



LORENA RAMALHO LEITE REIS

**REDES NEURAIAS ARTIFICIAIS PARA PROGNOSE DE
PRODUÇÃO EM POVOAMENTOS DE *Pinus taeda* NO
ESTADO DE SANTA CATARINA, BRASIL**

LAVRAS – MG

2021

LORENA RAMALHO LEITE REIS

**REDES NEURAIS ARTIFICIAIS PARA PROGNÓSE DE PRODUÇÃO EM
POVOAMENTOS DE *Pinus taeda* NO ESTADO DE SANTA CATARINA, BRASIL**

Trabalho de Conclusão de Curso
apresentado à Universidade
Federal de Lavras, como parte
das exigências do Curso de
Engenharia Florestal, para a
obtenção do título de Bacharel.

Prof. Dr. Natalino Calegario

Orientador

Doutor Daniel Dantas

Coorientador

Lavras – MG

2021

AGRADECIMENTOS

Primeiramente a mim, por não ter desistido e conseguido chegar até aqui, mesmo neste período conturbado de pandemia e situações adversas. Agradeço por ter continuado.

À minha família: minha mãe e ao meu pai (*in memorian*) por sempre me darem apoio; meus irmãos, que fazem minha vivência ser mais leve e divertida; e meus sobrinhos, motivos da minha alegria.

À Universidade Federal de Lavras, por toda estrutura disponibilizada para o desenvolvimento de minha graduação e a vivência, por meio da qual eu mais me desenvolvi em toda minha vida.

Aos meus professores Faustinho e Calegario, que me orientaram não só no TCC, mas em muitas questões da vida fora da UFLA. Ao meu coorientador Daniel, sempre atencioso e paciente.

À empresa Klabin S. A. por ser uma das melhores empresas que tive a oportunidade de conhecer, que possibilitaram a confecção deste estudo com a disponibilização da base de dados.

Às minhas amigas e amigos que, mesmo distantes, sempre estiveram por perto e foram essenciais para que eu não tivesse desistido no meio do caminho. São elas: Aline, Ana, Bruna, Larissa, Sabrina e Wakana; e eles: Fernando, Matheus e Víctor.

Agradeço, em especial, ao meu amigo Marcos (*in memorian*), um grande exemplo de ser humano para mim. Me ensinou a importância de viver o agora, mas planejar o futuro, e dar o meu melhor para os outros sem esperar nada em troca.

E a todos que, de alguma forma, contribuíram para a realização deste trabalho.

RESUMO

O objetivo geral deste trabalho foi aplicar e avaliar redes neurais artificiais (RNA) na projeção volumétrica de povoamentos de *Pinus taeda* localizados no estado de Santa Catarina, assim como comparar com os resultados obtidos pelo Modelo de crescimento e produção de Clutter (1993), comumente utilizado em empresas de base florestal. O treinamento e aplicação das RNA do tipo Multilayer Perceptron (MLP) foi realizado no *software* Rstudio versão 1.3.1093. Para todos os cenários, foram aplicados o algoritmo de treinamento *Resilient propagation* (RPROP+) e a função de ativação logística (ou sigmoide), de intervalo de 0 a 1, sendo necessária a normalização das variáveis de entrada e saída, além da transformação das variáveis qualitativas em sistema binário para enquadrarem na amplitude desejada. Para a obtenção do volume futuro a partir das RNA, foram utilizadas parcelas permanentes coletadas entre os anos de 2010 a 2018, tendo como variáveis de entrada quantitativas, idade, área basal, índice de sítio, volume, declividade, temperatura, e como variáveis qualitativas (categórica), material genético e tipo de solo. As estimativas do volume futuro obtidas com o emprego das RNA e pelo modelo de regressão foram avaliadas com base no coeficiente de correlação, na raiz do erro quadrado médio percentual (*Root Mean Square Error – RMSE %*), *Bias*, e no gráfico de dispersão dos erros percentuais. O treinamento da rede MPL foi realizado com 3 conjuntos diferentes de entrada, além de 5 cenários com proporções variadas de amostra para treinamento da rede e para validação, totalizando 15 redes de prognose. Para todos os testes em análise, as RNA apresentaram resultados superiores quando comparadas à equação de Clutter. Estes resultados confirmam a aplicabilidade prática de redes neurais artificiais para projeção do volume futuro para povoamentos de *Pinus Taeda*.

Palavras-chave: Volume. Estimativa. Inteligência artificial.

SUMÁRIO

PRIMEIRA PARTE	6
1. INTRODUÇÃO	6
2. REVISÃO DE LITERATURA.....	8
2.1. MODELAGEM DE CRESCIMENTO E PRODUÇÃO FLORESTAL.....	9
2.2. REDES NEURAIAS ARTIFICIAIS	11
2.2.1. DEFINIÇÃO.....	11
2.2.2. PROCESSOS DE APRENDIZAGEM	15
2.2.3. APLICAÇÃO DE RNA NO SETOR FLORESTAL.....	17
3. CONSIDERAÇÕES GERAIS	20
REFERÊNCIAS	21
SEGUNDA PARTE – ARTIGO	27
1. INTRODUÇÃO	28
2. MATERIAIS E MÉTODOS.....	29
2.1. DADOS.....	29
2.1.1. INVENTÁRIO FLORESTAL CONTÍNUO (IFC)	29
2.1.2. EDAFOCLIMÁTICOS.....	29
2.2. ORGANIZAÇÃO DOS DADOS	29
2.3. DEFINIÇÃO DOS TESTES.....	30
2.4. TREINAMENTO DAS REDES NEURAIAS ARTIFICIAIS	31
2.5. AJUSTE DO MODELO DE CLUTTER.....	33
3. RESULTADOS E DISCUSSÕES	34
4. CONCLUSÃO.....	40
REFERÊNCIAS	41
TERCEIRA PARTE.....	43

ESTRUTURAÇÃO DESTE TRABALHO DE CONCLUSÃO DE CURSO (TCC)

Este TCC foi segmentado em três partes. A primeira apresenta uma introdução sobre o assunto abordado no trabalho, o objetivo geral e os objetivos específicos, o referencial teórico e considerações, finalizando com as referências bibliográficas. A segunda parte trata do artigo científico originado do estudo deste trabalho, contendo em sua estrutura: introdução, metodologia, resultados e discussão, considerações finais e as referências utilizadas. Por fim, a última parte apresenta as considerações finais e sugestões para futuros trabalhos.

