



**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FLORESTAIS**

BRUNO ANDRÉ HOYOS FURTADO BENTES

**APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS PARA ESTIMATIVA DE
VOLUME DE ÁRVORES EM PÉ EM UMA FLORESTA OMBRÓFILA DENSE NA
AMAZÔNIA ORIENTAL**

**Belém-PA
Fevereiro de 2018**



**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FLORESTAIS**

BRUNO ANDRÉ HOYOS FURTADO BENTES

**APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS PARA ESTIMATIVA DE
VOLUME DE ÁRVORES EM PÉ EM UMA FLORESTA OMBRÓFILA DENSE NA
AMAZÔNIA ORIENTAL**

Dissertação apresentada à Universidade Federal Rural da Amazônia para obtenção do título de Mestre em Ciências Florestais na área de concentração Manejo de Ecossistemas Florestais.

Orientador: D. Phil. José Natalino Macedo Silva

**Belém-PA
Fevereiro de 2018**

Bentes, Bruno André Hoyos Furtado

Aplicação de redes neurais artificiais para estimativa de volume de árvores em pé em uma floresta ombrófila densa na Amazônia Oriental / Bruno André Hoyos Furtado Bentes. – Belém, PA, 2018.

59 f.

Dissertação (Mestrado em Ciências Florestais) – Universidade Federal Rural da Amazônia, 2018.

Orientador: José Natalino Macedo Silva.

1. Redes Neurais Artificiais. 2. Inteligência Artificial. 3. Análise de Regressão. 4. Floresta natural. 5. Casca de Madeira - Volume. I. Silva, José Natalino Macedo, (orient.) II. Título

CDD – 006.32

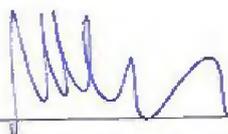
**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA
PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIAS FLORESTAIS**

BRUNO ANDRÉ HOYOS FURTADO BENTES

**APLICAÇÃO DE REDES NEURAIIS ARTIFICIAIS PARA ESTIMATIVA DE
VOLUME DE ÁRVORES EM PÉ EM UMA FLORESTA OMBRÓFILA DENSA NA
AMAZÔNIA ORIENTAL**

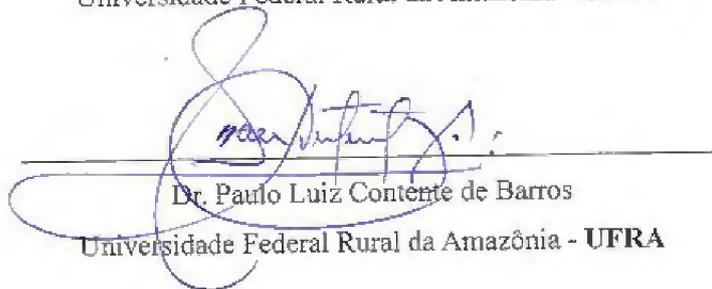
Exame de dissertação apresentada à Universidade Federal Rural da Amazônia como parte das exigências do Curso de Mestrado em Ciências Florestais, para obtenção do título de **Mestre**.

BANCA EXAMINADORA



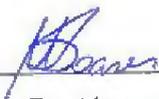
D. Phil. José Natalino Macedo Silva

Universidade Federal Rural da Amazônia - UFRA



Dr. Paulo Luiz Contente de Barros

Universidade Federal Rural da Amazônia - UFRA



Dr. Alvaro Augusto Vieira Soares

Universidade Federal Rural da Amazônia – UFRA Parauapebas



Dr. Rodrigo Geron Mendes Nascimento

Universidade Federal Rural da Amazônia – UFRA

**A minha mamãezinha, Maria Jacira Martins Furtado
Ao meu pai Lafayette de Farias Bentes Filho
Dedico**

**A Nic perfeita, a Carol alho, ao príncipe e a Lelê Bentes
Ofereço**

**A COEO
Consagro**

**A minha Família
Agradeço**

AGRADECIMENTO

Agradeço ao meu **Deus**, pela sempre presença na ausência de todos e por me proporcionar viver sonhar e realizar.

À minha amada **Mãe** Jacira por ser a flor mais linda da minha vida, e por me dar vida todos os dias;

Ao meu **Pai** Lafayette, por me amar de forma singular e incondicional;

Aos meus **filhos**: Nicole Bentes, Anna Carolina Bentes e Bruno Hoyos Bentes;

Ao meu **Irmão** Fabio Bentes e meus primos Daniel e Gabriel e irmão de coração Murilo, que são exemplo de família, por terem proporcionado os tesouros da minha vida (**sobrinhos**);

A minha **namorada** Ellen Ribeiro Bentes, por estar presente em todos os momentos;

Aos meus **amigos do peito** Marcos Almeida e Fabio Leão, por me ajudarem na conquista deste sonho;

Aos meus **amigos especiais** Carlos Richelli, Fabrícia e suas famílias;

Aos meus **grandes amigos** de curso especialmente, Mario, Larissa Badalo, Raniere, Thays, André, Ivy, Helton, Heitor, Alberto primo, Wagner, Graciliano OC, Wander, Denes, Andreza, Arllen, Vivian, Laís, Mayra, Luiz pdc, Luiz Carioca, Bruninho, Igor, Pradinho, Anhô, Paulinha e Odneila;

Ao **Professor** Dr. Fernando Cristovam Jardim (in memoriam) pela dedicação as Ciências Florestais;

Ao meu **orientador** D. Phil. José Natalino Macedo Silva por acreditar nesse engenheiro pé duro;

Ao meu **orientador de estagio** Robson e a **banca examinadora**: Professores Paulo Contente, Rodrigo Geroni e Álvaro Soares;

Ao **LABECOS**, por todo aprendizado e espaço disponibilizado ao longo do curso. Em especial ao Prof. Assis;

Ao **PPGCF** por me permitir realizar essa pesquisa e conquistar o título de Mestre;

A **Capes** pela bolsa de estudos;

A **empresa** Isabel Madeiras do Brasil Ltda. por ter cedido os dados para pesquisa;

A **empresa** DAP Florestal por fornecer o software Neuro 4 para os estudantes de forma gratuita.

A **empresa** AMPE Ambiental por toda experiência laboral, importante nesta conquista.

Aos que **prometi** sintam-se homenageados pois foram ferramentas importantes nessa conquista.

A todos que colaboraram para a realização deste estudo.

LISTA DE FIGURAS

Figura 1: Representação de um neurônio biológico	18
Figura 2: Representação de um neurônio artificial.....	19
Figura 3: Funções de ativação: a) função de limiar, b) função linear por partes e c) função sigmóide (Haykin, 2001)	20
Figura 4: Mecanismo de aprendizado supervisionado com professor.....	23
Figura 5: Distribuição de resíduos dos modelos utilizados para o ajuste da equação para determinação do volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará	44
Figura 6: Gráfico de Valores observados por cubagem rigorosa e estimados por treinamento de RNA para volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará.....	48
Figura 7: Análise gráfica da comparação do erro padrão e média dos volumes observados e estimados pelos diferentes métodos aplicados para determinar o volume de árvore em pé na UMF Água Azul II	51

LISTA DE TABELAS

Tabela 1: Número de ocorrências de indivíduos por classes de diâmetro, das árvores amostra, medidas na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará.....	31
Tabela 2: Estatística descritiva, número de indivíduos (N) por espécie utilizada na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará.....	31
Tabela 3: Modelos testados para estimar o volume das árvores em pé coletados na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará.....	34
Tabela 4: Variáveis utilizadas no treinamento das redes neurais artificiais (RNA) na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará.....	39
Tabela 5: Estatística descritiva da equação com fator de forma médio calculado para UMF Água Azul II e 0,7 calculado do Heinsdijk & Bastos (1963) para Amazônia.....	42
Tabela 6: Coeficientes dos modelos testados com a variáveis DAP e Hc com seus respectivos, R ² ajustado, Syx, Syx%, Syxr, Syxr%, DMP, pa estimativa do volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará.....	42
Tabela 7: Caracterização das Redes Neurais Artificiais (RNA) para a estimativa do volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará.....	46
Tabela 8: Precisão das redes neurais artificiais selecionadas para volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará.....	47
Tabela 9: Resultados estatístico e validação das estimativas do volumes de árvores em pé na UMF Água Azul II, por meio dos teste F, teste t e χ^2 a 95% de probabilidade.....	49
Tabela 10: Comparação entre o Volume e as predições de volume, por fator de forma médio, regressão linear e treinamento de RNA de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará, obtidas por diferentes métodos.....	50

SUMÁRIO

RESUMO	11
ABSTRACT	12
1. INTRODUÇÃO	13
1.1. QUESTÃO CIENTÍFICA	15
1.2. HIPÓTESE	15
1.3. OBJETIVOS	15
2. REVISÃO DE LITERATURA	16
2.1. MODELOS VOLUMÉTRICOS.....	16
2.2. REDES NEURAIS ARTIFICIAIS	17
2.2.1. Neurônios.....	18
2.2.2. Arquiteturas.....	21
2.2.3. Aprendizado	21
a) Supervisionado	23
b) Não supervisionado	25
2.2.4. Redes Multilayer Perceptron.....	25
2.2.5. Algoritmo de treinamento das redes MLP - Backpropagation	27
2.3. APLICAÇÃO DA RNA EM FLORESTA NATURAL.....	28
3. MATERIAL E MÉTODOS	30
3.1. FONTE DOS DADOS DE ÁRVORES AMOSTRAS.....	30
3.2. DISTRIBUIÇÃO DE FREQUÊNCIA DAS ÁRVORES AMOSTRADAS POR CLASSE DE DIÂMETRO.....	31
3.2.1. Estatística Descritiva.....	32
3.3. EQUAÇÃO DE VOLUME COM FATOR DE FORMA MÉDIO DA UMF.....	33
3.4. EQUAÇÃO DE VOLUME ESTIMADO QUE UTILIZA O FATOR DE FORMA 0,7 CALCULADO POR HEINSDIJK & BASTOS (1963)	33
3.5. MODELOS VOLUMÉTRICOS TESTADOS	34
3.6. CRITÉRIOS USADOS PARA AVALIAÇÃO DA QUALIDADE DE AJUSTE DOS MODELOS REGRESSÃO.....	34
3.7. TREINAMENTO REDES NEURAIS ARTIFICIAIS	38
3.8. VALIDAÇÃO DAS EQUAÇÕES E RNAS	41
4. RESULTADOS E DISCUSSÃO	42
4.1. EQUAÇÃO DE VOLUME COM FATOR DE FORMA MÉDIO E EQUAÇÃO DE VOLUME ESTIMADO COM FATOR DE FORMA 0,7.....	42
4.2. MODELOS TESTADOS POR REGRESSÃO	42
4.3. NORMALIDADE DOS RESÍDUOS	45

4.4. TREINAMENTO DE RNAs	45
4.5. VALIDAÇÃO DA EQUAÇÃO COM FATOR DE FORMA MÉDIO, EQUAÇÃO COM FATOR DE FORMA 0,7 USADO POR HEINSDIJK & BASTOS, MODELO SHUMAKER E HALL E RNA	49
4.6. ESCOLHA DOS MÉTODOS PARA ESTIMATIVA DO VOLUME	50
5. CONCLUSÕES	52
REFERENCIAS.....	54

RESUMO

Este estudo teve como objetivo comparar as metodologias utilizadas para mensurar o volume com casca de madeira em pé em florestas naturais, por meio da análise de regressão, fator de forma médio, fator de forma 0,7 proposto por Heinsdijk & Bastos (1963) e treinamento de Redes Neurais Artificiais em uma floresta ombrófila densa no município de Breu Branco-PA. Foram cubadas 576 árvores pela metodologia de Smalian e utilizadas as informações de 476 árvores, para estimativa dos coeficientes das equações volumétricas e fator de forma médio, também foram utilizadas 100 árvores diferentes das 476 arvores para validação e normalidade dos dados. Foram ajustados três modelos de simples entrada e três de dupla entrada. Os critérios utilizados para seleção dos melhores modelos foram R^2_{aj} , $Sy_x\%$, $Sy_{xr}\%$, RQME, AIC, VP e distribuição dos resíduos. Foram treinadas 3000 RNA para cada análise de estimativa de volume da árvore, utilizando variáveis quantitativas (DAP e Altura comercial) e qualitativa (espécie). Para a melhor estimativa da RNA foram utilizados os parâmetros RQME, Coeficiente de Correlação, Viés e análise de gráfico. Todas as estimativas de volumes foram comparado com o Volume cubado pela metodologia Smalian e verificou-se que neste estudo não obtiveram diferença significativa a uma probabilidade de 95% entre elas.

Palavras-Chave: Relações quantitativas, Regressão, Inteligência Artificial, Floresta natural.

ABSTRACT

The objective of this study was to compare the methodologies used to measure the volume of standing bark in natural forests, using regression analysis, mean shape factor, 0.7 form factor proposed by Heinsdijk & Bastos (1963) and artificial neural networks training in a dense ombrophilous forest in the municipality of Breu Branco-PA. A total of 576 trees were sown by the Smalian methodology and the information from 476 trees was used to estimate the coefficients of the volumetric equations and average form factor. 100 trees were also used, different from the 476 trees for data validation and normality. Three models of single entry and three of double entry were adjusted. The criteria used to select the best models were R^2_{aj} , $Sy_x\%$, $Sy_{xr}\%$, RQME, AIC, VP and waste distribution. Three thousand RNAs were trained for each analysis of tree volume estimation, using quantitative variables (DBH and commercial height) and qualitative (species). RQME parameters, Correlation Coefficient, Bias and graph analysis were used for the best RNA estimation. All volume estimates were compared to the volume cubed by the Smalian methodology and it was found that in this study there was no significant difference at a 95% probability between them.

Key words: Quantitative relations, Regression, Artificial Intelligence, Natural forest.

1. INTRODUÇÃO

O Brasil é um país com aproximadamente 58% do seu território coberto por florestas naturais e plantadas - o que representa a segunda maior área de florestas do mundo, atrás apenas da Rússia. São estimados 485,8 milhões de hectares de florestas naturais (BOLETIM SNIF 2016) e 7,84 milhões de hectares de florestas plantadas (IBA, 2017). A estimativa da área de floresta apresentado pelo Serviço Florestal Brasileiro se baseia na classificação dos diferentes tipos de formações vegetacionais (fitofisionomias) como floresta ou não, de acordo com as definições do Manual Técnico da Vegetação Brasileira - IBGE compatibilizadas com a definição da FAO.

As florestas naturais brasileira se distribuem em biomas continentais, sendo que a Amazônia, o Cerrado e a Caatinga detém 93% do total da cobertura florestal do país (BOLETIM SNIF 2016). O Brasil é o maior produtor de madeira no mundo, com uma produção total estimada em 30 milhões de metros cúbicos de toras, sendo que a região amazônica concentra a produção de madeira nativa, ao passo que a produção proveniente de plantações ocorre no sudeste e sul do país (ITTO, 2014). Segundo Industria Brasileira de Árvores - IBÁ (2016), cerca de 91% de toda a madeira utilizada para fins industriais no país advêm de florestas plantadas, e apenas 9% da madeira industrial é produzida de matéria-prima originária de florestas naturais.

O manejo florestal tanto do ponto de vista ambiental, como do econômico é a atividade produtiva que garante a legalidade e sustentabilidade da produção de madeiras ao longo do tempo, o mesmo acontece no ponto de vista econômico (SANTANA, 2012; REIS et al., 2015). Na busca por essa sustentabilidade, o planejamento assume papel fundamental. Entre as diversas fases do planejamento, o inventário florestal é a base para tomada de decisão sobre o uso da planta (SANQUETA et al., 2014). Neste contexto, uma boa estimativa dos volumes das árvores é condição primordial para fazer uma contra avaliação do estoque de madeira e avaliação do potencial produtivo da floresta (RIBEIRO et al., 2014).

Nesse contexto, as equações de volume de madeira tornaram-se uma ferramenta obrigatória para a estimativa do estoque de madeira em planos de manejo (RIBEIRO et al., 2014). O tema volumetria já foi bastante difundido por vários pesquisadores como, por exemplo, Spurr (1952), Husch (1971). No Brasil, os estudos de equações de volume foram tratados mais profundamente por diversos autores, e. g. Paula Neto (1977), Siqueira (1977), Higuchi et al., (1982) e Machado (1979). Na Amazônia, exemplos de estudos dessa natureza são os trabalhos de Baima et al., (2001), Menezes & Silva, (2003), Barreto (2014), Silva et al., (1984), Ribeiro et al., (2014). Recentemente, Figueiredo Filho et al., (2014) lançaram um

compendio reunindo os esforços nacionais no desenvolvimento de equações volumétricas e afilamento do tronco para diferentes regiões brasileiras.

O Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis – IBAMA, desde 31 de dezembro de 2002, por meio da Instrução Normativa nº 30, começou a exigir que a partir do 2º Plano Operacional Anual - POA o cálculo do volume de árvores em pé, só seria aceito mediante a utilização de equação de volume desenvolvida especificamente na área do PMFS.

Mais tarde a IN nº 30 foi revogada, entretanto o IBAMA determinou em sua IN nº 5, de 11 de dezembro de 2006, que: “o volume comercial das árvores derrubadas para aproveitamento, seria estimado por meio de equações volumétricas, desenvolvidas no Plano de Manejo Florestal Sustentável – PMFS”. O Ministério do Meio Ambiente ratificou essa exigência por meio da Resolução do CONAMA nº 406, de 2 de fevereiro de 2009, em seu Art. 10, o que era determinado pela IN nº 5/2006, ampliando por força de lei, as possibilidades de quantificação dos estoques volumétricos de árvores em pé, com precisão.

No Estado do Pará a Secretaria de Meio Ambiente e Sustentabilidade - SEMAS/PA, ratificou aquela exigência pela IN 5 de 10 de setembro de 2015, que em seu Art. 24º, parágrafo único, estabelece: “A partir do segundo Plano de Operação Anual - POA somente será aceito pela SEMAS/PA o volume de árvores em pé, calculado mediante equação de volume desenvolvida especificamente para a Unidade de Manejo Florestal - UMF”.

A grande variabilidade na forma do fuste tem significativa implicação para a determinação do volume. Nesse sentido, para a estimativa do volume em inventários florestais são utilizados valores padronizados de fatores de formas sem levar em consideração as diferenças existentes entre as espécies, o que resulta em estimativas tendenciosas do estoque madeireiro (LANSSANOVA et al., 2013).

Como alternativa a equações de volume, o uso da inteligência artificial, as Redes Neurais Artificiais _RNA tem sido usada mais recentemente. Essa ferramenta vem aprimorando técnicas de biometria e inventário florestal permitindo maior acurácia e precisão nas estimativas de crescimento e produção, dando suporte à tomada de decisão no manejo de florestas (PENG & WEN, 1999; CASTELLANOS et al., 2007).

A grande vantagem do uso desta tecnologia refere-se à modelagem de problemas complexos, a qual é determinada pelo mecanismo de treinamento (aprendizado) supervisionado (estímulo: entrada/saída) ou não supervisionado (baseado somente em estímulos de entrada), ou seja, uma vez concluído o treinamento, a rede armazena, em sua estrutura, o modelo do problema abordado, a princípio, não importando quão complexo ele é (HAYKIN, 2008).

Este estudo teve como objetivo avaliar a aplicação das RNA no cálculo de volumes individuais, e comparar os resultados obtidos com modelos matemáticos ajustados por métodos de regressão linear de simples e múltipla entrada e equação de volume com uso de fator de forma 0,7 para estimativa do volume de árvores em pé em uma floresta ombrófila densa na Amazônia Oriental.

1.1. QUESTÃO CIENTÍFICA

No presente trabalho pretende-se responder à seguinte questão científica:

A aplicação do treinamento de Redes Neurais Artificiais melhora a precisão da estimativa de volumes de árvores em pé em floresta natural amazônica?

1.2. HIPÓTESE

O uso de RNA para o cálculo dos volumes de árvores individuais em uma floresta natural da Amazônia resulta em uma maior precisão que os métodos tradicionais de regressão e uso do fator de forma;

1.3. OBJETIVOS

a. Geral

Avaliar a eficiência da técnica de RNA diante da metodologia tradicional, que utiliza modelos de regressão de simples e dupla entrada, equação utilizando o fator de forma médio e fator de forma 0,7 para estimativa dos volumes de árvores em pé em uma floresta ombrófila densa na Amazônia oriental.

b. Específicos

- Ajustar modelos volumétricos para estimar o volume de árvores individuais;
- Ajustar o fator de forma médio para estimar o volume de árvores individuais;
- Treinamento da RNA para melhorar a precisão de estimativa dos volumes de árvores individuais;
- Comparar os volumes determinados por cubagem rigorosa, com as metodologias de regressão linear, equação com fator de forma médio, equação com fator de forma 0,7 e com treinamento de RNA.

2. REVISÃO DE LITERATURA

2.1. MODELOS VOLUMÉTRICOS

Os modelos volumétricos ajustados por regressão proporcionam melhores métricas de qualidade de ajuste para estimativa de volume em inventários florestais em comparação ao método mais usado na Amazônia para a estimativa do volume de madeira da floresta, que utiliza o volume do cilindro multiplicado por um fator de forma médio igual a 0,7 que foi difundido por Heinsdijk & Bastos (1963), a melhor qualidade de ajuste se deve, entre outras coisas, à seleção criteriosa para cubagem de árvores amostra. ((HIGUCHI et al., 2004; CAMPOS & LEITE, 2009).

Nos cálculos dos volumes das árvores em inventários florestais, em muitos casos, são utilizados valores padronizados de fatores de formas sem levar em consideração as diferenças existentes entre as espécies, o que resulta em estimativas tendenciosas do estoque madeireiro (LANSSANOVA et al., 2013). Segundo Barros & Silva Júnior (2009), o ajuste de equações volumétricas desenvolvidas especificamente para as áreas de Planos de Manejo Florestal Sustentável se constitui em um procedimento mais eficiente, econômico e com precisão aceitável para a quantificação da produção volumétrica.

Segundo Schneider et al., (2009), os modelos ajustados são avaliados por meio de vários critérios estatísticos, permitindo-se chegar a uma equação com a melhor precisão possível. Leite e Rezende (2010) divide os modelos lineares em duas categorias, modelos lineares de simples e múltiplas entradas, os modelos lineares simples são aqueles que apresentam apenas uma variável independente para estimar com precisão o volume das árvores. No caso essa variável é o diâmetro medido a 1,30 m do solo (DAP); os modelos lineares múltiplos apresentam duas ou mais variáveis independentes, ou seja, além do DAP inclui-se a altura comercial (HC), podendo ser combinada ou não com o DAP.

O modelo de volume de madeira em função do DAP e HC proposto por Schumacher & Hall é um dos mais difundidos na área florestal, devido às suas propriedades estatísticas, facilidade de ajuste e uso, além da geração de estimativas não tendenciosas. (LEITE & ANDRADE, 2002; CAMPOS & LEITE, 2013).

Desta maneira, as técnicas estatísticas surgem nesse cenário como importantes fontes de produção de conhecimento, principalmente para estimação do volume comercial, em que o uso de equações de volume em inventário florestal vem se constituindo em operação

rotineira para cálculo de volume de madeira em pé das árvores, por meio da relação DAP e HC (VALENTE et al., 2011).

2.2. REDES NEURAIS ARTIFICIAIS

As RNA's visam, em sua maioria, solucionar problemas complexos, modelando sistemas através de conexões que possam simular o sistema nervoso humano, abrangendo a capacidade que o mesmo possui de aprender e agir perante as mais adversas situações apresentadas, adquirindo conhecimento por meio da experiência e da observação (HAYKIN, 1998).

Segundo Akkoç (2012), as RNA's são técnicas computacionais baseadas na estrutura neural de organismos inteligentes e que adquirem conhecimento por intermédio de experiências. Para Tafner & Xerez (1996) uma rede neural artificial pode ser definida como uma rede massivamente paralela de elementos interconectados em suas organizações hierárquicas que estão preparadas para iterar com objetos do mundo real do mesmo modo que um sistema nervoso biológico o faz.

Segundo Graupe (2013) os princípios importantes apresentados na década de 40 por McCulloch e Pitts, orientaram os desenvolvimentos atuais de RNA, como a estrutura fixa de sinapse, a existência de conexões inibitórias que previnem a ativação de neurônios, a necessidade de saída de cada neurônio ser binária, entre outros.

Segundo Haykin (2001), uma rede neural é um processador maciçamente e paralelamente distribuído, constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso. De forma mais geral, uma RNA é uma máquina que é projetada para modelar a maneira com o cérebro realiza uma tarefa particular ou função de interesse, utilizando componentes eletrônicos ou programação em um computador.

Segundo Valença (2005) a RNA tenta reproduzir as funções das redes biológicas, buscando implementar em seu comportamento uma similaridade com o sistema biológico básico e sua dinâmica de aprendizado, assim entender melhor como funciona o sistema nervoso. Aperfeiçoando as redes e buscar ideias e inspirações para a pesquisa em neuro-computação.

Segundo Haykin (2001) e Braga et al., (2007) as principais características das RNA são as que seguem:

1 - Capacidade de "aprender" através de exemplos e reconhecer elementos similares que não foram apresentados no conjunto de dados;

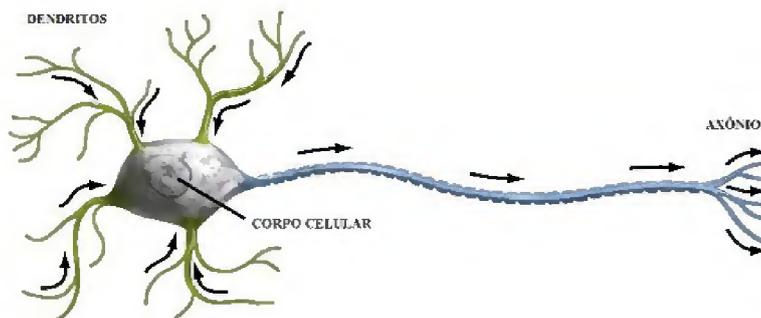
- 2 – Bom desempenho em tarefas pouco ou mal definidas, onde falta o conhecimento explícito de como resolvê-las;
- 3 – Robustez na presença de informações falsas ou ausentes;
- 4 – Escolha dos elementos no próprio conjunto de treinamento, no contexto de classificação dos padrões;
- 5 – Fornecem informações sobre quais padrões selecionar em função do grau de confiança apresentado e a tolerância à falha.

2.2.1. Neurônios

A RNA consiste em um modelo computacional biologicamente inspirado, constituído de elementos de processamento simples, capaz de gerar respostas a determinado conjunto de dados, desde que seja treinada (BINOTI, 2012). Os elementos computacionais ou unidades processadoras, conhecidas como neurônios artificiais, são modelos bem simplificados dos neurônios biológicos (SILVA et al., 2010).

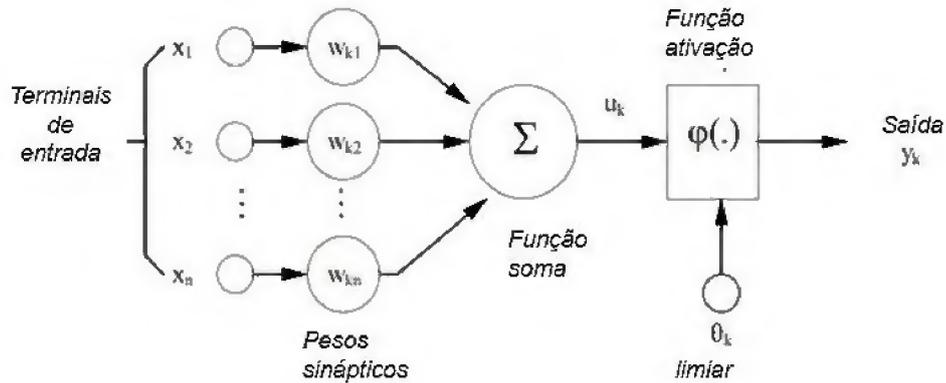
Os neurônios biológicos podem ser descritos em três seções: o corpo da célula, os dendritos e o axônio. Os dendritos têm por função receber as informações, ou impulsos nervosos, oriundas de outros neurônios e conduzi-las até o corpo celular. A informação é processada, e novos impulsos são gerados. Esses impulsos são transmitidos a outros neurônios, passando por meio do axônio até os dendritos dos neurônios seguintes. O ponto de contato entre a terminação axônica de um neurônio e o dendrito de outro é chamado de sinapse. É pelas sinapses que os neurônios se unem funcionalmente, formando redes neurais (Figura 1). As sinapses funcionam como válvulas, e são capazes de controlar a transmissão de impulsos, isto é, o fluxo da informação entre os neurônios na rede neural. (BRAGA et al., 2000).

Figura 1: Representação de um neurônio biológico



Fonte: GUYTON & HALL (2006)

Figura 2: Representação de um neurônio artificial



Fonte: HAYKIN (2001)

A representação de neurônio artificial é mostrada na figura 2, trata-se da representação de um modelo proposto por McCulloch e Pitts na década de 40. Neste modelo as entradas (w_{kn} ; U_k) são combinadas, usando uma função ϕ para produzir um estado de ativação do neurônio. Essa função vai produzir a saída do neurônio (y_k), para um valor auxiliar U_k é geralmente usado para representar uma polarização, valor abaixo do qual a saída é nula. Esse neurônio resultou em um modelo matemático com n terminais de entrada (x_1, x_2, \dots, x_n) e apenas um terminal de saída y . (SILVA et al., 2010).

Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados (w_1, w_2, \dots, w_n), cujos valores podem ser positivos ou negativos, dependendo serem as sinapses correspondentes serem inibitórias ou excitatórias, o efeito de uma sinapse particular i no neurônio pós-sináptico é dado por $x_i \cdot w_i$, os pesos determinam "em que grau" o neurônio deve considerar sinais de disparo que ocorrem naquela conexão (VALENÇA, 2011). A saída de um neurônio artificial é, então, dada pela Eq.01 (SILVA et al., 2010).

$$S_i = \sum_j^n w_{kn} x_j \quad \text{Eq. 01}$$

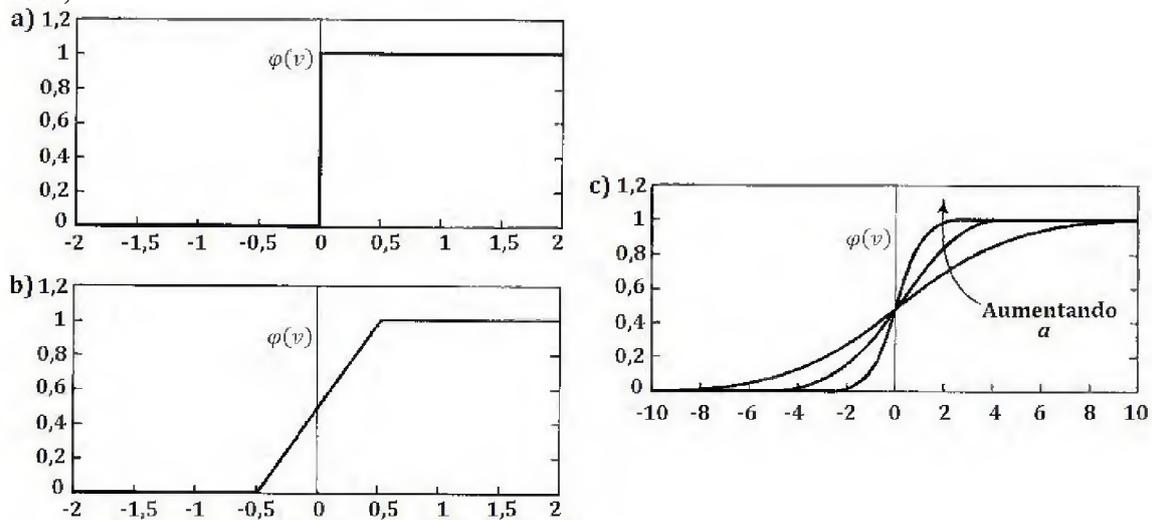
em que: S_i - saída linear do neurônio i ; x_j - entradas do neurônio; w_{kn} - pesos do neurônio.

Existem três tipos básicos de função de ativação: função de limiar, função linear por partes e função sigmóide (Figura 3). O neurônio com a função de ativação de limiar terá saída igual a 1 se $v \geq 0$ ou a 0 se $v < 0$. Tal neurônio é conhecido como o modelo de McCulloch e Pitts. Na função linear por partes, a saída será 1 se $v \geq +1/2$, se $+1/2 > v > -1/2$ e 0 se $v \leq -1/2$. Enquanto que a função de ativação sigmóide (Figura 3) é a mais comum na construção de redes neurais artificiais. (SILVA et al., 2010).

$$r = \frac{1}{1 + \exp(-ar)} \quad \text{Eq. 02}$$

onde a é o parâmetro que determina a inclinação da função sigmóide. Uma função sigmóide assume um intervalo contínuo de valores entre 0 e 1.

