados em formato usual na área de mensuração florestal;

comprovação da eficiência das RNA para quantificação do volume de árvores; simplicidade de aplicação da RNA selecionada à dados de parcelas de inventário.

Foram visualizadas as seguintes possibilidades de melhoria no software:

- permitir uma maior eficiência na manipulação dos dados observados, seleção de variáveis dentre outros;
- integração do sistema com ferramentas de cadastro e inventário florestal, tornando o sistema mais versátil e dinâmico;
- geração de relatórios em formato padronizado de uma maneira mais amigável ao usuário;
- integração com SGBD (Sistemas de Gerenciamento de Bancos de Dados), como o MySQL, Firebird, PostgreSQL, dentre outros, permitindo a exportação e publicação de informações em bancos de dados corporativos.

5. CONCLUSÕES

O sistema NeuroForest – Volumetric é eficiente e de uso fácil para quantificação do volume de árvores em inventário florestal.

O emprego do NeuroForest – Volumetric permite significativa redução no tempo para processar dados de cubagem e inventário, sem perda de exatidão.

6. REFERÊNCIAS BIBLIOGRÁFICAS

BISHOP, C.M. **Neural networks for pattern recognition**. Oxford University Press, 1995. 482 p.

BRAGA, A. de P.; CARVALHO, A. P. de L. F. de; LUDEMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro: Editora LTC, 2007. 226 p.

BULLINARIA, J. A. **Introduction to Neural Computation**. Notas de aula. Disponível em: <http://www.cs.bham.ac.uk/~jxb/inc.html>. 2010.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3. ed. rev. e ampl. Viçosa: UFV, 2009. 548 p.

GOLDSCHMIDT, R.; PASSOS, E. **Data mining: um guia prático**. Rio de Janeiro: Elsevier, 2005. 255 p.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 900 p.

HEATON, J. **Programming Neural Networks with Encog3 in Java**. 2 nd ed. Heaton Research, Incorporated, 2011. 240 p.

HUSCH, B.; BEERS, T. W.; KERSHAW, J. A. Jr. **Forest Mensuration**. 4th ed. New Jersey: John Wiley and Sons, 2003. 443 p.

LINDEN, Ricardo. **Algoritmos Genéticos: Uma importante ferramenta da Inteligência Computacional**. Rio de Janeiro: Editora Brasport, 2006.

RODRIGUES, F. L.; LEITE, H. G.; SANTOS, H. N.; SOUZA, A. L.; RIBEIRO, C. A. A. S. Metaheurística simulated annealing para solução de problemas de planejamento florestal com restrições de integridade. **Revista Árvore**, v.28, n.2, p.247-256, 2004.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais: para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010. 399 p.

SOARES, C. P. B.; PAULA NETO, F.; SOUZA, A. L. de. **Dendrometria e inventário florestal**. Viçosa: Ed. UFV, 2006. 276 p.

SOLA, J.; SEVILLA, J. Importance of input data normalization for the application of neural networks to complex industrial problems. **Nuclear Science**, v. 44, n. 3, p. 1464-1468, 1997.

VALENÇA, M. **Fundamentos das redes neurais**: exemplos em Java. 2 ed. ver. ampl. Olinda: Livro Rápido, 2010. 386 p.

CONCLUSÕES GERAIS

O modelo desenvolvido neste estudo, para estimação da altura de árvores em povoamentos de eucalipto com a utilização de redes neurais artificiais, para situações em que há introdução de um novo material genético, e quando já se conhece a tendência de crescimento em altura (medições em parcelas de inventário), é eficiente e resulta em grande exatidão.

O uso das RNA para estimar altura das árvores é eficiente e permite reduzir o número de medições em altura sem perda de precisão, reduzindo o custo do inventário florestal.

A metodologia que utiliza RNA para a redução do número de árvores a ser cubadas durante o processo de geração de equações volumétricas, é eficiente para a estimação do volume de árvores, sendo indicada para a obtenção do volume total com e sem casca de povoamentos de eucalipto, possibilitando a redução dos custos para a construção de equações volumétrica.

A aplicação de redes neurais artificiais para a projeção de parâmetros da distribuição Weibull, é mais eficiente do que o ajuste pelo método da máxima verossimilhança.

O modelo de distribuição diamétrica baseado em um modelo de autômatos celulares unidimensionais e redes neurais artificiais considerando ou não a simulação de desbastes, permite estimar distribuições diamétricas futuras com exatidão.

O sistema NeuroForest – Volumetric é eficiente e de uso fácil para quantificação do volume de árvores em inventário florestal.

O emprego do NeuroForest – Volumetric permite significativa redução no tempo para processar dados de cubagem e inventário, sem perda de exatidão.

A técnica de RNA apresentou resultados satisfatórios para a estimação do volume de árvores com diferentes configurações, ou seja, variando o número de neurônios na camada oculta, as funções de ativação, os algoritmos e parâmetros de treinamento. Apesar disso, RNA simples, com apenas um neurônio na camada oculta, são suficientes para estimar o volume de árvores com exatidão.