PRIMEIRA PARTE

1. INTRODUÇÃO

Os plantios de pinus ocupam cerca de 18% do território brasileiro (IBÁ, 2020) e fornecem matéria prima versátil para o mercado por possuir diversas finalidades, como construção civil, movelaria, confecção de laminados e indústria de papel e celulose. As características físicas, como a cor clara do lenho e a composição de fibras longas, que garantem alta resistência, o rápido crescimento e a adaptabilidade são fatores que garantem a competitividade do gênero (EMBRAPA, 2014). Para atender as demandas internas e externas do país, é necessário que o manejo do reflorestamento de pinus esteja alinhado com o potencial produtivo do local, material genético e tratos silviculturais, visando a máxima produção de madeira de qualidade em menor tempo e área mínima.

Conhecer a dinâmica do crescimento da floresta é imprescindível para o manejo, porque além de evidenciar o potencial de cada unidade de manejo, também permite estimar o volume presente e projetar o estoque futuro de madeira, auxiliando o gestor “a tomar decisões de quando, quanto, onde e como colher a madeira nos povoamentos florestais” (ALCÂNTARA, 2015). Para simulação do crescimento e produção florestal, foram desenvolvidos vários métodos ao longo do tempo, que são modelos de regressão lineares ou não lineares que estimam a variável dependente a partir de variáveis do plantio e do ambiente.

Com o avanço da Inteligência Artificial, várias metodologias vêm sendo aplicadas no setor florestal em substituição de métodos tradicionais, como as Redes Neurais Artificiais (RNA). Estas são modelos computacionais inspirados no funcionamento de aprendizagem do cérebro humano, por meio de padronização de estímulos externos e geração de respostas. Têm apresentado bons resultados e superioridade na precisão de estimativas (GORGENS, 2006) não só para crescimento e produção florestal, mas também em estudos de densidade básica da madeira em florestas nativas e plantadas, descrição do perfil do tronco, produtividade em SAFs, estoque de carbono, entre outros.

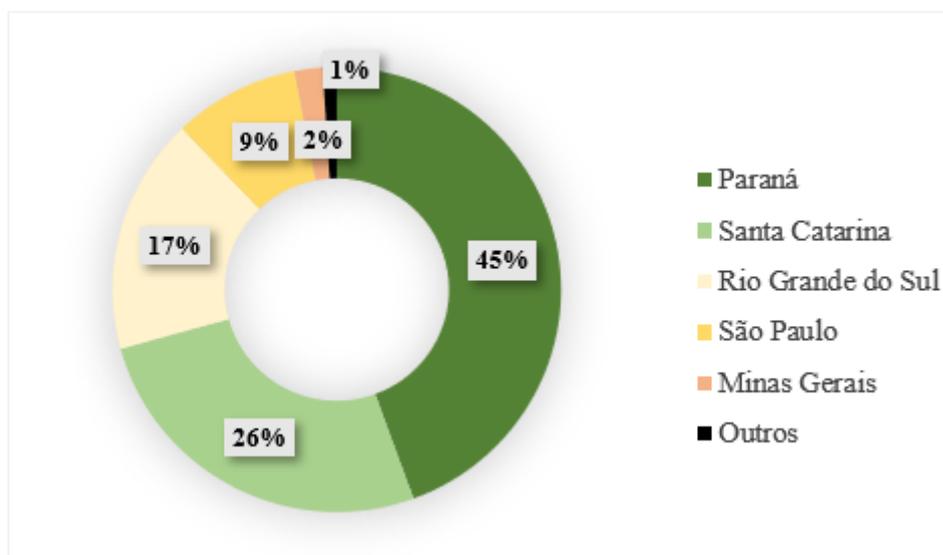
Além dos resultados positivos dos estudos com este mecanismo, outro fator que estimula a utilização de redes é a possibilidade de utilização de variáveis edafoclimáticas e categóricas, não convencional em modelos de regressão tradicionais.

Desta forma, o objeto de estudo deste trabalho é desenvolver modelos de predição de volume para povoamentos de pinus, a partir de testes com diferentes combinações de variáveis de entrada; testar redes neurais artificiais para projeção de volume para projeção de produção em plantios florestais de pinus, a fim de gerar mais uma alternativa para auxílio do manejo florestal e tomadas de decisão pelos gestores florestais, além de visar o avanço nos estudos de modelagem de dados do setor florestal brasileiro.

2. REVISÃO DE LITERATURA

No Brasil, o setor florestal tem sido um importante indicador de desenvolvimento econômico, social e ambiental, por promover mudanças econômicas locais (oportunidade de trabalho e geração de renda para a população local), contribuir com a adaptação e mitigação das mudanças climáticas e diminuir a pressão sobre florestas nativas. Em 2019, a extensão de áreas com reflorestamento era de 9,0 milhões de hectares, e aproximadamente 1,64 milhão de hectares eram florestas de pinus, o que equivale à 18% do total, sendo que a maior parte está localizada na região sul do país (Figura 1), onde o Paraná possui 45%, Santa Catarina 26% e o Rio Grande do Sul 17% das áreas (IBÁ, 2020).

Figura 1 – Distribuição dos plantios de *Pinus spp.* no Brasil por estado, em 2019.



Fonte: adaptado IBÁ (2020).

O gênero *Pinus* pertence à ordem Coniferae, do grupo das Gymnospermae. As espécies mais plantadas no território nacional são: *Pinus taeda*, *Pinus elliotti* (que toleram geadas), *Pinus caribaea*, *Pinus oocarpa* e *Pinus tecunumanii* (estes toleram déficit hídrico) (IPEF, 2016). A madeira pode ser utilizada na indústria de papel e celulose, construção civil, laminação, confecção de móveis, geração de energia e outros fins, de acordo com Medrado(2005, citado por NETTO, 2008, p.20).