Figura 3: Funções de ativação: a) função de limiar, b) função linear por partes e c) função sigmóide (Haykin, 2001)



Fonte: HAYKIN (2001)

2.2.2. Arquiteturas

Um dos aspectos importantes sobre a eficiência da RNA, é sua arquitetura, isto é, a quantidade de neurônios e a forma com que são conectados. Biologicamente, as redes neurais são organizadas de forma tridimensional por componentes microscópicos, contendo bilhões de neurônios. Já em uma RNA, existe uma forte restrição do número de neurônios e camadas devido ao enorme esforço computacional requerido (HAYKIN, 2001).

Uma rede neural pode ter uma ou várias camadas. As redes que possuem uma única camada são as que possuem um nó entre uma entrada e uma saída da rede. Esse tipo de rede é indicado para a solução de problemas linearmente separáveis. Já as redes multicamadas possuem mais de uma camada entre as já existentes camadas de entrada e saída (HAYKIN, 2001).

Segundo Haykin (2001) o neurônio ou modelo neuronal é uma unidade de processamento de informações que é fundamental para operação de uma RNA. A forma como está estruturada está intimamente ligada ao algoritmo de aprendizagem usado para treinar a rede. O algoritmo de aprendizagem utilizado no projeto de RNA, como sendo estruturado, a classificação passa a ser considerada como “Camadas”, onde os neurônios estão organizados na forma de camadas. Na forma mais simples de uma rede em camadas, temos uma camada de entrada de nós de fonte que se projeta sobre uma camada de saída de neurônio, mas não vice-versa. Esta rede é chamada de rede de camada única, sendo que a designação “camada única” se refere à camada de saída de neurônios. (HAYKIN, 2001).

2.2.3. Aprendizado

Uma RNA aprende acerca do seu ambiente por meio de um processo iterativo de ajuste aplicados a seus pesos sinápticos em níveis de viés. Idealmente, a rede se torna mais estrita sobre seu ambiente após cada interação do processo de aprendizagem (HAYKIN, 2001).

Segundo Cabacinha & Castro (2009), a capacidade de aprendizagem é uma característica marcante da RNA. Ela aprende por meio de um processo iterativo de ajuste de pesos e limiares (viés). As RNA possuem a capacidade de aprendizado conexionista, ou seja, não se procura obter regras como na abordagem simbólica da Inteligência Artificial (IA), mas sim determinar a intensidade de conexões entre neurônios (BINOTI, 2012). As RNAs são aproximações brutas das redes e neurônios encontrados no cérebro humano.

Dentre as características de similaridade das RNA com o cérebro humano, está a capacidade de aprender algo por meio de exemplos do ponto de vista prático, uma RNA consiste em várias unidades de processamento simples (neurônios artificiais) conectadas entre si e dispostas em camadas formando um sistema computacional paralelo para executar uma determinada tarefa (BINOTI, 2012; BULLINARIA, 2014).

Quanto à funcionalidade de uma RNA em multicamadas, Haykin (2001) e Valença (2011), afirmam que o processamento realizado por cada neurônio é definido pela combinação dos processamentos realizados pelos neurônios da camada anterior que estão conectados a ele. Quando se segue da primeira camada intermediária em direção à camada de saída, as funções implementadas se tornam cada vez mais complexas. Estas funções definem como é realizada a divisão do espaço de decisão.

Um conjunto de procedimentos bem definidos para adaptar os parâmetros de uma RNA para que a mesma possa aprender uma determinada função é chamado de algoritmo de aprendizado (HAYKIN, 2001). Existem vários algoritmos de aprendizado, cada qual com suas vantagens e desvantagens. Estes algoritmos basicamente se diferem na maneira pela qual o ajuste dos pesos é feito (GORGENS et al., 2009; BINOTI et al., 2012).

É evidente que uma RNA extrai seu poder computacional, de uma estrutura maciça e paralelamente distribuída, segundo sua habilidade de aprender e, portanto, de generalizar (HAYKIN, 2001). A generalização se refere ao fato de a RNA produzir saídas adequadas para as entradas que não estavam presentes durante o treinamento (aprendizagem). Estas duas capacidades de processamentos de informação tornam possível para a RNA resolver problemas complexos, de grande escala, que são atualmente intratáveis (HAYKIN, 2001).

Na prática, as RNA não podem fornecer uma solução trabalhando individualmente, em vez disso, elas precisam ser integradas em uma abordagem consistente de engenharia de sistemas. Especificamente um problema complexo de interesse é decomposto em um número de tarefas relativamente simples e atribui-se à RNA um subconjunto de tarefas que coincidem às suas capacidades inerentes (HAYKIN, 2001).

O aprendizado ocorre quando a rede obtém uma solução generalizada para uma classe de problemas. Esta generalização é obtida por meio de um conjunto de regras bem definidas para a solução de um problema, onde a rede assimila uma determinada função. Estas regras são denominadas de algoritmo de aprendizado ou de treinamento (HAYKIN, 2001).

Segundo Haykin (2001), para o aprendizado das redes neurais, foram propostos diversos métodos de treinamento, sendo estes subdivididos em dois paradigmas principais: o aprendizado supervisionado e o não supervisionado. Para estes modelos existem vantagens e

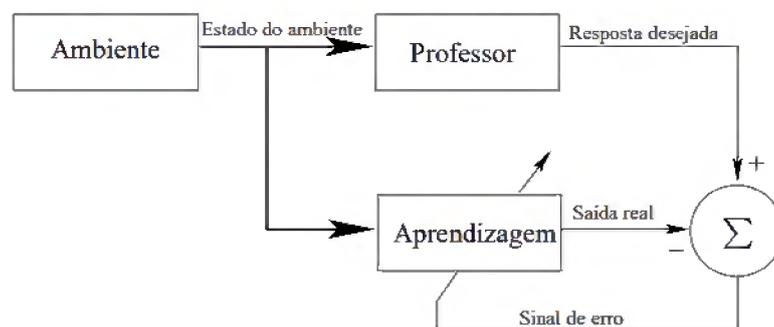
desvantagens que serão expostas a seguir. As RNA possuem a capacidade de aprender por exemplos, determinando a intensidade de conexões entre os neurônios pertencentes à rede. Logo, um conjunto de procedimentos definidos para ajustar os parâmetros de uma RNA, a fim que a mesma possa aprender uma determinada função, é chamado de algoritmo de aprendizado.

A designação de uma RNA, na resolução de um determinado problema, passa inicialmente por um processo de aprendizagem, onde a rede procura extrair informações relevantes de padrões de informação apresentados a ela, modelando uma representação própria (HAYKIN, 2001).

a) Supervisionado

Para Ferneda (2006), uma das propriedades mais importantes de uma RNA é a capacidade de aprender por intermédio de exemplos e fazer inferências sobre o que aprendeu, melhorando gradativamente o seu desempenho. O aprendizado supervisionado é o mais comum no treinamento das RNA's, tanto de neurônios com pesos como de neurônios sem pesos. É chamado de aprendizado supervisionado porque a entrada e saída desejadas para a rede são fornecidas por um supervisor chamado de professor, externo (Figura 4). O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos (BRAGA et al., 2000; CASTRO et al., 2013).

Figura 4: Mecanismo de aprendizado supervisionado com professor



Fonte: HAYKIN (2001)

A desvantagem do aprendizado supervisionado é que, na ausência do professor a rede não conseguirá aprender novas estratégias para situações não cobertas pelos exemplos do treinamento da rede. Os exemplos mais conhecidos de algoritmos para aprendizado supervisionado são a regra delta e a sua generalização para redes de múltiplas camadas, o algoritmo backpropagation (VALENÇA, 2005).

Nas RNA supervisionadas, os pesos terão os seus valores iniciais colocados aleatoriamente, e serão ajustados, por um algoritmo de aprendizagem pela rede na próxima iteração ou ciclo. O ajuste sináptico é dependente do valor esperado e do sinal atual de saída. Desta maneira, o método de aprendizado tenta minimizar o fluxo corrente de erros de todos os elementos em processamento, modificando continuamente os pesos até que a rede alcance uma certa precisão (HAYKIN, 2001).

Segundo Haykin (2001), com o aprendizado supervisionado, as RNA devem ser treinadas antes de serem usadas, o treinamento consiste da apresentação dos sinais de entrada e saída à rede. O treinamento é considerado completo quando a rede neural alcança um certo nível de performance. Este nível significa que a rede alcançou uma precisão estatística conforme as produções de saída necessárias para uma dada sequência de entradas e quando não há mais a necessidade de aprendizado, não há mais mudanças significativas dos pesos (COUTINHO, 2007).

Segundo Haykin (1999) o conjunto de treinamento precisa ser suficientemente grande para conter as informações necessárias, afim de que, a RNA aprenda os moldes e as relações importantes. Se a rede é treinada somente com um exemplo em um determinado tempo, todos os pesos serão colocados meticulosamente para este fato, os quais poderiam sofrer alterações drásticas no aprendizado de um próximo fato. Conforme Coutinho (2007) em um aprendizado, o sistema precisa aprender com todos os fatos em conjunto, provendo posteriormente o melhor ajuste dos pesos para todo o conjunto de fatos.

Segundo Braga et al., (2007), a rede é apresentada a diversos exemplos de um determinado problema, do qual extrai características durante o processo de aprendizagem, para posteriormente aplicar o conhecimento adquirido em outros dados do mesmo tipo de problema.

Haykin (2001) afirma que o pós treinamento supervisionado, a RNA precisa passar por uma análise sobre seu potencial em apresentar soluções com os dados que ainda não foram apresentados, pois se o resultado de saída do sistema não for razoável para este novo conjunto de dados, presume-se que o treinamento da rede ainda não foi suficiente.

Esta avaliação é crítica para assegurar que a rede simplesmente não memorizou um dado conjunto de dados, mas sim aprendeu os modelos/padrões gerais envolvidos na aplicação. É importante ressaltar que às vezes o problema da generalização é devido à má qualidade dos dados usados para o treinamento e não um problema da rede. (HAYKIN, 2001)

b) Não supervisionado

No aprendizado não supervisionado, a rede aprende sem a necessidade de um conjunto de treinamento. Ela não suporta influência externa para ajustar os seus pesos sinápticos, pois há um monitoramento de desempenho interno da mesma, analisando as regularidades e/ou tendências dos sinais de entrada, e conseqüentemente adaptando-se automaticamente às necessidades da rede (HAYKIN, 2001).

Segundo Haykin (2001), o aprendizado não supervisionado possui características de auto nivelção, sem um suposto auxílio para determinar se o aprendizado converge ou não para o caminho certo. A rede possui mecanismos, mais precisamente, informações, de como se organizar. Esta propriedade e percepção da rede devem-se à topologia e às regras de aprendizado adotados pela RNA.

Em uma RNA com o algoritmo de aprendizado não supervisionado deve ter enfatizado a cooperação entre as camadas de unidades de processamento. A competição entre estas unidades é a base de aprendizado da rede. Normalmente, quando a competição pelo aprendizado ocorre de fato, somente os pesos pertencentes à unidade de processamento vencedora são ajustados (HAYKIN, 2001).

2.2.4. Redes Multilayer Perceptron

Segundo Hornik et al., (1989) na arquitetura Redes Multilayer Perceptron existe múltiplas camadas de neurônios, consisti em conectar cada neurônio a todos os neurônios da camada seguinte. Cada neurônio é muito similar ao Perceptron proposto originalmente. Uma importante diferença é que a saída de cada um deles não é mais binaria, ela é dada pela aplicação de uma função de ativação não linear ao resultado da soma ponderada da entrada dos pesos da RNA.

Segundo Carvalho et al., (1998), o processo imposto a cada neurônio pertencente à RNA é resultante da união do processamento realizado pelos neurônios da camada anterior, que por sua vez estão atribuídos a este nodo da próxima camada. À medida com que cada camada intermediária da rede se aproxima da camada de saída, há uma delimitação do espaço de decisão dos dados que está recebendo.

Para uma RNA com duas camadas intermediárias, teríamos a primeira camada oculta, delimitando o espaço de padrões de treinamento através das “retas traçadas” pelos neurônios, a segunda camada forma regiões convexas, onde o número de lados que compõe tal

região é determinado pela quantidade de unidades conectadas a este neurônio, que por sua vez combina as retas que surgiram da camada anterior. Cada neurônio da camada de saída forma regiões, provenientes das combinações das regiões convexas (CARVALHO et al., 1998).

A RNA, Perceptron Multicamadas (MLP) utiliza o algoritmo de treinamento Backpropagation. Este algoritmo possui um procedimento sistemático de adaptação de massa baseado no método do gradiente descendente, de forma que, o erro quadrático da RNA torne-se mínimo (WIDROW & LEHR, 1990).

Considerando-se o neurônio que compõe uma rede Multilayer Perceptron de índice i da rede, e utilizando-se o método do gradiente descendente (WIDROW & LEHR, 1990), o ajuste de V é formulado pela Equação (3):

$$V_i = (k + 1) = V_i(k) + 2\gamma\beta_i X_i \quad \text{Eq. 03}$$

em que: γ - parâmetro de controle da estabilidade ou taxa de treinamento; X_i : vetor contendo as entradas do neurônio i ; V_i : vetor contendo os pesos do neurônio i ; $V_i = [w_{0i} \ w_{1i} \ w_{2i} \ \dots \ w_{ni}]^T$.

A Equação (4) descreve β_i quando o elemento i encontra-se na última camada, então

$$\beta_i = sgm'(s_i)\varepsilon_i \quad \text{Eq. 04}$$

no qual: $sgm'(s_i)$: derivada da função sigmóide com relação a s_i ; $i\varepsilon = d_i - y_i$; d_i : saída desejada do neurônio i . Se o elemento i encontrar-se nas demais camadas, β_i é determinado pela Equação (5):

$$\beta_i = sgm'(s_i)\varepsilon_i \sum_{k \in R(j)} w_{jk}\beta_k \quad \text{Eq. 05}$$

no qual: $sgm'(s_i)$ - derivada da função sigmóide com relação a s_i . $R(j)$ - conjunto dos índices dos elementos que se encontram na fileira seguinte à fileira do elemento i e que estão interligados ao elemento i .

2.2.5. Algoritmo de treinamento das redes MLP - Backpropagation

O algoritmo de redes do tipo multicamadas chamado de Retropropagação do Erro (Error Back-propagation), utiliza o paradigma do aprendizado supervisionado, por meio de pares de entrada e saída. Os pesos são ajustados buscando minimizar o erro, ocorrendo em duas fases, forward e backward (BULLINARIA, 2004). A grande contribuição desse algoritmo foi solucionar o problema da definição do erro dos neurônios das camadas intermediárias, possibilitando, com isso, a atualização dos pesos intermediários da rede (DIAMANTOPOULOU, 2006).

O Backward Propagation of Errors – Backpropagation é utilizado para treinamento de RNA de múltiplas camadas, usados em conjunto com algumas técnicas de otimização para cálculo de uma função de erro dos pesos da rede e estes são ajustados de forma a minimizar o erro da função.

De acordo com Haykin (2001), Valença (2011) e Pandorfi et al., (2011), o algoritmo Backpropagation para o treinamento das redes são organizados primeiramente com as escolhas do pesos, depois as polarizações e demais parâmetros. Atribuem-se os padrões de entrada do conjunto de treinamento definindo as entradas e saídas, para depois calcular o erro para os neurônios da camada de saída, subtraindo a saída desejada da saída calculada, depois ajustando-se os pesos da camada de saída. De forma simplificada pode se reunir as etapas do treinamento de uma rede usando Backpropagation da seguinte forma:

1 – Inicialização dos pesos e viés, definição dos valores da taxa de aprendizado, momento do critério de parada;

2 – Apresentação à rede das amostras de treinamento. Para cada amostra, é realizada a série de cálculos descrita nos passos 3 e 4;

3 – Propagação: supondo que cada amostra de treinamento seja representada por $(x(n), d(n))$, com vetor $x(n)$ sendo apresentado na entrada da rede e o vetor da resposta desejada $d(n)$ apresentada na camada de saída da rede, são calculadas as saídas locais induzidas e os sinais funcionais da rede prosseguindo camada por camada;

4 – Retropropagação: cálculo dos gradientes locais. Ajuste dos pesos sinápticos da rede na(s) camada (s) oculta (s) de acordo com a regra delta generalizada;

5 – Iteração: iterar as computações para frente e para trás de acordo com os passos 2 e 3 apresentando novas coletas de amostras de treinamento para rede até que o critério de parada seja atingido.

Cabe ressaltar que existem vários algoritmos de aprendizado para as redes MLP, e estes normalmente possuem um aprendizado supervisionado. Pode-se ainda classificar os mesmos em dois grupos: estáticos; dinâmicos (HAYKIN, 2001).