Mesmo não tendo ocorrência natural na América do Sul, ele se adaptou muito bem no Brasil devido às condições edafoclimáticas, e possui elevada produtividade comparada com outros países por conta do manejo diferenciado e do melhoramento

genético (SHIMIZU, 2008). Diferentemente dos plantios de eucalipto, os de pinus obtiveram aumento de produtividade de 30,1 m³/ha de 2018 para 31,3 m³/ha em 2019. Entre 2009 e 2019, houve um aumento de 79,6% das exportações de celulose e um aumento de consumo aparente de celulose em 60% (IBÁ, 2020). De acordo com Cordeiro e Neves (2020), a pandemia por Covid-19 acelerou a queda de consumo de papel, que já estava em declínio nos últimos anos, mas aumentou para celulose por conta do aumento da demanda por produtos de higiene pessoal e embalagens para e-commerce (devido ao crescimento de compras de produtos pela internet).

Para que o setor florestal atenda as demandas internas e externas, é necessário que o planejamento das florestas seja feito com a maior exatidão possível para que se tenha o volume atual das florestas e seja possível calcular os volumes futuros de madeira.

2.1. Modelagem de crescimento e produção florestal

De acordo com Campos e Leite (2002, p. 243), “os três principais elementos necessários para o manejo de florestas equiâneas são: classificação de terras, prescrição de tratamentos silviculturais com definição de regime de corte e predição do crescimento e da produção”. Estes modelos são equações que incluem uma ou mais variáveis independentes e quantificáveis em nível de indivíduo, classe diamétrica ou povoamento total da floresta e têm o objetivo de descrever o estoque futuro de madeira (prognose) para ajustar o manejo da floresta, a partir dos valores de Incremento Corrente Anual (ICA), Incremento Médio Anual (IMA), Incremento Periódico (IP) ou Incremento Periódico Anual (IPA), por exemplo. Na modelagem, o que se busca são modelos que descrevam a dinâmica de uma floresta com maior nível de precisão (CARVALHO et al., 2011).

Para que o modelo de prognose seja bom, é essencial ser flexível o suficiente para permitir que um amplo número de alternativas de manejo possa ser simulado (VANCLAY, 1995), a fim de se encontrar uma solução ótima para o uso dos recursos florestais disponíveis. Ou seja, ser possível a estimação do volume de madeira a ser retirada periodicamente das florestas para alimentar as empresas e indústrias florestais de forma planejada. A escolha da melhor abordagem de modelagem deve levar em conta custos e benefícios, segundo Buongiorno e Gilles (2003, citado por MAGALHÃES, 2014, p.2), além da facilidade de uso, a precisão das previsões e o realismo biológico, segundo Young e Giese (2003, citado por MAGALHÃES, 2014, p.2).

De acordo com Carvalho et al. (2011), antes de se conhecer o estoque presente e futuro de um povoamento, são necessários serem feitos anteriormente: relação hipsométrica, modelagem das alturas das árvores médias dominantes e codominantes, classificação do sítio, área basal, modelagem do volume por árvore individual e, enfim, a predição e projeção da produção, expressa em volume por unidade de área. As equações existentes mais utilizadas para estimar crescimento e produção são, em sua maioria, modelos de regressão linear e não linear (GARCIA, 1988). No Brasil, os modelos mais utilizados são em nível de povoamento e estimam a produção por meio de variáveis como idade, área basal e índice local (ALCÂNTARA, 2015). Como exemplos, podem ser citados os de Buckman (1962) e Clutter (1963). Do ponto de vista da mensuração, será selecionado o modelo que tiver maior exatidão e ausência de tendência (SOARES, PAULA NETO; SOUZA, 2006). Entretanto, modelos volumétricos são altamente dependentes de variáveis categóricas, como local, material genético, espaçamento, tipo de solo, entre outros, que comumente apresentam relações de não linearidade com o volume, fazendo com que a introdução destes parâmetros seja uma tarefa complexa (BINOTI; BINOTI; LEITE, 2014a). Outra limitação frequente é a adição de novos clones no modelo e a consequente exclusão dos clones antigos, resultando no descarte de dados da base (MARTINS et al., 2015). Também existe a possibilidade de modelos que “podem apresentar tendenciosidades e níveis de erro inaceitáveis para os objetivos do manejo florestal”, segundo Binoti et al. (2014b, p. 284).

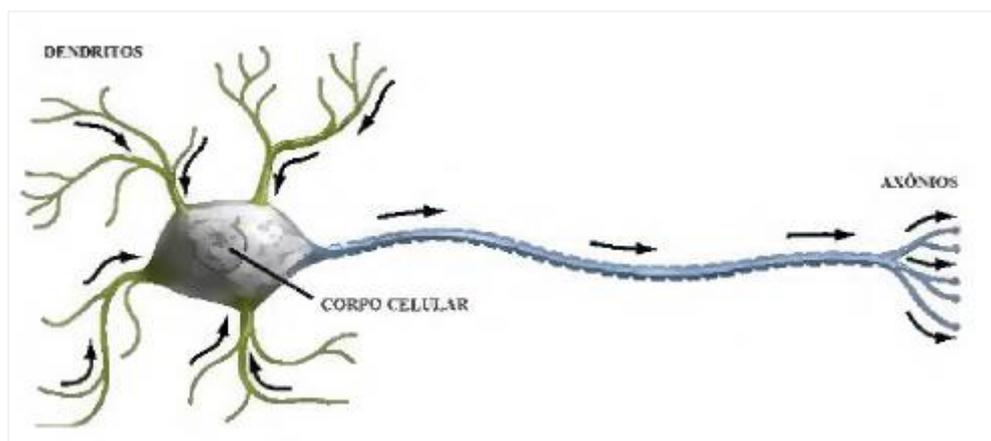
Entretanto, com o avanço de programas computacionais e Inteligência Artificial (IA), novos processos matemáticos passaram a ser incorporados nos estudos de modelagem florestal, resolvendo diversos problemas comuns nos modelos volumétricos tradicionais, a fim de fornecer estimativas mais precisas e alternativas de redução de custo para o processo de prognose de produção e crescimento (MAGALHÃES, 2014). Dentre estes avanços, estão o Aprendizado de Máquina, Máquinas de Vetores de Suporte e Redes Neurais Artificiais (RNA), que possuem resultados promissores, visto que podem incluir dados categóricos aos modelos, não ser necessária estratificação e apresentar melhor precisão das estimativas (MAGALHÃES, 2014).