Os algoritmos de aprendizado estáticos não alteram a estrutura da rede, mudando somente o valor dos pesos sinápticos. Já os dinâmicos, podem mudar a arquitetura da rede, isto é, podem alterar o número de camadas, o número de neurônios da camada intermediária e o número de conexões da rede.

Segundo Patterson (1995), o algoritmo backpropagation pode ser aplicado a qualquer RNA que usufrui de uma função de ativação diferencial e aprendizado supervisionado, suas correções são realizadas nos pesos das conexões, na fase do treinamento, os sinais/padrões de entrada são apresentados a RNA em uma determinada ordem, cada padrão de treinamento é propagado adiante, camada após camada, até a produção do sinal/padrão de saída. A saída computada pela rede é então comparada com uma saída desejada, esta comparação irá gerar um valor que determinará o erro.

Segundo Carvalho et al., (1998) e Patterson (1995), o erro gerado pela comparação da saída com a desejada será utilizado como uma realimentação para as conexões, que resultará no ajuste dos pesos sinápticos de cada camada num sentido oposto à propagação dos sinais de treinamento, ajustando os pesos para as camadas ocultas de acordo com as camadas subsequentes. Deste modo, os erros computados na camada de saída serão usados para ajustar os pesos entre a última camada oculta com a camada de saída. Assim, o erro calculado de uma camada oculta será usado para ajustar os pesos da camada oculta anterior. Este processo será repetido até que a primeira camada oculta seja ajustada.

O treinamento do backpropagation é definido por duas fases para o algoritmo. Cada fase percorre um sentido da rede. A primeira fase, chamada de forward, define a saída da rede para um determinado padrão de dados de entrada. A segunda e última, está incumbida de utilizar a saída desejada/esperada e a saída fornecida pela última camada da rede para ajustar os pesos sinápticos da RNA (CARVALHO et al., 1998).

2.3.APLICAÇÃO DA RNA EM FLORESTA NATURAL

Na área de Mensuração e Manejo Florestal, a predição se dá em diversos níveis, de árvores individuais a povoamentos; e em diferentes características, imprescindíveis ao manejo, como volume, perfil do tronco, distribuição de altura e diâmetro (GORGENS et al., 2014). Ozçelik et al., (2010) asseguram que as RNA's oferecem uma série de vantagens, incluindo a

capacidade de detectar implicitamente relações não-lineares complexas entre variáveis de entrada e saída, o que é muito útil na modelagem do volume de árvores.

A aplicação da RNA em florestas justifica-se por diversas propriedades úteis e capacidades como, tolerância a falhas e ruídos, aprendizagem e generalização a partir de dados de treinamento, paralelismo maciço que as tornam muito eficientes, uniformidade de análise e projeto e analogia neurobiológica (BRAGA et al., 2000; HAYKIN, 2001; BULLINARIA, 2009). No início da década de 1990 Guan & Gerter (1991), utilizaram RNA para modelar a mortalidade regular de árvores individuais. As redes foram mais eficientes do que o emprego da regressão logística, sendo consideradas pelos autores como um procedimento que resolve os problemas da estimação de parâmetros e da seleção do modelo. Simultaneamente, em 1995 os mesmos autores publicaram um trabalho sobre o uso de RNA para modelar a probabilidade de sobrevivência das árvores individuais, obtendo resultados satisfatórios (GUAN & GERTER, 1991).

Soares et al., (2012) objetivaram especificamente a estimativa dos diâmetros das árvores utilizando uma rede neural Perceptron de Múltiplas Camadas a partir de apenas três medidas na base da árvore. Os autores concluíram que os valores dos diâmetros preditos pela RNA foram bastante satisfatórios em relação aos valores dos diâmetros reais. O mesmo ocorrendo nos trabalhos de Diamantopoulou, (2005); Diamantopoulou, (2006); Gorgens et al., (2009); Silva et al., (2009); Diamantopoulou & Milios, (2010).

Diamantopoulou (2005) comprovou a superioridade das RNA em relação aos algoritmos de regressão não linear para estudar o volume de casca em *Pinus sp* e concluiu que as RNA têm grande potencial para muitas aplicações em modelagem florestal, sendo útil como alternativa às tradicionais. Utilizando dados de cubagem de diferentes empresas florestais, Gorgens (2006), testou algumas formas de pré-processamento dos dados e arquiteturas de RNA para estimar volume de árvores de eucalipto (*Eucalyptus spp.*) e teca (*Tectona grandis L.f.*). Silva (2008) também obteve resultados satisfatórios, por meio de RNA, pois em seu estudo conseguiu estimar de maneira precisa os diâmetros sem casca e do cerne ao longo do fuste para árvores de teca.

Gorgens et al., (2009) empregaram as RNAs com o objetivo de construir uma rede que estimasse com eficiência o volume de árvores, concluíram que modelagem por RNA foi perfeitamente viável.

O uso da inteligência artificial por meio de RNA na pesquisa florestal ainda é recente, porém se apresenta como uma nova ferramenta em diversos ramos da Ciência Florestal apresentando excelentes resultados e sendo fundamental na tomada de decisões. Na pesquisa

florestal, o uso de RNA engloba diversas áreas, principalmente aquelas relacionadas à mensuração e manejo florestal, destacando-se os trabalhos de Silva (2008), Binoti (2010), Binoti (2012), Reis et al., (2016) e Cosenza et al., (2017).

3. MATERIAL E MÉTODOS

3.1.FONTE DOS DADOS DE ÁRVORES AMOSTRAS

Foram utilizados dados de 576 árvores amostra, coletados na UMF Água Azul II, situada no município de Breu Branco, Pará, cedidos para a presente pesquisa pela empresa Isabel Madeiras do Brasil Ltda – IBL.

Utilizou-se o método de Smalian para calcular o volume rigoroso das árvores amostra. Os dados foram coletados durante a operação de derruba na UMF Água Azul II. Mediram-se as circunferências em seções de dois metros ao longo do tronco, a circunferência a 1,30 metros da base do tronco (CAP) e o comprimento do fuste até a primeira bifurcação grossa (altura comercial). Todas as circunferências foram transformadas para diâmetro.

A seguinte expressão foi usada para calcular o volume de cada seção do tronco:

$$V_s = \frac{AS_1 + AS_2}{2} \cdot L \quad \text{Eq. 06}$$

Onde V_s = Volume da seção; L = comprimento da seção.

$$AS = \frac{\pi \cdot D^2}{4} \quad \text{Eq. 07}$$

A última seção ou seja, aquela da ponta do tronco junto a copa, comprimento menor que dois metros e foi usado a expressão acima.

O volume rigoroso do tronco (V) foi o obtido pela soma dos volumes individuais de cada seção do tronco.

3.2.DISTRIBUIÇÃO DE FREQUÊNCIA DAS ÁRVORES AMOSTRADAS POR CLASSE DE DIÂMETRO.

Tabela 1: Número de ocorrências de indivíduos por classes de diâmetro, das árvores amostra, medidas na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará

CLASSES DIAMÉTRICAS	N Arv	(%)
50-60	36	6,25%
60-70	131	22,74%
70-80	137	23,78%
80-90	119	20,66%
90-100	63	10,94%
100-110	46	7,99%
110-120	18	3,13%
>120	26	4,51%
TOTAL	576	100%

Quanto maior a amostragem, menor é o erro padrão e mais próximo da normalidade o conjunto de amostras se encontrará, independente de sua distribuição inicial, ou seja, quanto maior o número de amostras, mais próximo da realidade o conjunto de dados estará.

Para testar a normalidade foram usados os dados de 100 árvores amostra do conjunto de dado (Tabela 1 e 2) e aplicado o teste de Kolmogorov-Smirnov, segundo Nadal et. al. (2003).

Tabela 2: Estatística descritiva, número de indivíduos (N) por espécie utilizada na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará

Espécie	Num Ind	DAP (méd)(m)	Hc (médio) (m)	Total V (m³)
Amapá	2	79,15	18,75	13,81
Amapa amargoso	7	86,59	18,46	60,85
Amapa doce	6	74,20	20,03	36,37
Amarelão	1	83,00	19,30	7,45
Amesclão	7	77,37	19,64	45,74
Andiroba	6	71,78	17,93	28,84
Angelim amargoso	2	74,80	15,00	9,97
Angelim pedra	5	123,80	17,00	77,90
Angelim vermelho	25	104,42	17,37	312,44
Barrote	1	62,40	12,00	3,44
Breu sucububa	2	74,00	21,15	12,57
Casca-seca	2	62,15	16,65	8,27
Coco pau	1	69,00	21,20	7,35
Copaíba	10	77,33	20,71	76,15
Cumarú	2	82,75	24,30	18,26
Cupiúba	11	75,10	15,38	55,75
Currúpixa	8	104,05	16,29	85,52
Envira quiabo	12	68,68	18,08	58,21
Escorrega macaco	5	78,04	16,36	27,41
Fava atañã	4	76,88	21,41	31,13
Fava bolota	16	88,09	16,24	116,10

Faveira branca	4	69,93	18,45	20,78
Freijó	1	69,00	24,00	5,11
Guajará bolacha	6	69,42	17,79	29,26
Inharé	2	74,50	24,66	12,97
Jatobá	18	89,54	28,23	205,63
Louro abacate	2	61,25	17,55	8,63
Louro faia	3	83,07	22,86	30,00
Louro tamaquaré	22	74,35	18,73	134,01
Louro vermelho	13	89,05	19,02	98,72
Maçaranduba	137	81,88	20,70	1.033,49
Mamorana	1	80,30	23,00	8,56
Maparajuba	48	70,87	19,94	290,96
Marupá	1	70,30	26,80	5,91
Matá-matá branco	1	57,00	13,00	2,72
Melancieiro	38	86,26	19,06	264,20
Melancieiro Rosa	2	95,75	20,10	19,57
Mirindiba	2	78,50	19,75	14,17
Muiracatiara	12	79,32	27,78	107,27
Mururé	2	77,75	24,20	14,11
Não identificada	2	69,20	17,90	8,82
Oiticia	2	75,20	24,50	14,00
Orelha de macaco	4	77,75	13,84	19,30
Para-pará	2	61,80	21,75	8,76
Pau jacaré	2	58,10	16,48	7,30
Piquiá	10	102,38	16,22	93,15
Piquiarana	10	88,84	17,35	76,66
Roxinho	7	79,74	18,20	48,48
Sapucaia	2	100,65	17,65	21,31
Sucupira amarela	6	80,13	24,27	54,24
Sucupira babona	4	76,28	24,43	27,52
Sumaúma	6	82,37	17,99	45,67
Sumaúma rosa	4	69,80	20,30	24,81
Tachirana	2	89,10	21,10	19,68
Tanibuca	3	83,10	16,57	21,39
Tatajuba	10	104,88	22,61	133,61
Taurí	28	87,16	18,26	241,56
Timborana	16	74,43	17,14	89,72
Ucuúba preta	1	68,20	20,60	4,92
Uxirana	2	66,15	14,60	9,03
Virola	3	62,20	25,09	14,22
Total Geral	576	82,28	19,71	4.383,70

3.2.1. Estatística Descritiva

	<i>DAP (médio)(m)</i>	<i>Hc (médio) (m)</i>
Média	78,84	19,60
Erro padrão	1,63	0,45
Mediana	77,37	19,02
Amplitude	66,80	16,23
Mínimo	57,00	12,00
Máximo	123,80	28,23

3.3. EQUAÇÃO DE VOLUME COM FATOR DE FORMA MÉDIO DA UMF

O fator de forma médio para o volume do fuste comercial de cada árvore amostra foi determinado por meio da seguinte equação:

$$ff = \frac{V_{real}}{VC_{1,3h}} \quad \text{Eq. 08}$$

em que: ff = Fator de forma; V_{real} = Volume rigoroso da árvore em m^3 , determinado pelo processo de Smalian; $VC_{1,3h}$ = Volume do cilindro (diâmetro medindo a altura de 1,3 m do solo).

A estimativa dos volumes totais e comerciais usando o fator de forma é dada pela equação:

$$V_{ff} = \frac{\pi \cdot DAP^2}{4} \cdot H_c \cdot ff \quad \text{Eq. 09}$$

em que: V = volume estimado (m^3); DAP = Diâmetro a altura do Peito (diâmetro medindo a altura de 1,3 m do solo) e; ff = fator de forma médio da UMF; H_c = altura comercial.

3.4. EQUAÇÃO DE VOLUME ESTIMADO QUE UTILIZA O FATOR DE FORMA 0,7 CALCULADO POR HEINSDIJK & BASTOS (1963)

As equações que melhor se ajustaram à UMF no Estado do Pará, foram utilizadas também como parâmetro para comparação com o método que utiliza o volume do cilindro multiplicado por um fator de forma 0,7 (HEINSDIJK & BASTOS, 1963). Esta forma de cálculo, é descrita pela seguinte expressão:

$$V_{ff} = \pi \cdot \frac{DAP^2}{4} \cdot H_c \cdot 0,7 \quad \text{Eq. 10}$$

Onde, Dap é o diâmetro a altura do peito; H_c é a altura comercial em m e 0,7 é o fator de correção ou fator de forma proposto por Heinsdijk & Bastos (1963) utilizado na Amazônia.

3.5. MODELOS VOLUMÉTRICOS TESTADOS

Foram testados seis modelos de simples e dupla entrada, conforme tabela 3.

Tabela 3: Modelos testados para estimar o volume das árvores em pé coletados na Unidade de Manejo Florestal Água Azul II, situada no município de Breu Branco, Pará

Modelo	Modelo Aritmético	Autor
1	$V = \beta_0 + \beta_1 dap + \varepsilon$	Berkhout
2	$\log V = \log \beta_0 + \beta_1 \log dap + \varepsilon$	Husch
3	$\log V = \log \beta_0 + \beta_1 \log dap + \beta_2 Hc + \varepsilon$	Brenac
4	$\log V = \log \beta_0 + \beta_1 \log(dap^2 Hc) + \varepsilon$	Spurr modificado
5	$\log V = \log \beta_0 + \beta_1 \log dap + \beta_2 \log Hc + \varepsilon$	Schumacher e Hall modificado
6	$V = \beta_0 + \beta_1 dap^2 + \beta_2 d^2 h + \beta_3 Hc + \varepsilon$	Stoate

Em que: V = volume; dap = diâmetro à altura do peito; Hc = altura comercial; β_0 , β_1 , β_2 , β_3 = coeficientes; ε = componente aleatório (erro do modelo).

3.6. CRITÉRIOS USADOS PARA AVALIAÇÃO DA QUALIDADE DE AJUSTE DOS MODELOS REGRESSÃO

A seleção do melhor modelo de regressão foi baseada nos seguintes critérios: coeficiente de determinação ajustado em porcentagem ($R_{aj}^2\%$), Erro padrão da Estimativa ($S_{y,x}$), Erro padrão da Estimativa percentual ($S_{y,x}\%$), Erro padrão Estimado recalculado ($S_{y,x,r}$) utilizando o Índice de Furnival (IF), Erro padrão da Estimativa percentual recalculado utilizando o Índice de Furnival ($S_{y,x,r}\%$), Desvio Médio Percentual (DMP%), Raiz do Erro Quadrático Médio (RQME), Critério de Informação de Akaike (AIC) e valor ponderado dos escores estatísticos (VP), além da análise gráfica dos resíduos, para verificar a presença de tendenciosidade nas estimativas do volume (THOMAS et al., 2006; MOSER & OLIVEIRA, 2017).

a) R^2 ajustado (R^2 ajust.)