2.2. Redes Neurais Artificiais

2.2.1. Definição

O cérebro humano possui aproximadamente 10 bilhões de células complexas, os neurônios. Um neurônio (Figura 2) pode ser dividido em três partes, sendo: dendritos, os terminais de entrada onde o estímulo externo é captado; o corpo, onde a informação é processada; e os axônios, que são terminais de saída do impulso nervoso para outro neurônio (FERNEDA, 2006). Estes neurônios “se comunicam” através de sinapses, que é a região entre os axônios da célula nervosa X com os dendritos da célula Y e é este movimento que faz com que haja reação diante de acontecimentos (CARVALHO, 2009), como tirar a mão do fogo e comer quando sente fome. Cada neurônio forma, em média, entre mil e dez mil sinapses (HAYKIN, 2001). Um cérebro pode ser formado por uma rede com mais de 60 trilhões de ligações sinápticas, sendo capaz de realizar processamentos com mais eficiência que um computador convencional por conseguir realizar simultaneamente estes processos complexos (VALENÇA, 2008).

Figura 2 – Estrutura do neurônio biológico.



Fonte: GUYTON, HALL (2006).

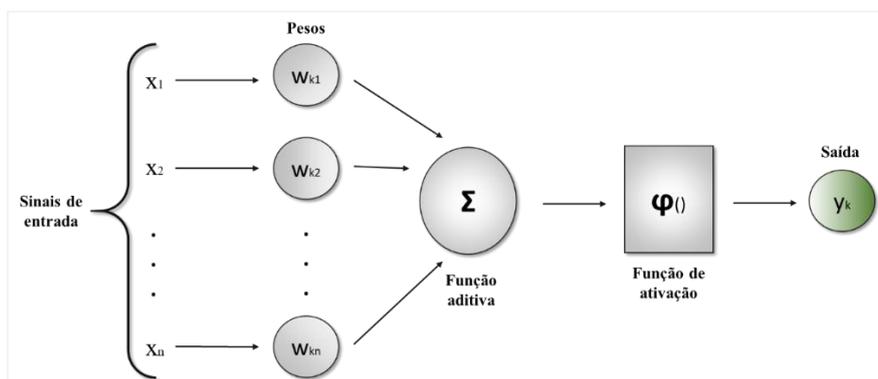
As redes neurais artificiais (RNA) possuem este nome porque tentam imitar o funcionamento do cérebro humano na maneira como o conhecimento é adquirido (processo de aprendizagem): por meio de reconhecimento e associação dos padrões (estímulos do ambiente); e determinação das forças de conexão entre os neurônios, que definem o peso de cada aprendizado (HAYKIN, 2001). De acordo com a literatura, uma RNA pode ser definida como um sistema computacional baseado em modelos matemáticos, com elementos de processamento distribuídos paralelamente, denominados

neurônios artificiais ou nodos, dispostos em camadas – sendo uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída – e conectados entre si de forma específica (pesos sinápticos) para desempenhar determinada função (BINOTI et al., 2010, 2014c; BRAGA et al., 2000; BULLINARIA, 2014; DANTAS et al., 2018; MARTINS et al., 2015; HAYKIN, 2001).

A Figura 3 apresenta um modelo de neurônio artificial que contém os seguintes elementos básicos (ALCÂNTARA, 2015; HAYKIN, 2001):

- Terminal de entrada (x_1, x_2, \dots, x_j): valores assumidos pelas variáveis da base de dados e contém sinais de entrada. São normalizados para melhorar a eficiência computacional;
- Pesos sinápticos ($w_{k1}, w_{k2}, \dots, w_{kn}$): caracterizam os terminais de entrada e podem variar entre valores negativos e positivos, dependendo se são sinapses inibitórias ou excitatórias;
- Somador (Σ) para acumular os sinais de entrada, constitui um combinador linear dos sinais; e
- Função de ativação (φ), que restringe a amplitude da saída de um neurônio, e normalmente ocorre em um intervalo fechado unitário, entre 0 e 1, ou alternativamente, entre -1 e 1.

Figura 3 – Modelo de neurônio artificial.



Fonte: adaptado de HAYKIN (2001).

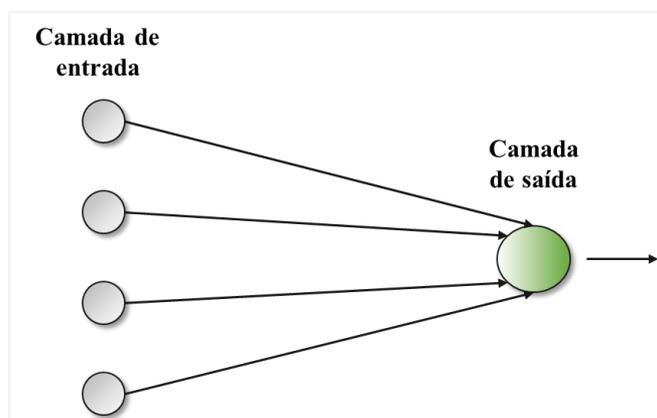
A saída de um neurônio artificial, desta forma, é dada pela equação 1 a seguir, em que S_i é a saída do neurônio, x_j é a camada de entrada e w_{kn} são os pesos sinápticos (BENTES, 2018):

$$S_i = \sum_j^n W_{kn} X_j \quad (1)$$

Segundo Alcântara (2015), as redes neurais artificiais podem ser classificadas quanto ao tipo de conexões, conectividade e número de camadas. Quanto ao tipo de conexão, existem redes que são alimentadas adiante (*feedforward*) e redes cíclicas (*feedback*), ou seja, que são realimentadas (HAYKIN, 2001). Em relação à conectividade, de acordo com Bullinaria (2014), as redes podem ser classificadas como totalmente conectadas (quando todos os neurônios de uma camada se conectam com todos os neurônios da camada posterior) e parcialmente conectadas (quando não há conexão em alguma parte da rede). Por fim, a classificação quanto ao número de camadas é dividida em três estruturas: Linear (*Perceptron*), Perceptron de Múltiplas Camadas (MLP – *Multilayer Perceptron*) e Redes de funções de base radial (RBF – *Radial base function*) (BINOTI et al., 2010).