O coeficiente de determinação (R^2) informa a porcentagem da variação dos dados observados em torno da média que está sendo explicada pelo modelo ajustado. Os valores de R^2 foram transformados (SCOLFORO, 1993), para que fosse possível comparar os valores encontrados, já que os graus de liberdade da regressão diferem entre os modelos testados. A expressão matemática utilizada para sua obtenção é dada por

$$R_{Aj}^2 \% = 1 - \left[(1 - R^2) \frac{n - 1}{n - P - 1} \right] \quad \text{Eq. 11}$$

Onde: $R_{aj}^2 \%$. = Coeficiente de Determinação Ajustado; n = número de dados observados; P = número de coeficientes do modelo. Por esse critério, quanto mais próximo de um for o valor do Coeficiente da Determinação, melhor será o ajuste da linha de regressão.

b) Erro padrão da estimativa ($S_{y.x}$)

Erro padrão da estimativa indica a precisão do ajuste de um modelo matemático. Esse critério informa o quanto, relativamente, o modelo erra em média ao estimar a variável dependente (MACHADO et al., 2008).

$$S_{y.x} = \sqrt{QM_{res}} \quad \text{Eq. 12}$$

Onde $S_{y.x}$ = erro padrão da estimativa; QM_{res} = quadrado médio do resíduo, obtido na análise de variância.

c) Erro padrão da estimativa percentual ($S_{y.x} \%$)

$$S_{y.x} \% = \left(\frac{S_{y.x}}{\bar{Y}_{real}} \right) \cdot 100 \quad \text{Eq. 13}$$

Onde, $S_{y.x} \%$ erro padrão da estimativa percentual; \bar{Y}_{real} a média aritmética do volume real.

d) Erro padrão da estimativa percentual recalculado ($S_{y.xr} \%$)

$$S_{y.xr} \% = \frac{S_{yxr}}{\bar{Y}_{real}} \cdot 100 \quad \text{Eq. 15}$$

Onde, $S_{y.xr} \%$ = erro padrão da estimativa percentual recalculado; $S_{y.xr}$ = erro padrão da estimativa recalculado; \bar{Y}_{real} = média aritmética do volume real

e) Desvio Médio Percentual, (DMP%)

Indicador da qualidade de ajuste de um modelo de regressão, de acordo com (DRAPER & SMITH, 1966), expressa o quanto na média o modelo está subestimando ou superestimando a variável dependente de forma global. Quanto menor o valor do desvio médio percentual, melhor será o ajuste do modelo calculado por:

$$DMP_{\%} = \frac{\sum \left(\frac{O - P}{O} \right)}{n} \cdot 100 \quad \text{Eq. 16}$$

Onde, O = valores observados; P = valores estimados, n = número de observações, onde $DMP_{\%}$ negativo será superestimado e positivo subestimado.

f) Raiz do erro quadrático médio (RQME)

Raiz do quadrado médio do erro (JANSSEN & HEUBERGER, 1995).

$$RQME = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2} \quad \text{Eq. 17}$$

Onde, P_i são os valores estimados, O_i são os valores observados e n é o número de observações.

g) Critério de informação de Akaike (AIC)

O critério de informação de Akaike, está fundamentado na Teoria da informação, refletindo a perda de informação associada às predições de um modelo e os valores observados. Assim quanto menor o valor de AIC, mais explicativo é o modelo. Por si, o valor de AIC é abstrato e adimensional, porém aplicável na comparação de modelos. (MOSER & OLIVEIRA, 2017).

$$AIC = n \cdot \ln(SQ_{res}) - n \cdot \ln(n) + 2p \quad \text{Eq. 19}$$

Onde, AIC = critério de informação de Akaike; n = número de observações da amostra; p = número de parâmetros do modelo; SQ_{res} = Soma de quadrado de resíduos.

h) Valor ponderado (VP)

Esse critério de seleção leva em consideração todas as variáveis estatísticas anteriormente descritas, e por isso sintetiza os resultados, facilitando o processo de seleção do melhor modelo matemático. O valor ponderado (VP) dos escores dos parâmetros estatísticos, segundo Thomas et al., (2006), leva em consideração todas as variáveis estatísticas consideradas como critérios descritos anteriormente, sintetizando os resultados e facilitando o processo de seleção do melhor modelo matemático.

O Valor Ponderado foi determinado atribuindo-se valores ou pesos aos parâmetros estatísticos. As estatísticas foram ordenadas de acordo com a sua eficiência, sendo atribuído peso 1 para a equação mais eficiente, 2 para a segunda e assim sucessivamente (ranking), conforme metodologia descrita por Thiersch (1997). Após essa classificação individual, efetuou-se o somatório da pontuação para cada modelo matemático, sendo que a equação de volume que recebeu a menor soma foi recomendada como mais adequada para uso.

i) Análise Gráfica dos Resíduos

Ao final do processo de seleção dos modelos matemáticos, foi analisada a distribuição gráfica dos resíduos, com o objetivo de verificar se o modelo selecionado apresentava alguma tendência na estimativa da variável dependente o volume.

Azevedo et al., (2007) citam que os resíduos obtidos pelo processo de ajustamento são importantíssimos na metodologia de diagnóstico por observação. A grande vantagem do uso das técnicas gráficas é que elas, além de indicarem os problemas de heterogeneidade de variância, a não-linearidade dos modelos e a auto-correlação dos resíduos, indicam também a forma mais adequada de resolvê-los, como, por exemplo, a inclusão de novos componentes, o que pode melhorar, significativamente, um modelo. O resíduo padronizado é obtido como segue:

$$Res. Pad_i = \frac{\hat{e}_i}{S_{yx}} \quad \text{Eq. 22}$$

Onde, $S_{y.x}$ = erro padrão da estimativa; $\hat{\epsilon}_i = Y_i - \hat{Y}_i$ onde o resíduo é calculado pela diferença entre os valores observados (Y_i) e estimados (\hat{Y}_i).

3.7. TREINAMENTO REDES NEURAIAS ARTIFICIAIS

Na construção e treinamento das RNAs, utilizou-se o sistema Neuro 4, que permite construir RNA totalmente personalizáveis ou escolher entre várias arquiteturas de redes neurais pré-construídas, possibilitando obter estimativas com elevada precisão de características qualitativas e quantitativas de povoamentos florestais e processos industriais (BINOTI et al., 2015).

Para arquitetura da rede foi utilizada a Feedforward de camadas múltiplas, o Perceptron multicamadas. A funções de ativação utilizada foi a sigmoïdal para camada oculta e saída. Considerou-se o treinamento Supervisionado, empregando o algoritmo da retropropagação do erro (Backpropagation), como regra de aprendizagem por correção de erro. Aplicou-se o algoritmo de mínimo quadrado médio, para a regra de aprendizado, o qual tem por objetivo encontrar um conjunto de pesos e polarizações que reduzam a função erro (HAYKIN, 2001).

A rede neural Perceptron Multicamadas com algoritmo Backpropagation foi usada para realizar a previsão dos volumes de árvores em pé na floresta estimada. No processo de modelagem, foram treinadas 3000 RNA. Como variável de entrada foram combinadas ou não as variáveis dendrométricas (DAP; DAP²; 1/DAP; Hc; Log DAP; Log DAP²; Log Hc; ESPÉCIES) e para variável de saída (Volume) conforme Tabela 4.

Para compor o conjunto de dados de entrada da RNA foram utilizados os valores de diâmetro a altura do peito (DAP), altura comercial (Hc), espécies florestais (ESPÉCIES) e volume calculado por cubagem rigorosa (V) conforme item 3.1. A base de dados foram as 476 árvores amostra, medidas na UMF Água Azul II.

As redes treinadas foram divididas em três tipos: uma teve na camada de entrada um neurônio e como variável quantitativa, o “DAP”, a outra dois neurônios como variáveis quantitativas, “DAP e Hc”, e a última rede com três neurônios referente o “DAP, Hc e uma variável qualitativa ESPÉCIE”.

Tabela 4: Variáveis utilizadas no treinamento das redes neurais artificiais (RNA) na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará

RNA	Variáveis de Entrada (quantitativa e qualitativa)	Número de treinamentos	Saída
1-300	DAP; Hc; ESPECIE	300	Volume
1-300	LOGDAP; LOGHc; ESPECIE	300	Volume
1-300	DAP; ESPECIE	300	Volume
1-300	LOGDAP; LOGHc	300	Volume
1-300	DAP; Hc	300	Volume
1-300	LOGDAP ² ; LOGHc	300	Volume
1-300	DAP	300	Volume
1-300	DAP; 1/DAP; DAP ²	300	Volume
1-300	LOGDAP	300	Volume
Total		3000	

DAP: Diâmetro a altura do peito (cm); Hc: Altura comercial (m); espécies: espécies florestais inventariadas.

As arquiteturas escolhidas foram do tipo 1-8-1; 2-8-1; 3-8-1, variando a entrada com um, dois ou três neurônios, todas com oito camadas ocultas, e apenas uma saída. O intervalo do número de neurônios de acordo com o número de variáveis de entrada e saída é dado pela seguinte equação:

$$2 \cdot \sqrt{n} + n_2 \leq n_1 \leq 2 \cdot n + 1 \quad \text{Eq. 24}$$

Onde, n é o número de entradas da rede, n_1 é a quantidade de neurônios na camada escondida e n_2 é a quantidade de neurônios na camada de saída. O máximo definido pelo método foi para evitar memorização dos dados de entrada (overfitting) ou extrair informação insuficientemente no treinamento (underfitting). (HAYKIN, 2001).

Para cada treinamento foram retidas duas RNAs. Dessas escolheu-se as duas melhores por treinamento pela correlação entre os valores estimados e observados para serem comparadas entre si. Na escolha e comparação das RNAs a serem usadas na estimativa do volume foram avaliadas, no treinamento e na generalização, a estatística de correlação entre os valores estimados e reais ($Y_{\hat{Y}}$), e a raiz quadrada do erro quadrático médio (RQME) e o Índice de Akaike (AIC). Foi também realizada a análise gráfica da dispersão dos erros percentuais (Erro %) em relação aos valores observados:

$$Erro\% = \frac{(\hat{Y} - Y)}{Y} \cdot 100 \quad \text{Eq. 25}$$

Onde, *Erro%*: erros percentuais; Y : valores observados, \hat{Y} : valores estimados

a) Correlação

$$Y_{Y\wedge Y} = \frac{cov(Y, \hat{Y})}{\sqrt{S^2(Y)S^2(\hat{Y})}} \quad \text{Eq. 26}$$

Onde, $Y_{Y\wedge Y}$: correlação entre os valores estimados e reais; Y : valores observados, \hat{Y} : valores estimados, S^2 : variância, cov : covariância e n o número de observações

b) $RQME\%$: raiz quadrada do erro quadrático médio

$$RQME\% = 100 \cdot \bar{Y}^{-1} \sqrt{n^{-1} \sum_{i=1}^n (Y_i - \hat{Y})^2} \quad \text{Eq. 27}$$

Onde, $RQME\%$: raiz quadrada do erro quadrático médio; Y : valores observados, \hat{Y} : valores estimados, S^2 : variância, \bar{Y} : média aritmética dos valores observados e n o número de observações

c) Critério de informação de Akaike (AIC)

O critério de informação de Akaike, está fundamentado na Teoria da informação, refletindo a perda de informação associada às predições de um modelo e os valores observados. Assim quanto menor o valor de AIC, mais explicativo é o modelo. Por si, o valor de AIC é abstrato e dimensional, porém aplicável na comparação de modelos. (MOSER & OLIVEIRA, 2017).

$$AIC = n \cdot \ln(SQ_{res}) - n \cdot \ln(n) + 2p \quad \text{Eq. 19}$$

Onde, AIC = critério de informação de Akaike; n = número de observações da amostra; p = número de parâmetros do modelo; SQ_{res} = Soma de quadrado de resíduos.

Os volumes estimados pelas equações obtidas usando os diferentes métodos de ajuste e as arquiteturas ajustadas com as RNA foram comparados com os correspondentes volumes observados. Foram gerados gráficos do volume estimado em função do observado e elaborados visando identificar possíveis discrepâncias ou tendências nas estimativas de volume.

3.8.VALIDAÇÃO DAS EQUAÇÕES E RNAS

Das 576 árvores disponíveis no banco de dados 100 árvores de diferentes classes diamétricas, e que não fizeram parte dos ajustes dos modelos de regressão e equação utilizando o fator de forma médio, foram usados para validação do volume estimado pelas Equações e pelas RNAs. Aplicou-se o Teste-F (KLEIJNEN, 1987), Teste-t (CONOVER, 1980) e o teste Qui-Quadrado (SCHNEIDER et al. 2009) a 95% de probabilidade.

O teste não paramétrico de Kolmogorov-Smirnov foi utilizado para avaliação da normalidade dos resíduos, e a estatística de teste é dada por (NADAL et al., 2003; PEIXOTO, 2007):

$$d_{max} = \frac{y_i}{n} - Z_i \quad \text{Eq. 28}$$

Onde: n = tamanho da amostra, sendo $i=1,2,3,\dots,n$; Z_i = probabilidade acumulada da distribuição normal padronizada.

Aceita-se a hipótese (H_0) de que os resíduos se distribuem normalmente se $d_{max} < d_{crítico}$. O valor de crítico é obtido a partir de tabela, com nível de significância de 1% e $n=100$ (tamanho da amostra). O valor de Z_i , foi obtido usando a função estatística do BioEstat 5.0.

Onde:

Hipótese nula (H_0): é a hipótese assumida como verdadeira onde não difere estatisticamente dos demais.

Hipótese alternativa (H_1): é considerada quando a hipótese nula (H_0) não tem evidência estatística.

Para as RNAs foi separado 100 árvores diferentes do treinamento, para validação, adaptando a abordagem clássica de Haykin (1994) e Prechelt (1994). Para essa validação foi usada a arquitetura 2-8-1.

4. RESULTADOS E DISCUSSÃO

4.1. EQUAÇÃO DE VOLUME COM FATOR DE FORMA MÉDIO E EQUAÇÃO DE VOLUME ESTIMADO COM FATOR DE FORMA 0,7.

Forma utilizados 476 árvores amostra para estimar o volume de árvore em pé, utilizando o fator de forma médio calculado para este trabalho e o fator de forma 0,7 calculado do Heinsdijk & Bastos (1963) para Amazônia.

Tabela 5: Estatística descritiva da equação com fator de forma médio calculado para UMF Água Azul II e 0,7 calculado do Heinsdijk & Bastos (1963) para Amazônia

METODOLOGIA	TOTAL	MÉDIA	DESV. MÉDIO	VAR	ERRO PADRÃO	FATOR DE FORMA
Vff	3611,57	7,59	3,89	15,13	0,18	0,705249
V0,7	3584,69	7,53	3,86	14,9	0,18	0,7

4.2. MODELOS TESTADOS POR REGRESSÃO

O volume observado para as 476 árvores amostra cubadas na unidade de manejo no município de Breu Branco, PA, foi de 3178,242 m³ com um valor médio de 7,3441 m³ por árvore. Os diâmetros variaram de 50 a 153 cm.

Os coeficientes e as estatísticas dos modelos testados, estimados a partir das variáveis independentes DAP e Hc, são apresentados na Tabela 6.

Tabela 6: Coeficientes dos modelos testados com a variáveis DAP e Hc com seus respectivos, R²ajustado, Syx, Syx%, Syxr, Syxr%, DMP, pa estimativa do volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará

Mod	b0	b1	b2	b3	Raj ² %	Syx%	Syxr%	DMP%	RQME%	AIC	VP
1	-7,98619	0,18975			78,78	22,62		-3,54	1.343,58	638.272,72	20
2	-2,83006	1,92438			77,49		22,84	-2,14	1.370,42	490,50	19
3	-2,51494	1,71310	0,00107		77,47		23,28	-2,14	1.420,21	488,81	20
4	-3,76017	0,90163			90,08		16,29	-0,92	696,90	-495,44	10
5	-3,74325	1,91301	0,72634		91,09		15,28	-0,82	611,75	-576,41	7
6	-0,21894	0,00034	0,00003	0,02853	90,31	15,28		-1,84	610,89	289.516,93	12

Onde, b0, b1, b2, b3 = coeficientes; R_{aj}²% = Coeficiente de determinação ajustado; S_{y.x} = erro padrão da estimativa; S_{y.x}% = erro padrão da estimativa em percentual; S_{y.x.r} = erro padrão da estimativa recalculado; Syxr % = erro padrão da estimativa recalculado em porcentagem; Raiz do erro quadrático médio (RQME), AIC = Critério de informação Akaike; VP = Valor Ponderado

Os valores ponderados, que resumem os critérios usados na escolha dos melhores modelos, revelaram que os piores ajustes foram apresentados pelos modelos 1, 2 e 3 (Berkhout, Hush e Brenac). Os melhores ajustes segundo aquele indicador foram apresentados pelos modelos 4, 5 e 6 (Tabela 6).