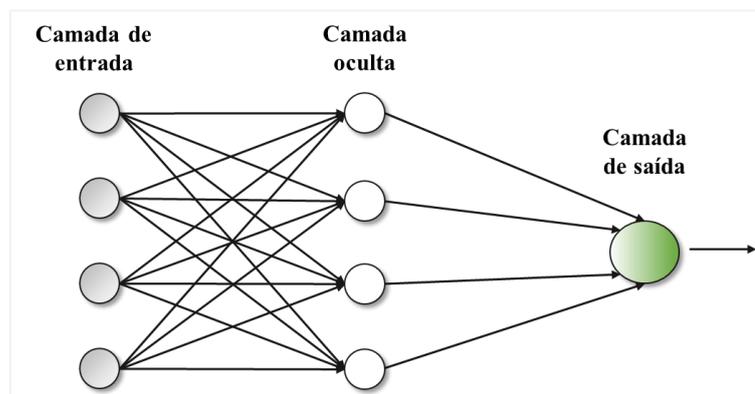
A Rede Linear possui apenas uma camada para processamento, a mesma camada de saída, e é utilizada em situações mais simples no qual os problemas são linearmente separáveis (GRANATYR; PACHOLOK, 2018; MINSKY; PAPERT, 1969). A Rede de Múltiplas Camadas, como o próprio nome diz, é estruturada com uma ou mais camadas intermediárias para processamento, além das camadas de entrada e saída, sendo mais utilizadas para base de dados extensas e com variáveis contínuas e (FREIMAN, 2004). A RBF possui três camadas, sendo a de entrada, uma oculta e a de saída. As funções de ativação são não lineares de base radial, e geralmente é aplicada a base gaussiana (BINOTI et al., 2010; BRAGA; CARVALHO; LUDEMIR, 2007; HAYKIN, 2001).

Figura 4 – Rede *feedforward* com uma camada de processamento, a camada de saída.



Fonte: Do autor (2021).

Figura 5 – Rede MPL *feedforward* com duas camadas de processamento, sendo uma camada oculta e a camada de saída.



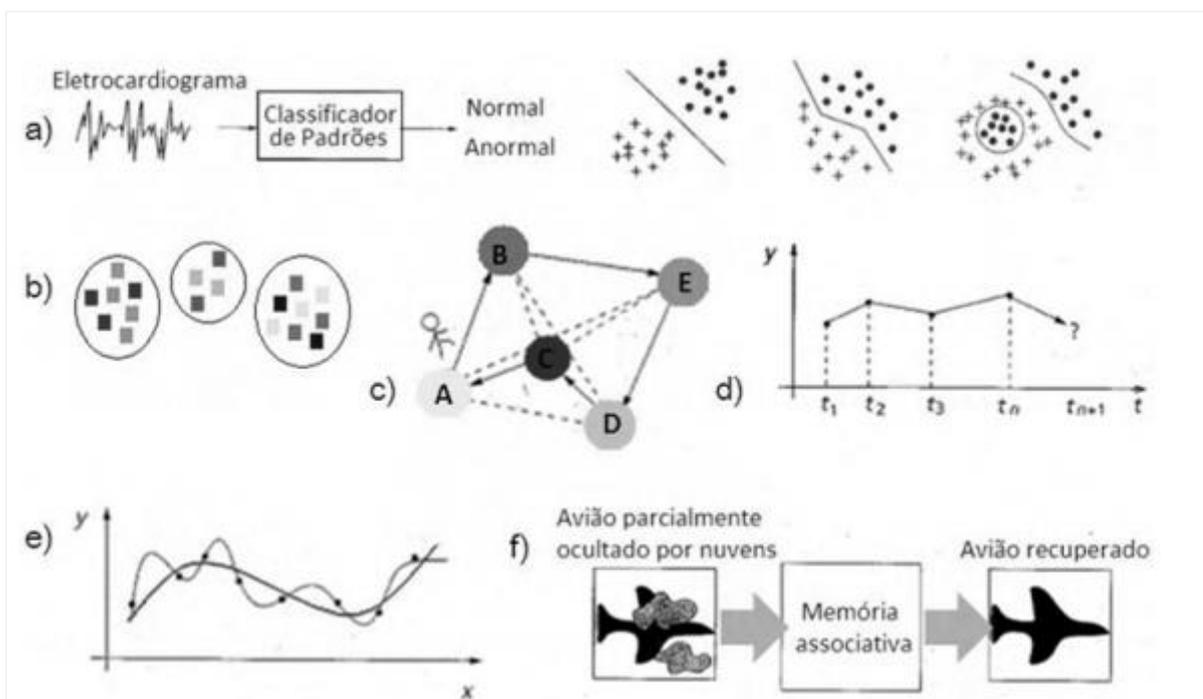
Fonte: Do autor (2021).

Do ponto de vista da modelagem, os estudos por redes se justificam porque são capazes de resolver problemas complexos, mesmo com poucos dados; tolerância à ruídos e falhas (*outliers*); introdução de variáveis categóricas; poder de aproximação de uma função sem ser necessária a especificação do modelo; detectar relações não lineares complexas entre variáveis de entrada e saída; caráter dinâmico, com possibilidade de ser retreinada mesmo que os dados de entrada mudem (ALCÂNTARA, 2015; BRAGA et al., 2000; BRAGA; CARVALHO; LUDEMIR, 2007; BULLINARIA, 2014; HAYKIN, 2001; LOPES et al., 2020; MAGALHÃES, 2014; ÖZÇELIK et al., 2013). De acordo com Morais (2010), “a solução de problemas através de RNAs é bastante atrativa, já que a arquitetura de paralelismo natural cria a possibilidade de um desempenho superior ao dos modelos convencionais”. Entretanto, o dimensionamento do número de camadas é uma dificuldade, pois mesmo existindo métodos empíricos para dimensioná-la (BULLINARIA, 2003), o melhor método ainda é por tentativa e erro (GÖRGENS, 2006). O excessivo número de neurônios pode prejudicar o poder de memorização, consequentemente diminuindo o poder de validação (processo denominado *overfitting*), e o escasso número de neurônios pode não ser suficiente para a resolução do problema (processo conhecido como *underfitting*) (BINOTI et al., 2014c; GÖRGENS, 2006; SILVA; SPATTI; FLAUZINO, 2010; VENDRUSCOLO et al., 2015).