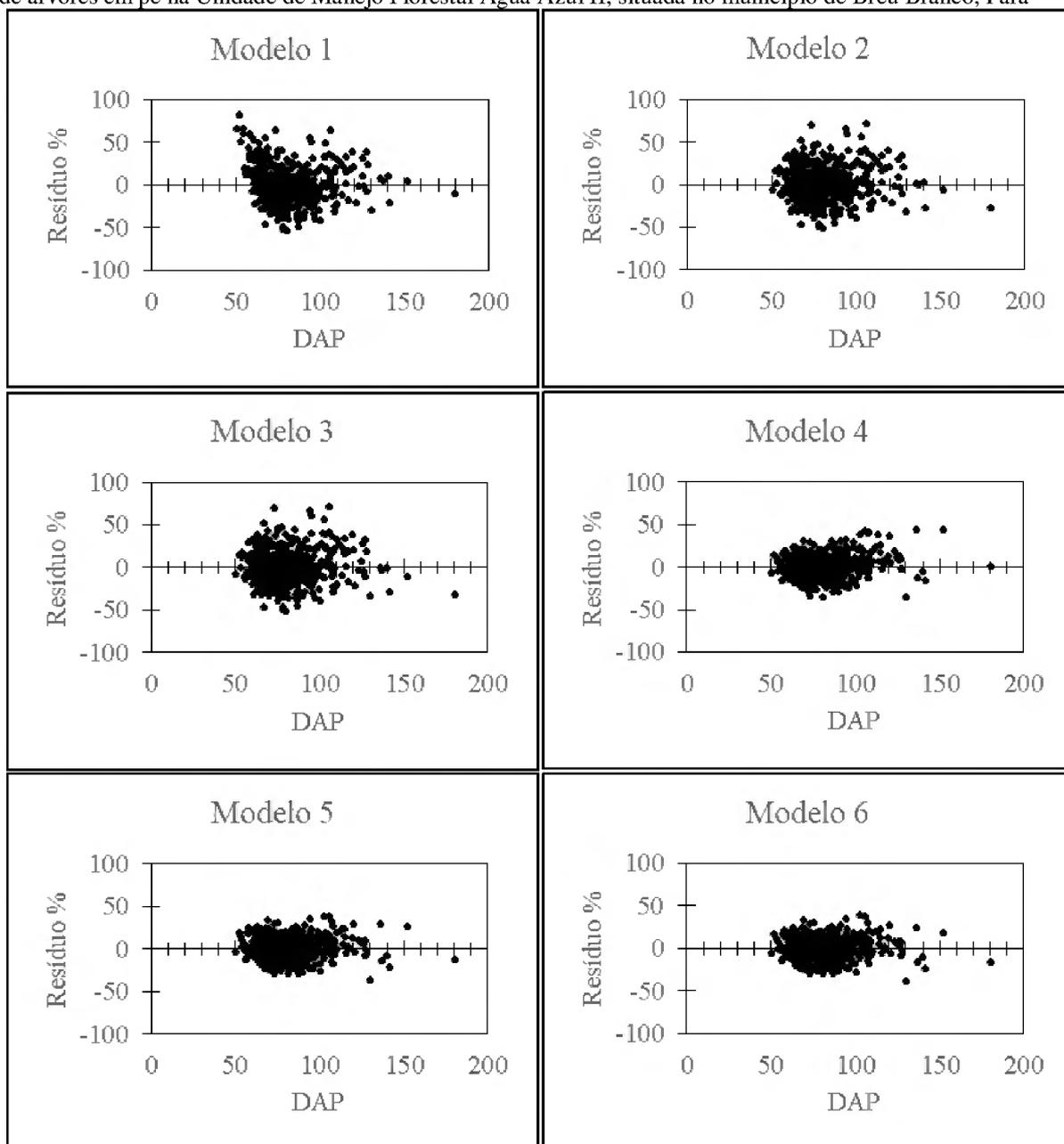
De todos os modelos testados, os melhores foram 4 e 5 os que apresentaram os menores erros de estimativa $S_{y.x}\%$ e $S_{y.x.r}\%$ (15,28%) (Tabela 6). O modelo 5, além disso apresentou o menor valor para DMP%(-0,82%) e maior $R_{aj}^2\% = 0,91$, porém é prudente verificar o desempenho do melhor modelo de regressão linear de simples e dupla entrada, com os valores de RQME, AIC e VP (Tabela 6) (MOSER & OLIVEIRA, 2017).

Segundo Moser & Oliveira (2017) quanto menor o valor de RQME e AIC mais explicativo será o modelo. O modelo 4 seria escolhido pela comparação simples do AIC e o modelo 6 pelo RQME. Então é necessária ranquear os modelos para que não haja nenhuma dúvida quanto à escolha do modelo a ser utilizado. Para essa escolha foi utilizado o score o VP (Tabela 6).

De acordo com o valor ponderado a posição estatística (VP), o modelo 5 é o mais satisfatório para estimativa da variável Volume Comercial (Tabela 6), pois apresentou o menor valor para esse critério. Este resultado está relacionado aos valores de coeficiente de determinação ajustado em porcentagem ($R_{aj}^2\%$), Erro padrão da Estimativa percentual ($S_{y.x}\%$), Erro padrão da Estimativa percentual recalculado utilizando o Índice de Furnival ($S_{y.x.r}\%$), Desvio Médio Percentual (DMP%), Raiz do Erro Quadrático Médio (RQME) e Critério de Informação de Akaike (AIC) e valor ponderado dos escores Estatístico (VP).

Por outro lado também é prudente verificar a distribuição gráfica dos resíduos percentuais. Segundo Draper & Smith (1981), mesmo que todos os indicadores de ajuste sejam bons, a análise gráfica de resíduos é quesito fundamental na escolha de uma equação de regressão aplicada à Ciências Florestais.

Figura 5: Distribuição de resíduos dos modelos utilizados para o ajuste da equação para determinação do volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará



Fonte: o autor

Na Figura 5 são mostradas as destruições dos resíduos dos seis modelos testados, de simples entrada (modelos 1,2 e 3), da variável combinada (modelo 4) e de dupla entrada (modelos 5 e 6).

A análise dos gráficos de dispersão dos resíduos é utilizada para detectar a independência dos resíduos, a presença de dados discrepantes e o comportamento da variância (MOSER et al., 2017). Essa análise permite verificar que todos os modelos tendem a superestimar os volumes das árvores com grandes diâmetros, a partir de aproximadamente 120 cm no caso dos modelos 2 e 3, e a partir de 150 cm nos casos restantes (Figura 5). Isso pode ser

explicado pelo número reduzido de árvores amostra com grandes diâmetros. Para melhorar as estimativas nessa situação, seria necessário coletar dados suficientes para superar esse problema. Melhor seria desenvolver uma equação específica para grandes diâmetros.

Os modelos 4, 5 e 6 apresentaram uma distribuição dos resíduos mais uniformes em torno de $\pm 25\%$ da linha de regressão. O modelo 5 por apresentar os melhores indicadores entre os modelos testados (tabela 6). Foi selecionado o modelo mais adequado para estimar o volume de madeira na UMF estudada.

4.3. NORMALIDADE DOS RESÍDUOS

O teste de aderência de Kolmogorov-Smirnov (K-S), a 1% de significância, dada pela equação (28). Resultou em aceitação da hipótese de normalidade (H_0), onde $d_{max} = 0,1393 < d_{crítico} = 0,1520$.

4.4. TREINAMENTO DE RNAs

Após o treinamento das 3000 RNAs para, estimar o volume individual das árvores, foram escolhidas 18 RNAs que apresentaram melhores desempenhos nas estimativas para cada conjunto de entrada (Tabela 7). RNAs apresentaram Coeficiente de Correlação variando de 0,8988 a 0,9662, conforme observa-se na Tabela 7.

As estimativas volumétricas nas RNAs apresentaram baixos valores de viés RQME e variância tornando os modelos confiáveis. Todas as RNAs testadas com diferentes entradas apresentaram correlação maiores maiores que 0,89, RQME baixo de 1,6 e viés próximo de zero (Tabela 8). A RNA sem variável qualitativa que apresentou maior correlação, menor Viés, RQME, Variância e AIC foram as RNA 30 e 204, com viés próximo a zero, RQME próximo a 1, e AIC igual 101,6261, demonstra grande capacidade de aprendizado das RNAs em apresentar uma resposta satisfatória para variável predita (HAYKIN, 2001; SILVA et al., 2010).

Tabela 7: Caracterização das Redes Neurais Artificiais (RNA) para a estimativa do volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará

RNA	TIPO	Arquitetura	Pesos	inputs	outputs	Ciclos	Correlação	Função de Ativação	
								Oculto	Saída
RNA 10	MLP	3-8-1	11	DAP; HC; ESPECIE	V	3000	0,9662	Sigmoidal	Sigmoidal
RNA 39	MLP	3-8-1	11	DAP; HC; ESPECIE	V	3000	0,9662	Sigmoidal	Sigmoidal
RNA 227	MLP	3-8-1	11	LOGDAP; LOGHC; ESPECIE	V	3000	0,9652	Sigmoidal	Sigmoidal
RNA 37	MLP	3-8-1	11	LOGDAP; LOGHC; ESPECIE	V	3000	0,9652	Sigmoidal	Sigmoidal
RNA 156	MLP	2-8-1	10	DAP; ESPECIE	V	3000	0,9395	Sigmoidal	Sigmoidal
RNA 29	MLP	2-8-1	10	DAP; ESPECIE	V	3000	0,9395	Sigmoidal	Sigmoidal
RNA 30	MLP	2-8-1	10	LOGDAP; LOGHC	V	3000	0,9540	Sigmoidal	Sigmoidal
RNA 204	MLP	2-8-1	10	LOGDAP; LOGHC	V	3000	0,9540	Sigmoidal	Sigmoidal
RNA 12	MLP	2-8-1	10	DAP; HC	V	3000	0,9537	Sigmoidal	Sigmoidal
RNA 21	MLP	2-8-1	10	DAP; HC	V	3000	0,9537	Sigmoidal	Sigmoidal
RNA 51	MLP	2-8-1	10	LOGDAP ² ; H	V	3000	0,9440	Sigmoidal	Sigmoidal
RNA 63	MLP	2-8-1	10	LOGDAP ² ; H	V	3000	0,9440	Sigmoidal	Sigmoidal
RNA 211	MLP	1-8-1	9	DAP	V	3000	0,8996	Sigmoidal	Sigmoidal
RNA 10	MLP	1-8-1	9	DAP	V	3000	0,8996	Sigmoidal	Sigmoidal
RNA 209	MLP	2-8-1	10	DAP; 1/DAP; DAP ² ; LOGDAP	V	3000	0,8996	Sigmoidal	Sigmoidal
RNA 228	MLP	2-8-1	10	DAP; 1/DAP; DAP ² ; LOGDAP	V	3000	0,8996	Sigmoidal	Sigmoidal
RNA 246	MLP	1-8-1	11	LOGDAP	V	3000	0,8988	Sigmoidal	Sigmoidal
RNA 8	MLP	1-8-1	11	LOGDAP	V	3000	0,8988	Sigmoidal	Sigmoidal

Tipo de RNA: MLP = Multilayer Perceptron, DAP = Diâmetro a altura do peito, Hc = altura comercial, espécies, V = Volume.

As RNA que apresentaram os piores resultado de correlação foram as de número 211, 10, 228, 209, 246 e 8, porém apresentaram Viés, RQME, variância, AIC, com bom ajuste métrico. Portanto, na área florestal é altamente aplicável a análise do gráfico de dispersão dos valores observados os valores preditos para verificação da qualidade dos valores preditos pela RNA (MOSER & OLIVEIRA, 2017).

Tabela 8: Precisão das redes neurais artificiais selecionadas para volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará

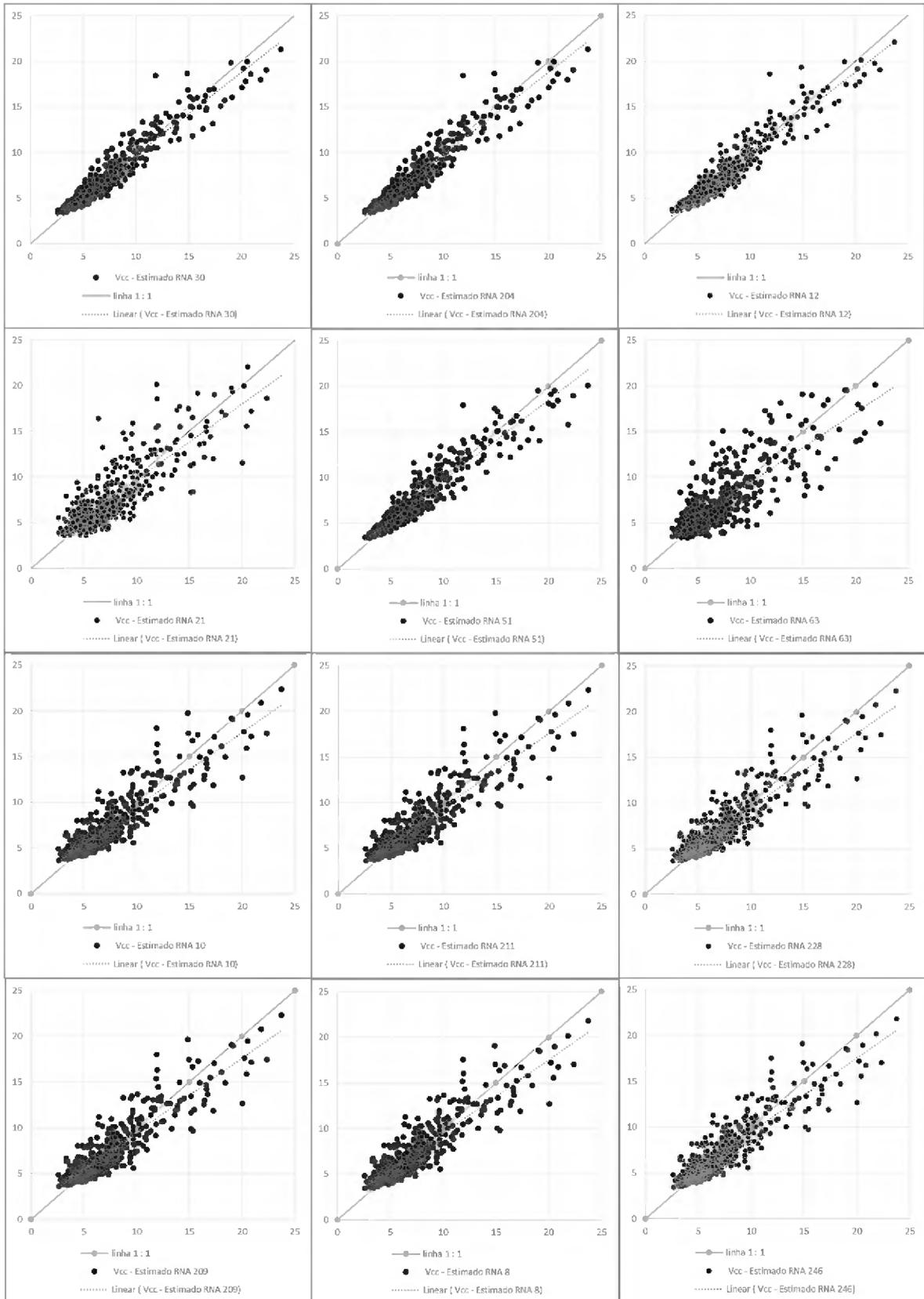
RNA	Bias	RQME	Variância	SQR	Correlação	AIC	VARIÁVEL
RNA 156	-0,0023	1,2471	1,5584	740,2654	0,9395	230,1972	D; ESPECIE
RNA 29	-0,0023	1,2471	1,5584	740,2654	0,9395	230,1972	D; ESPECIE
RNA 227	-0,0131	0,9490	0,9023	428,6584	0,9652	- 27,8646	LOGD; LOGHC; ESPECIE
RNA 37	-0,0131	0,9490	0,9023	428,6584	0,9652	- 27,8646	LOGD; LOGHC; ESPECIE
RNA 39	-0,0004	0,9381	0,8819	418,8860	0,9662	- 38,8418	D; HC; ESPECIE
RNA 10	-0,0004	0,9381	0,8819	418,8860	0,9662	- 38,8418	D; HC; ESPECIE
RNA 30	-0,0045	1,0895	1,1895	565,0426	0,9540	101,6261	LOGD; LOGHC
RNA 204	-0,0045	1,0895	1,1895	565,0426	0,9540	101,6261	LOGD; LOGHC
RNA 21	0,0024	1,0955	1,2026	571,2296	0,9537	106,8097	D; HC
RNA 12	0,0024	1,0955	1,2026	571,2296	0,9537	106,8097	D; HC
RNA 63	-0,0082	1,1941	1,4288	678,7118	0,9440	188,8747	LOGD2; H
RNA 51	-0,0082	1,1941	1,4288	678,7118	0,9440	188,8747	LOGD2; H
RNA 211	0,0051	1,5934	2,5443	1.208,5349	0,8996	459,5112	D
RNA 10	0,0051	1,5934	2,5443	1.208,5349	0,8996	459,5112	D
RNA 228	0,0018	1,5935	2,5445	1.208,6251	0,8996	465,5467	D; 1/D; D ² ; LOGD
RNA 209	0,0018	1,5935	2,5445	1.208,6251	0,8996	465,5467	D; 1/D; D ² ; LOGD
RNA 246	-0,0118	1,5980	2,5589	1.215,5541	0,8988	462,2678	LOGD
RNA 8	-0,0118	1,5980	2,5589	1.215,5541	0,8988	462,2678	LOGD

Bias= Viés, RQME=Raiz quadrada do erro médio, SQR = somatória do quadrados do resíduo, AIC = Índice de Akafke

As RNAs 30 e 204, se apresentaram superiores ao melhor modelo volumétrico em relação a todas as estatísticas de qualidade de ajuste (Tabela 7 e 8), o que confirma com outros trabalhos de que as RNA apresentam uma alternativa promissora na substituição de modelos tradicionais de regressão para modelagem florestal (BINOTI, 2012; CAMPOS & LEITE, 2013).