As principais aplicações de RNA se aplicam em problemas de aproximação de função, classificação de padrões, agrupamento de dados, predição (séries temporais), otimização, recuperação de conteúdo e controle (JAIN; MAO; MOHIUDDIN, 1996; HAYKIN, 2001; BINOTI et al., 2010; BINOTI et al., 2014c). A projeção de parâmetros

de uma função de Volume que descreva o crescimento de um povoamento florestal se enquadra no problema de aproximação de função, que consiste em projetar uma rede neural que aproxime a função desconhecida $f(x)$, que descreve o mapeamento dos pares de entrada-saída $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ de um conjunto de n padrões de treinamento (BINOTI et al., 2014c).

Figura 6 – problemas que podem ser resolvidos com RNAs: a) classificação de padrões, b) agrupamento de dados, c) otimização, d) predição (séries temporais), e) aproximação de função e f) recuperação de conteúdo.



Fonte: BINOTI (2010).

2.2.2. Processos de aprendizagem

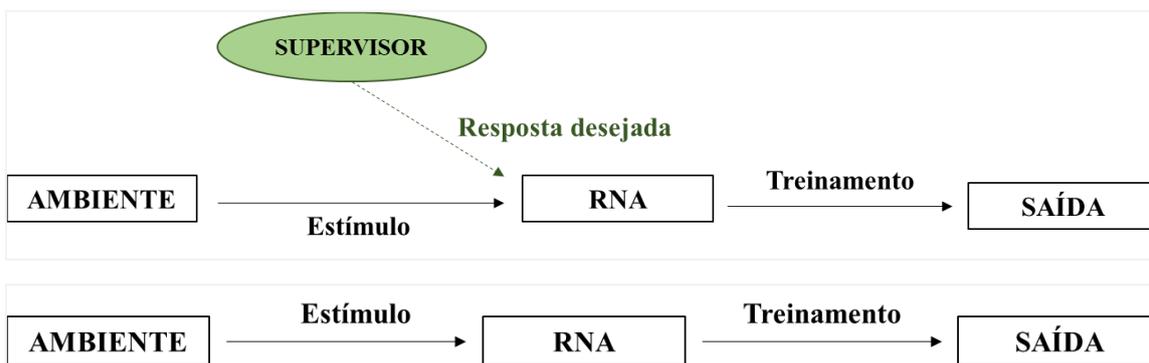
Como introduzido anteriormente, as redes têm a capacidade de aprender, e este processo pode ser definido da seguinte forma, de acordo com Haykin (2001):

“Aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede será inserida.”

Os conjuntos de processos pré-definidos para que a rede aprenda são denominados algoritmos de aprendizado e, apesar de existirem diversos tipos, todos buscam minimizar o erro e levarem ao menor tempo possível para processamento dos dados (MORAIS,

2010). São diferenciados pela forma que os pesos sinápticos são ajustados, podendo ser categorizados em treinamento supervisionado e treinamento não supervisionado (HAYKIN, 2001; MORAIS, 2010). O método supervisionado é o mais comum no treinamento de RNA e possuem este nome porque o supervisor, também chamado de professor, fornece à rede as saídas desejadas de acordo com os padrões de entrada, com o objetivo de indicar para a rede quais ajustes nos pesos devem ser feitos para minimizar o erro entre o valor estimado e o desejado (BENTES, 2018; JAIN; MAO; MOHIUDDIN, 1996; MORAIS, 2010). O método não supervisionado, no entanto, não conta com as respostas desejadas, apenas com os padrões de entrada (JAIN; MAO; MOHIUDDIN, 1996). Desta forma, a rede processa esse conjunto de padrões e busca classificá-las automaticamente. Este tipo de aprendizado precisa de dados redundantes para que seja possível encontrar os padrões (FERNEDA, 2006).

Figura 7 e 8 – Esquemas de aprendizado supervisionado e não supervisionado, respectivamente.

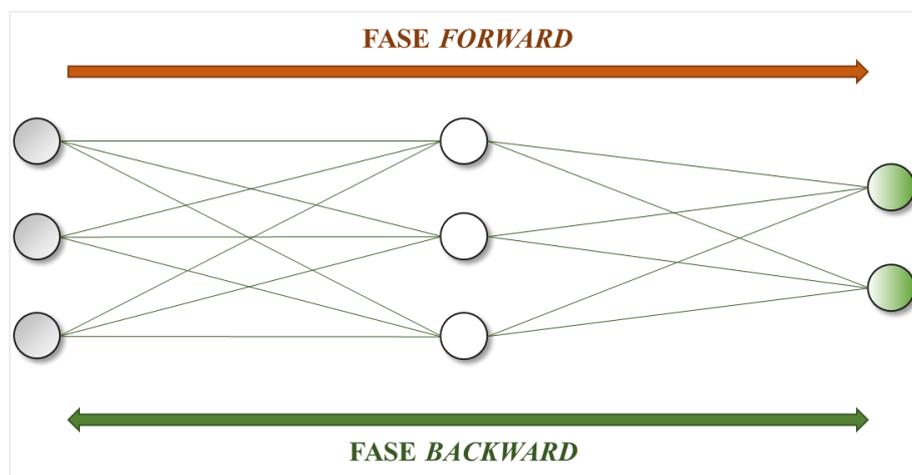


Fonte: adaptado de MORAIS (2010).

O algoritmo de retropropagação do erro (*ErrorBackPropagation*) é um dos processos de aprendizagem supervisionado mais conhecidos (MORAIS, 2010), e o mais utilizado para treinar redes de múltiplas camadas (BENTES, 2018; FERNEDA, 2006; HAYKIN, 2001; HEATON, 2011; VALENÇA, 2008). Desenvolvido por Rumelhart, Hinton e Williams no final da década de 80 (RUMELHART; HINTON; WILLIAMS, 1986), a ideia básica deste algoritmo é definir os erros nas respostas intermediárias de acordo com a diferença nas respostas desejadas e as respostas fornecidas pela rede, a fim de atualizar os pesos para minimização deste erro (DIAMANTOPOULOU; GEORGIU; PAPAMICHALL, 2006). Na prática, o algoritmo funciona em duas etapas, *forward* e *backward*, respectivamente. A primeira serve para relacionar e entender os padrões entre os pares de entradas e saídas e a segunda compara as respostas desejadas e as fornecidas

pela rede, calcula o erro e ajusta os pesos, a fim de melhorar as futuras respostas com a diminuição do erro (FERNEDA, 2006; MORAIS, 2010). O modelo MLP possui grande poder de validação por conta desta forma de treinamento (FREIMAN, 2004).

Figura 9 – Funcionamento do algoritmo *backpropagation*.



Fonte: Do autor (2021).

2.2.3. Aplicação de RNA no setor florestal

De acordo com Dantas et al. (2018), o uso de redes tem sido estudado em diversas áreas, como engenharia de sistemas e computação, medicina, ciências econômicas e outras, porém ainda é pouco difundida na área florestal, mesmo que os estudos feitos até hoje apresentem bons resultados, principalmente para a área de modelagem. Chiarello et al. (2019) apresentam que estes estudos têm aumentado continuamente no setor, sendo que 32% dos artigos analisados em sua pesquisa são sobre previsão de volume total comercial, sendo que 78% deste fazem o uso de MLP e 63% provêm de pesquisadores brasileiros.

Os estudos a seguir são exemplos de algumas das pesquisas nas diversas áreas da engenharia florestal com a aplicação de redes neurais artificiais que apresentaram resultados consideráveis.

Binoti, Binoti e Leite (2014a), com o objetivo de avaliar formas de treinamento de RNA para estimar volume de árvores e comparar com a estimação volumétrica tradicional pelo modelo de Schumacher e Hall (1933), observaram que houve pouca diferença entre as estimativas por RNA ou pelo modelo tradicional, porém o

processamento com RNA reduziu o tempo gasto com ajustes e avaliações de modelos volumétricos por estrato.

Martins et al. (2015) conseguiram definir configurações adequadas de redes para prognose da produção florestal de plantios de eucalipto, concluindo que os algoritmos de aprendizagem *Resilient Propagation*, *Quick Propagation* e *Scaled Conjugate* são ótimos.

Lopes et al. (2020) estudaram modelos de regressão em Sistema Agroflorestal (SAF) para *Khaya ivonensis* (mogno africano) e provaram que tanto o modelo de regressão de Schumacher e Hall como RNA estimam com exatidão o volume das árvores, sendo maior em configuração MLP, com 7 neurônios na camada oculta, função sigmoideal nas camadas de entrada e saída, treinada pelo algoritmo *resilient propagation*, utilizando com variáveis de entrada o DAP, altura comercial, locais e a classe diamétrica. O emprego da RNA dispensa a necessidade de estratificação por local e permite o uso de uma menor quantidade de dados de árvores-amostra.

Filho e Tusset (2012) testaram o uso de redes na identificação da madeira de catorze espécies nativas, através de análise de cor da madeira e a taxa de acerto para a classificação foi de aproximadamente 66%.

Almeida (2015) ao aplicar as RNAs para estimativa de altura total e volume em povoamentos de *Araucaria angustifolia*, encontrou uma correlação de 0,863 e RQME de 13,94 para a variável altura total. Já para a variável volume encontrou uma correlação variando entre 0,997 e 0,998 e RQME de 1,09.

Lafetá et al. (2018) avaliou a eficiência das RNAs para estimar a eficiência do uso de nutrientes no tronco de indivíduos de *Eucalyptus* em um plantio de híbridos de *Eucalyptus grandis* W. Hill ex Maiden \times *Eucalyptus camaldulensis* Dehnh em Minas Gerais e concluiu que a modelagem por redes utilizando apenas uma amostra de casca no DAP é adequada para estimar o coeficiente de utilização biológica do tronco.

Ozçelik et al. (2017) estudaram a aplicação de redes para predição de biomassa para pinus, que sugere que o modelo de rede LMANN alcançou maior precisão do que a derivada do modelo de regressão não linear NLR.

Dantas et al. (2018) concluíram com o estudo “Comitê de redes neurais artificiais para estimação do volume individual de árvores de *Eucalyptus*” que utilizando um comitê

redes, foi possível reduzir o erro relativo médio das estimativas de volume. No entanto, apenas uma única rede neural artificial é perfeitamente aplicável ao problema, com desempenho satisfatório e menor custo operacional.

Cunha Neto et al. (2017) com o estudo de propor, construir e validar um modelo de estimação de volume de árvores em um povoamento clonal de híbridos de *Eucalyptus grandis* W. Hill ex Maiden x *Eucalyptus urophylla* S.T. Blake, no município de Paragominas/PA, provaram que RNAs são válidas e confiáveis para a estimativa individual de volume para o clone, portanto, podem ser utilizadas como alternativa para esta estimativa.

Vendruscolo et al. (2015) compararam a modelagem por regressão não linear e por redes neurais artificiais para a estimativa da altura de árvores de eucaliptos em Mato Grosso e mesmo ambas serem eficientes para a predição da altura de árvores, as redes neurais apresentaram critérios estatísticos levemente superiores em relação à regressão.

Silva et al. (2018) avaliaram a utilização das RNA para estimar a densidade básica de espécies do cerrado *sensu stricto* com a utilização de variáveis dendrométricas e o resultado foi satisfatório.

Binoti et al. (2012) realizaram um estudo de modelo de distribuição diamétrica baseado em um modelo de autômatos celulares para a simulação de desbaste e permitiu estimar esse tipo de distribuição futura com exatidão.

Dantas et al. (2020) avaliaram a qualidade de estimativas volumétricas para *Eucalyptus spp* usando um modelo de efeitos mistos, rede neural artificial (ANN) e máquina de suporte vetorial (SVM). O resultado confirmou que embora os modelos de efeitos fixos sejam amplamente utilizados no setor florestal brasileiro, há métodos mais eficazes para modelagem de variáveis dendrométricas.

Vendruscolo et al. (2015), pesquisaram a eficácia do uso de redes neurais artificiais e modelagem linear de efeito misto para estimativa de altura para *Tectona grandis* em função de variáveis numéricas e categóricas. Eles foram estatisticamente eficientes na previsão a altura individual de *Tectona grandis*, com desvios abaixo de 12%, mesmo com uma redução de 80% nos dados utilizados em treinamento e ajuste. A eficácia do uso da RNA foi superior em árvores de maiores diâmetros.