As RNAs com uma variável de entrada DAP(211, 10, 228, 209, 246 e 8) apresentaram maior tendência de subestimação em todas as classes, vivivelmente comprovada pela diferença da reta 1:1, possivelmente por não ter a influência da variável Hc (Figura 6).

Figura 6: Gráfico de Valores observados por cubagem rigorosa e estimados por treinamento de RNA para volume de árvores em pé na Unidade de Manejo Florestal Agua Azul II, situada no município de Breu Branco, Pará



Fonte: o autor

Os gráficos de dispersão do volume observado pelo predito, verificou-se as melhores disposições em relação a reta identidade foram as RNAs 30, 204 e 12, essa melhora se dá pela inclusão da variável HC, com ligeira subestimação entre as classes 15 e 20.

A RNA mais simples (98) que utilizou somente a variável D sem combinação e nem logaritimização, gerou também resultados bem próximos aos da RNA 4 que utilizou as variáveis D e Hc (Tabela 4), com uma diferença na correlação de 0,054 (Tabela 7) e no RQME de 1,537% (Tabela 8). Porém, apresentou maior tendenciosidade de subestimação em todas as classes (Figura 7), possivelmente por não ter a influência da variável Hc. Apesar disso foi a rede mais simples o que garante maior praticidade no seu uso.

Todos os modelos apresentaram estimativas estatísticas próximas de correlação, RQME e Viés (Tabela 7 e 8), Estes resultados estão de acordo com Binotti (2012) que demonstram a elevada capacidade das RNAs em estimar o volume de árvore da floresta em função do DAP e altura comercial, mesmo sem a inclusão da variável ESPÉCIE.

De acordo com as métricas de qualidade de ajuste, a RNA 30 apresentou melhor desempenho, quando comparado ao modelo 5 (Schumacher e Hall modificado), o que corrobora com Ribeiro et al., (2015), que em seu estudo concluiu que, de maneira geral as RNAs apresentam melhores resultados que equações de regressão para estimativa de volume.

4.5.VALIDÇÃO DA EQUAÇÃO COM FATOR DE FORMA MÉDIO, EQUAÇÃO COM FATOR DE FORMA 0,7 USADO POR HEINSDIJK & BASTOS, MODELO SHUMAKER E HALL E RNA

Os resultados do teste F e Teste t ($p > 0,05$), levaram a aceitar a hipótese nula de que não existe diferença significativa entre os volumes observados (V) e os estudados pelas equações e a RNA. Corroborando com estudos em florestas naturais realizadas por Binotti (2012) e Campos & Leite (2013) (tabela 9).

Tabela 9: Resultados estatístico e validação das estimativas do volumes de árvores em pé na UMF Água Azul II, por meio dos teste F, teste t e χ^2 a 95% de probabilidade

	gl	Média	Desvio Padrão	Erro Padrão	Tete F p-valor	Tete t p-valor	χ^2 calc	χ^2 Tab
V		4,656						76,594
V ff	99	8,711	5,387	0,539	0,682	- 0,472	13,128	
V0,7	99	8,620	5,330	0,533	0,772	- 0,327	12,093	
Equação SH	99	0,857	0,251	0,025	0,977	0,149	31,070	
RNA 4	99	8,637	4,580	0,458	0,738	- 0,335	14,986	

V = Volume com casca cubados rigorosamente pela metodologia de smalian, Vff= Volume utilizando o fator de forma médio, equação SH = Equação de volume gerada pelo método Shumacker e Hall modificada, RNA = rede neural artificial treinada

A validação pelos valores do teste F, teste t, deverão ser menor que 0,05 ($p>0,05$), levando-se a aceitar a hipótese nula de que não existe diferença significativa entre os modelos para estimativas volumétricas, corrobora com Lacerda et al., (2017), em seu estudo na Savana amazônica também obteve resultados semelhantes.

A validação dos modelos escolhidos, a partir do teste Qui-quadrado também mostra que não existe diferença estatisticamente significativa entre os valores estimados e os reais. Nesse estudo, todos os valores tabelado com graus de liberdade 99 e um nível de confiança de 95% foram inferiores ao tabelado: 13,128 para Volume estimado por fator de forma médio, 12,093 para Volume Estimado por fator de forma 0,7, 31,070 para Volume estimado por regressão linear e 14,986 para volume estimado por treinamento de rede neurais artificiais.

4.6.ESCOLHA DOS MÉTODOS PARA ESTIMATIVA DO VOLUME

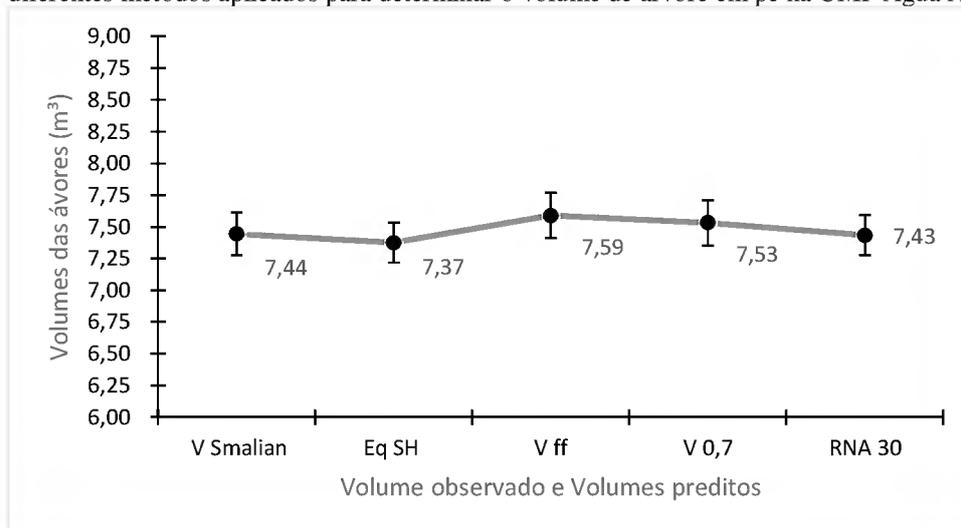
Para essa escolha foram comparados os valores estimados pelas equações e pela RNA com volumes rigorosos calculados pelo método de Smalian (Tabela 10). Foi verificado também que os dados são normais de acordo com Teste de Kolmogorov-Smirnov ($p>0,05$).

Tabela 10: Comparação entre o Volume e as predições de volume, por fator de forma médio, regressão linear e treinamento de RNA de árvores em pé na Unidade de Manejo Florestal Água Azul II, situada no município de Breu Branco, Pará, obtidas por diferentes métodos.

METODOLOGIA	Total	Médio	Desvio Médio	Variância	Erro Padrão	Teste F (valor-P)	Teste t (valor-P)
Volume Rig. Smalian	3.543,07	7,44	3,66	13,36	0,17		
Equação SH	3.510,01	7,37	3,43	11,78	0,16	0,76	0,28
V ff	3.611,57	7,59	3,89	15,13	0,18	0,56	0,56
V 0,7	3.584,69	7,53	3,86	14,90	0,18	0,72	0,31
V RNA 30	3.538,51	7,43	3,46	11,96	0,16	0,97	0,07

Conforme se observa na Tabela 10 não houve diferença significativa entre as médias dos volumes observados e estimados pelos métodos aplicados a uma probabilidade de 95% (Teste F e Teste t). Neste caso qualquer dos métodos poderia ser utilizados para estimar o volume de árvores em pé na UMF Água Azul II. No entanto, segundo Chaves et al., (2014) e Binoti (2012), é importante considera a análise dos gráficos de dispersão entre os valores observados e preditos, para detectar tendências de superestimação e de subestimação (Figura 7 e 8).

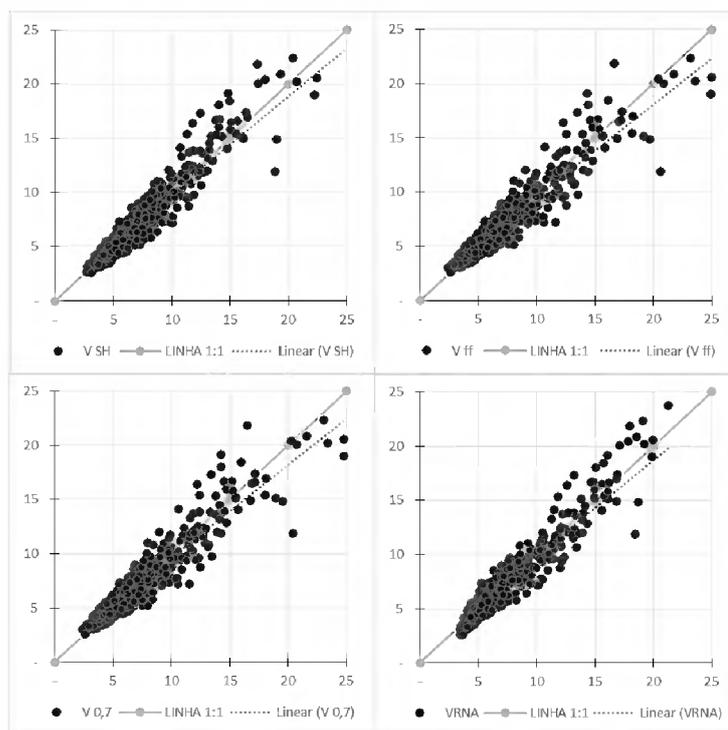
Figura 7: Análise gráfica da comparação do erro padrão e média dos volumes observados e estimados pelos diferentes métodos aplicados para determinar o volume de árvore em pé na UMF Água Azul II



Fonte: o autor

Observou-se na Figura 7 a pequena diferença entre os volumes medidos pelo método de Smalian e os estimados pelas diversas técnicas do estudo, o que torna perceptivo os resultados da tabela 10, onde não houve diferença significativa entre as medias dos volumes observados e preditos.

Figura 8: Análise gráfica do ajuste do Volume com casca cubados rigorosamente com Modelo de regressão Shumacher e Hall modificado, Volume do cilindro utilizando o fator de forma médio, Volume do cilindro utilizando o fator de forma 0,7 e com Volume estimado pela melhor Rede Neurais Artificiais



Fonte: o autor

Segundo Miguel et al., (2010) relatam que, para a obtenção de estimativas rápidas da variável volume, o fator de forma é uma ferramenta de grande valia, porém a sua precisão é inferior às equações de volume com uso de regressão linear, fazendo-se necessário ponderações e cuidados quanto ao seu uso.

Na figura 8, observar-se que a dispersão do gráfico do volume estimado com uso do fator de forma médio está ligeiramente mais distante que os demais da reta 1:1, verificou-se também uma superestimação do volume estimado pela RNA nos volumes entre 10 a 13 m³.

Verificou-se também que os Modelo 5 e o Modelo com uso de fator de forma médio não estimaram muito bem árvores muito grossas com volume entorno de 25 m³. E que os volumes por regressão linear subestima nas classes 12 a 19 m³.

De acordo com os parâmetros estatístico da tabela 7, e a análise dos gráficos de dispersão (observado vs. predito) e com a validação dos dados, onde verificou-se que houve significância a 95% (teste-F), para o treinamento de RNA 30, comprova a superioridade da metodologia de RNA em relação aos demais métodos de predição, o mesmo resultado obteve Pequeno et al., (2016) com relação a predição de diâmetro.

A RNA mostrou-se eficiente em relação a regressão linear, em todos os parâmetros estatístico e na análise dos gráfico de dispersão, o que corrobora com Lacerda et al., (2017), em seu estudo na Savana amazônica também obteve os mesmos resultados. Ribeiro et al., (2015) A aplicação de RNA é mais eficiente para estimar volume comercial de fuste na Floresta Nacional do Tapajós do que a equação do modelo único para múltiplos volumes.

5. CONCLUSÕES

As metodologias para estimativa de volume com casca proposta neste estudo não obtiveram diferença significativa a uma probabilidade de 95% entre elas.

Todas as RNAs com adição da variável qualitativa de entrada ‘ESPÉCIE’, obteve maior precisão estatística que a melhor RNA (30). Porém não foi comparada com os demais métodos por não possuir variável qualitativa nas suas fontes de dados. Os volumes estimados pelas RNA com duas entradas (DAP, Hc) tiveram melhores desempenhos comparado com uma única entrada (DAP).

A adição da variável qualitativa de entrada ‘ESPÉCIE’, resultou em maior precisão estatística que a melhor RNA (30) sem aquela variável na predição de volume.

As RNAs com uma variável DAP de entrada também apresentaram métricas estatísticas satisfatórias, podendo ser utilizadas para estimativa de volume de árvores em pé de uma floresta natural

REFERENCIAS

AKKOÇ, S. An empirical comparison of conventional techniques, neural networks and the three stage hybrid Adaptive Neuro Fuzzy Inference System (ANFIS) model for credit scoring analysis: The case of Turkish credit card data. **European Journal of Operational Research**, v. 222, n.1, p.168–178, 2012.

ALEKSANDER, I.; MORTON, H. **An introduction to neural computing**. London: Chapman & Hall, 1990.

AZEVEDO, C. P.; SANQUETTA, C. R.; SILVA, J. N. M.; MACHADO, S. A.; SOUZA, C. R.; OLIVEIRA, M. M. Predição da distribuição diamétrica de uma floresta manejada experimentalmente através de um Sistema de Equações Diferenciais. **Acta Amazônica**, v. 37, n. 4, p.521-532, 2007.

AZEVEDO, G. B.; SOUSA, G. T. O.; BARRETO, P. A. B.; CONCEIÇÃO JÚNIOR, V. Estimativas volumétricas em povoamentos de eucalipto sob regime de alto fuste e talhadia no sudoeste da Bahia. **Pesquisa Florestal Brasileira**, Colombo, v. 31, n. 68, p. 309-318, 2011.

BAIMA, A.V.; SILVA, S.M.S.; SILVA, J.N.M. A silvicultura tropical na Amazônia: Equações de volume para uma floresta tropical de terra-firme Moju-Pará. Belém: **EMBRAPA** Amazônia Oriental. DFID, 2001. p. 367-392.

BARONE, D. A. C. Projeto Revox. Disponível em: <http://www.ucs.tche.br/revox>, 1999. Acesso em: 22 dez. 2017.

BARRETO, W. F., LEÃO, F. M., de MENEZES, M. C. SOUZA, D. V. (2014). Equação de volume para apoio ao manejo comunitário de empreendimento florestal em Anapu, Pará. **Pesquisa Florestal Brasileira**, v. 34, n. 80, p. 321-329, 2014.

BARROS, P. L. C.; SILVA JÚNIOR, A. T. Equação de volume para árvores de uma floresta tropical densa no município de Anapu, oeste do estado do Pará, Amazônia oriental. **Revista Ciências Agrárias**, Belém, n.51, p.115-126, 2009.

BINOTI, D. H. B.; LEITE, H. G.; LOPES, P. F.; FERREIRA, E. Redes Neurais Artificiais – A realidade da silvicultura brasileira. **Revista Eletrônica Campo & Negócio**, 2015.

BINOTI, M. L. M. S. 2010. **Redes Neurais Artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. Dissertação de Mestrado em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, Minas Gerais. 54p.

BINOTI, M. L. M. S. **Emprego de redes neurais artificiais em mensuração e manejo florestal**. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, MG, 2012.

BRAGA, A. P; CARVALHO, A. P. L. F; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2.ed. Rio de Janeiro: LTC, 2007.

BRAGA, A.; LUDERMIR, T. B.; CARVALHO, A. C. P. L. F. **Redes Neurais Artificiais teoria e aplicações**, Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A., 2000.

BULLINARIA, J. A. **Committee machines**. Lecture notes in introduction of neural networks, 2004.

BULLINARIA, J. A. **Introduction to neural computation**, 2009.

BULLINARIA, J. A. **Introduction to Neural Computation**. 2014.

CABACINHA, C. D.; CASTRO, S. S. Relationships between floristic diversity and vegetation indices, forest structure and landscape metrics of fragments in Brazilian Cerrado. **Forest Ecology and Management**, Amsterdam, v. 257, p. 2157-2165, 2009.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3.ed. Viçosa, MG: Universidade Federal de Viçosa, 2009. 543p.

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

CARVALHO, André, LUDEMIR, Antônio. **Fundamentos de Redes Neurais Artificiais: 11ª Escola de Computação**. Imprinta Gráfica e Editora Ltda, 1998.

CASTELLANOS, A.; BLANCO, A. M.; PALENCIA, V. Applications of radial basis neural networks for area forest. **International Journal Information Theories and Applications**, v.14, p.218-222, 2007.

CASTRO, R. V. O.; SOARES C. P. B.; LEITE, H. G.; SOUZA, A. L.; NOGUEIRA, G. S.; MARTINS, F. B. Individual Growth Model for Eucalyptus Stands in Brazil Using Artificial Neural Network. **ISRN Forestry**, v. 2013, p. 112, 2013.

CHAVES, J., RÉJOU-MÉCHAIN, M., BÚRQUEZ, A., CHIDUMAYO, E., COLGAN, M. S., DELITTI, W. B., DUQUE, A., EID, T., FEARNESIDE, P. M., GOODMAN, R. C., et al. Improved allometric moles to estimate the aboveground biomass of tropical trees. **Global change biology**, v. 20(10), p.3177-3190, 2014.

COSENZA, D. N.; SOARES, A. A. V.; ALCÂNTARA, A. E. M.; SILVA, A. A. L.; RODE, R.; SOARES, V. P.; LEITE, H. G. Site classification for eucalypt stands using artificial neural networks based on environmental and management features. **Cerne**. v. 23, p. 310–320, 2017.

CONOVER, W. J. **Practical Nonparametric Statistics**. Wiley, New York, NY. 1980.

COUTINHO, G. **Reconhecimento facial baseado em Support Vector Machines**. Dissertação. UFLA, 2007.

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

DIAMANTOPOULOU, M. J.; MILIOS, E. Modelling total volume of dominant pine trees in reforestations via multivariate analysis and artificial neural network models. **Biosystems engineering**, v. 105, n. 3, p. 306–315, 2010.

DIAMANTOPOULOU, M. J. Tree-Bole Volume Estimation on Standing Pine Trees Using Cascade Correlation Artificial Neural Network Models. **Agricultural Engineering International: The CIGR EJournal**, VIII, 2006.

DRAPER, N.R. SMITH, H., 1966. **Applied Regression Analysis**. Wiley, New York, NY.

DRAPER, N.R.; SMITH, H. Applied regression analysis. New York: John Wiley & Sons. 1981. 407p.

FERNANDES, N. P.; JARDIM, F. C. S.; HIGUCHI, N. Tabelas de volume para a floresta de terra firme da Estação Experimental de Silvicultura Tropical. **Acta Amazonica**, n.13, v. 3-4, p. 537-545, 1983.

FERNEDA, E. Redes neurais e sua aplicação em sistemas de recuperação de informação. **Ciência da Informação**, v. 35, n. 1, p. 25–30. 2006.

FIGUEIREDO FILHO, A.; MACHADO, S. A.; MIRANDA, R. O. V. de; RETSLAFF, F. A. de S. **Compêndio de equações de volume e de afilamento de espécies florestais plantadas e nativas para as regiões fitogeográficas do Brasil**. Curitiba. 2014. 306 p.

GORGENS, E. B.; LEITE, H. G; SANTOS, H. N.; GLERIANI, J. M.; SOARES, C. P. B.; CEOLIN, A. Influência da arquitetura na estimativa de volume de árvores individuais por meio de redes neurais artificiais. **Revista Árvore**. Viçosa - MG, v. 38, n. 2, p. 289-295, 2014.

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

GORGENS, E. **Estimación do volume de árvores utilizando redes neurais artificiais**. 84f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, MG, 2006.

GRAUPE, D. **Principles, of artificial neural networks. S.I:** World Scientific, 2013.

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

GUYTON, A. C. & HALL, J. E. **Fisiologia médica**. 11^a Ed. Rio de Janeiro. Elsevier. 2006, 366p.

HAYKIN, S. **Neural Networks and Learning Machines**, Prentice Hall, Third Edition, 2008.

HAYKIN, Simon. **Neural Networks: a comprehensive foundation**. New York: Macmillan College Publish Company, 1994.

HAYKIN, S. **Neural Networks: a comprehensive foundation**. 2nd. Ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 898 p.

HIGUCHI, N.; RAMM, W. Developing bole wood volume equations for a group of tree species of central amazon (Brazil). **Commonw. For. Rev.** p. 33-41, 1985.

HIGUCHI, N.; SANTOS, J.; JARDIM, F.C.S. Tamanho de parcela amostral para inventários florestais. **Acta Amazonica**. p. 91-103, 1982.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural networks**, S.I., v2, n. 5, p. 359-366, 1989.

HUSCH, B. **Planning a Forest Inventory**. Forest Products, Rome, 1971. 120 p.

IBGE. Sistema IBGE de Recuperação Automática - SIDRA. Produção da Extração Vegetal e da Silvicultura, Disponível em:<<http://www.sidra.ibge.gov.br/bda/pesquisas/pevs/default.asp>>. Acesso 22 dez. 2017.

INDUSTRIA BRASILEIRA DE ÁRVORES (IBÁ). Relatório IBÁ 2016. 2016. Disponível em: http://iba.org/images/shared/Biblioteca/IBA_RelatorioAnual2016_.pdf. Acesso em: 17 dez.2017.

ITTO. Biennial review and assessment of the world timber situation 2013-2014. Yokohama. 2015. Disponível em http://www.itto.int/annual_review/. Acesso em 22 dez. 2017.

JANSSEN, P. H. M.; HEUBERGER, P. S. C. Calibration of process-oriented models. **Ecological Modelling**. p. 55-66, 1995.

KLEIJNEN, J. P. C. **Statistical Tools for Simulation Practitioners**. Marcel Dekker, New York, NY. 1987.

LACERDA, T. H. S.; CABACINHA, C. D.; JUNIOR, C. A. A.; MAIA, R. D.; LACERDA, K. W. S. Artificial neural networks for estimating tree volume in Brazilian Savana. **Cerne**. v. 23, n. 4, p. 483-491, 2017.

LANSSANOVA, L. R.; UBIALLI, J. A.; ARCE, J.E.; PELISSARI, A.; FAVALESSA, C.M.; DRESCHER, R. Avaliação de funções de afilamento para a estimativa de diâmetro de espécies florestais comerciais do bioma amazônico mato-grossense. **Floresta**, v. 43. p. 215-224, 2013.

LEITE F. S; REZENDE, A. V. Estimativa do volume de madeira partindo do diâmetro da cepa em uma área explorada de floresta Amazônica de terra firme. **Revista Ciência Florestal**, Santa Maria, v. 20, n. 1, p. 71-82 2010.

LEITE, H. G.; ANDRADE, V. C. L. Um método para condução de inventários florestais sem o uso de equações volumétricas. **Revista Árvore**, v.26, n.3, p.321-328, 2002.

LEITE, H. G.; OLIVEIRA, F. H. T. Statistical procedure to test the identity of analytical methods. **Communications in Soil Science and Plant Analysis**. 33, n.7/8, p.1105-1118, 2002.

MACHADO, A. A. 1989. **Diagnóstico em regressão linear**. Lavras: ESAL. 73 p.

MACHADO, R., CORRÊA, R. F., BARBOSA, R. T., BERGAMASCHI, M. A. C. M. Escore da Condição Corporal e sua Aplicação no Manejo Reprodutivo de Ruminantes, São Carlos: **Embrapa Pecuária Sudeste**, 2008. 16p. (Circular técnica, 57).

MACHADO, S. A. Tabela de volume para Pinus taeda na região de Telêmaco Borba-PR. **Revista Floresta**, n. 10, p. 29-35, 1979.

MAGNUSSON, M. K., AVROLA, A., HURSTI K. U. K., ABERG, L., & SJODEN, P. O. Choice of organic foods is related to perceived consequences for human health and to environmentally friendly behaviour. **Appetite**, 40, p. 109-117, 2003.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immament in nervous activity. **The bullentin of mathematical biophysics**, S.I, v.5, n.4, p. 115-133, 1943.

MELO, L. C.; BARRETO, P. A. B.; OLIVEIRA, F. G. R. B.; NOVAES, A. B. Estimativas volumétricas em povoamentos de *Pinus caribaea* var. *hondurensis* no Sudoeste da Bahia. **Pesquisa Florestal Brasileira**, v.33, n.76, p.379-386, 2013.

MENEZES, N. P.; SILVA, J. N. M. 2003. **Equações de volume para florestas de terra firme nos Município de Tailândia e Paragominas**. In: Seminário de Iniciação Científica da Ufra, 1. Seminário de Iniciação Científica da Embrapa Amazônia Oriental, 7, 2003, Belém, PA. Resumos Expandidos. Belém: Universidade Federal Rural da Amazônia.

MIGUEL, E. P.; CANZI, L. F.; RUFINO, R. F.; SANTOS, G. A. Ajuste de modelo volumétrico e desenvolvimento de fator de forma para plantios de *Eucalyptus grandis* localizados no município de Rio Verde – GO. **Enciclopédia Biosfera**, Goiânia, vol. 6, n. 11, 2010.

MOSER, P.; OLIVEIRA, L. Z. **Regressão linear aplicada à dendrometria: uma introdução e iniciação ao R.**– Blumenau: Edfurb, 2017. 152p.

MOSER, P.; VIBRANS, A. C. MC ROBERTS, R. E. NAESSET, E. GOBAKKEN, T. CHIRICI, G., MURA, M., & MARCHETTI, M. (2017). Methods for variable selection in LIDAR-assisted forest inventories. **Forestry**, 90(1): 112-124.

MUGASHA W. A., MWAKALUKWA E. E., LUOGA E., MALIMBWI R. E., ZAHABU E., SILAYO D. S., SOLA G.,5 CRETE P., HENRY M., Kashindye A. Allometric Models for Estimating Tree Volume and Aboveground Biomass in Lowland Forests of Tanzania. **International Journal of Forestry Research**. p. 13, 2016.

NADAL, C. A. JULIANO, K. A.; RATTON, E. Testes estatísticos utilizados para a validação de regressões múltiplas aplicadas na avaliação de imóveis urbanos. **Bol.Ciênc. Geod.**, v.9, n°2, p.243-262, 2003.

OZÇELIK, R.; DIAMANTOPOLOU, M. J.; BROOKS, J. R.; WIANT JR, H. V.; Estimating tree bole volume using artificial neural network models for four Species in Turkey. **Journal of Environmental Management**. v. 91, p. 742-753, 2010.

PANDORFI, N.; SILVA, I. J. O.; SARNIGHAUSEN, V. C. R.; VIEIRA, F. M. C.; NASCIMENTO, S. T.; GUISELINI, C. Uso de redes neurais artificiais para predição de índices zootécnicos nas fases de gestação e maternidade na suinocultura. **Revista Brasileira de Zootecnia**, v.40, n.3, p.676-681, 2011.

PATTERSON, D. W. **Artificial Neural Networks: Theory and Applications**. Prentice Hall, 1995.

PAULA NETO, F. Tabelas volumétricas com e sem casca para *Eucalyptus saligna*. **Revista Árvore**, n. 1, p.31-54, 1977.

PEIXOTO, A. P. N. **Regressão linear simples**. UNIFESP, São Paulo, 2007.

PENG, C.; WEN, X. **Recent applications of artificial neural networks in forest resource management: an overview.** In: AMERICAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE WORKSHOP, 1999, Orlando.

REIS, L. P.; LOPES, A.; SOUZA, D.; MAZZEI, L.; CAROLLINE, P.; GARCIA, H.; PEDRO, C.; SOARES, B.; MOREIRA, C.; ELETO, M.; FERNANDES, L.; ROBERTO, A. Prognosis on the diameter of individual trees on the eastern region of the amazon using artificial neural networks. **For. Ecol. Manage.** v. 382, p. 161–167, 2016.

PRECHELT, L. **A Set of Neural Network Benchmark Problems and Benchmarking Rules.** Universit at Karlsruhe, Technical Report 21/94, 1994.

REIS, D.R.; PLANGG, R.; TUNDISI, J.G.; QUEVEDO, D.M. Physical characterization of a watershed through GIS: a study in the Schmidt stream, Brazil. **Brazilian Journal of Biology**, v.75, p.S16-S29, 2015.

RIBEIRO, R. B. S.; GAMA, J. R. V.; MELO, L, O. Seccionamento para cubagem e escolha de equações de volume para a Floresta Nacional do Tapajós. **CERNE** | v. 20 n. 4. p. 605-612, 2014.

RIBEIRO, R. B. S.; GAMA, J. R. V.; SOUZA, A. L.; LEITE, H. G.; SOARES, C. P. B.; SILVA, G. F. Métodos para estimar o volume de fustes e galhos na Floresta Nacional do Tapajós. 2015.

SANQUETTA, C. R.; CORTE, A. P. D; RODRIGUES, A. L.; WATZLAWICK, L. F. **Inventários florestais: planejamento e execução.** 3 Ed. Curitiba: Multi-Graphic Gráfica e Editora. 406p. 2014.

SANTANA, L. E. A. S. **Otimização em comitês de classificadores:** uma abordagem baseada em filtros para seleção de subconjuntos de atributos. 2012. Tese (Doutorado em Ciência da Computação) – Universidade Federal do Rio Grande do Norte. 2012.

SCHNEIDER, P. R; SCHNEIDER, P. S. P; SOUZA, C. A. M. **Análise de Regressão aplicada à Engenharia Florestal.** 2. Ed- Santa Maria: FACOS, 2009. 294p.

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

SCOLFORO, (1993).

SCRINZI, G.; MARZULLO, L.; GALVAGNI, D. Development of a neural network model to update forest distribution data for managed alpine stands. **Ecological Modelling**. Amsterdam, v. 206, p. 331-346, 2007.

SERVIÇO FLORESTAL BRASILEIRO (SFB). **Florestas do Brasil em resumo:** dados de 2007 - 2012. Brasília: SFB, 2013. 188 p.

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

SILVA, J. N. M., CARVALHO, J. D., LOPES, J. C. A., & CARVALHO, M. D. (1984). Equações de volume para a Floresta Nacional do Tapajós. **Boletim de Pesquisa Florestal**, v. 8, 1984.

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

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.

SIQUEIRA, J. D. P. Tabelas de volume para povoamentos nativos de *Araucaria angustifolia* (Bert) O. Ktze, no sul do Brasil. **Floresta**, v. 8, n. 1, p. 7-12, 1977.

SISTEMA NACIONAL DE INFORMAÇÕES FLORESTAIS (SNIF). Boletim SNIF 2016. 2017, v 1, ed 2, p. 7.

SOARES, F. A. A. M. N.; FLÔRES, E. L.; CABACINHA, C. D.; CARRIJO, G. A.; VEIGA, A. C. P. Recursive diameter prediction for calculating merchantable volume of *Eucalyptus* clones without previous knowledge of total tree height using artificial neural networks. **Applied Soft Computing**. v. 12, p. 2030–2039, 2012.

SPURR, S. H. Forest Inventory. **The Ronald Press Company**, New York. 1952. 476 p.

TAFNER, M., XEREZ, M. R. I. **Rede Neurais Artificiais: Introdução e Principios de Neurocomputação.** EKO, 1996.

THIERSCH, A. **A eficiência das distribuições diamétricas para prognose de *Eucalyptus camaldulensis*.** 1997. 155 f. Dissertação (Mestrado em Engenharia Florestal) Universidade Federal de Lavras, Lavras, MG.

THOMAS, C.; ANDRADE, C. M.; SCHNEIDER, P. R.; FINGER, C. A. G. Comparação de equações volumétricas ajustadas com dados de cubagem e análise de tronco. **Ciência Florestal**, Santa Maria, v.16, n.3, p.319-327, 2006.

VALENÇA, M. **Aplicando redes neurais: um guia completo.** Olinda-PE: Livro Rápido, 2005. 284p.

VALENÇA, M. **Fundamentos das redes neurais: Exemplos em Java.** 2ª. Edição, Revisado e ampliado. Olinda-PE: Livro Rápido, 2011. 386p.

VALENTE, M. D. R.; QUEIROZ, W. T.; PINHEIRO, J. G.; MONTEIRO, L. A. S. Modelo de predição para o volume total de quaruba (*Vochysia inundata* ducke) via análise de fatores e regressão. **Revista Árvore**, Viçosa, v.35, n.2, p.307-317, 2011.

WIDROW, B. LEHR, M. A. **30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation, Proceedings of the IEEE**, v. 78, n. 9, p. 1415-1442, 1990.