3. CONSIDERAÇÕES GERAIS

Com este trabalho, foi possível analisar que a importância das aplicações de redes neurais artificiais vêm aumentando na área florestal e estão sendo bem exploradas para esta finalidade de projeção de volume, principalmente para espécies de eucalipto, entretanto são deficientes os estudos para espécies de pinus. O uso de RNA possibilita a inclusão de dados categóricos e apresentou bons resultados quando acrescentados dados edafoclimáticos na camada de entrada. Percéptrons multicamadas apresentam bons resultados por conta da variedade de algoritmos de aprendizagem, o qual o algoritmo de retropropagação e erro possui mais referências.

REFERÊNCIAS

- ALCÂNTARA, A. E. M. de. **Redes neurais artificiais para prognose do crescimento e da produção de povoamentos de eucalipto em Minas Gerais**. 2015. 58 p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal. Universidade Federal de Viçosa (UFV), Viçosa, 2015.
- ALMEIDA, T. G. de. **Comparação entre redes neurais artificiais e modelos tradicionais de regressão para estimar altura e volume em povoamentos de *Araucaria angustifolia* (Bertol.) O. Kuntze**. 2015. 56 p. Monografia (Bacharel em Engenharia Florestal) - Universidade Federal do Paraná (UFPR), Curitiba, 2015.
- BENTES, B. A. H. F. **Aplicação de redes neurais artificiais para estimativa de volume de árvores em pé em uma floresta ombrófila densa na Amazônia Oriental**. 2018. 59 p. Dissertação (mestrado) – Programa de Pós-Graduação em Ciência Florestal. Universidade Federal Rural da Amazônia (UFRA), Pará, 2018.
- BINOTI, M. L. M. da. S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. 2010. 54p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal. Universidade Federal de Viçosa (UFV), Viçosa, 2010.
- BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G.; SILVA, A.; SANTOS, A. C. de. A. Modelagem da distribuição diamétrica em povoamentos de eucalipto submetidos a desbaste utilizando autômatos celulares. **Revista Árvore**, Viçosa-MG, v.36, n.5, p.931-939, 2012.
- BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G. Configuração de redes neurais artificiais para estimação do volume de árvores. **Ciência da Madeira**, Pelotas, v.5, n.1, p.58-67, 2014a.
- BINOTI, M. L. M. S.; BINOTI, D. H. B; LEITE, H. G.; GARCIA, S. L. R.; FERREIRA, M. Z.; RODE, R.; SILVA, A. A. L. Redes neurais artificiais para estimação do volume de árvores. **Revista Árvore**, Viçosa, v.38, n.2, p.283-288, 2014b.

BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G.; SILVA, A. A. L.; PONTES, C. Utilização de redes neurais artificiais para a projeção da Distribuição diamétrica de povoamento equiâneos. **Revista Árvore**, Viçosa, v.38, n.4, p.747-754, 2014c.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDEMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2007. 260p.

BUCKMAN, R. E. **Growth and yield of red pine in Minnesota**. Washington. 1962. 50p.

BULLINARIA, J. A. **Introduction to neural computation**: notas de aula. 2014. Disponível em: <<http://www.cs.bham.ac.uk/~jxb/inc.html>>. 2014.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal**: perguntas e respostas. Viçosa: Editora UFV, 2002. 408 p.

CARVALHO, A. P. de L. F. de. Redes Neurais Artificiais. **Instituto de Ciências Matemáticas e de Computação**. 2009. Disponível em: <<https://sites.icmc.usp.br/andre/research/neural/>>. 2009. Acesso em: 13/04/2021.

CARVALHO, S. P. C.; CALEGARIO, N.; FONSECA, F.; BORGES, L. A. C.; MENDONÇA, A. R.; LIMA, M. P. Modelos não lineares generalizados aplicados na predição da área basal e volume de *Eucalyptus* clonal. **CERNE**, v. 17, n. 4, p. 541-548, 2011.

CHIARELLO, F.; STEINER, M. T. A.; OLIVEIRA, E. B.; ARCE, J. E.; FERREIRA, J. C. Artificial Neural Networks Applied in Forest Biometrics and Modeling: State of the Art. **CERNE**, v. 25, n. 2, p.140-155, 2019.

CLUTTER, J. L. Compatible growth and yield models for loblolly pine. **Forest Science**. Bethesda, v. 9, n. 3, p. 354-371, 1963.

CORDEIRO, J.; NEVES, M. Efeitos da pandemia por covid-19 no setor de celulose e papel. **O Papel**, p. 31-32, 2020.

CUNHA NETO, E. M. da.; CANTÃO, B. P.; CALLEGARI, D. M.; GOUVEIA, D. M.; ROCHA, J. E. C. da. Redes neurais artificiais na estimativa do volume individual de árvores de um híbrido de eucalipto, no município de Paragominas-PA. **II Congresso Internacional das Ciências Agrárias (COINTER)**. PDVAgro, 2017.

DANTAS, D.; ARRIEL, T. G.; CALEGARIO, N.; LACERDA, W. S. Comitê de redes neurais artificiais para estimação do volume individual de árvores de *Eucalyptus*. **Caderno de Ciências Agrícolas**, v. 10, n. 1, p. 01–06, 2018.

DANTAS, D.; CALEGARIO, N.; ACERBI JÚNIOR, F. W.; CARVALHO, S. P. C.; ISAAC JÚNIOR, M. A.; MELO, E. A. Multilevel nonlinear mixed-effects model and machine learning for predicting the volume of *Eucalyptus* spp. trees. **CERNE**, v. 26, n. 1, p. 48-57, 2020.

DIAMANTOPOULOU, M. J.; GEORGIU, P. E.; PAPAMICHALL, D. M. A time delay artificial neural network approach for flow routing in a river system. **Hidrology and Earth System Sciences Discussions**, v.3, p. 2735-2756, 2006.

EMBRAPA FLORESTAS. **Cultivo de Pinus**. 2014. Disponível em: <https://www.spo.cnptia.embrapa.br/conteudo?p_p_id=conteudoportlet_WAR_sistemas_deproducao16_1ga1ceportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=1&p_r_p_-6293187_sistemaProducaoId=3715&p_r_p_-996514994_topicoId=3228>

FERNEDA, E. Redes neurais e sua aplicação em sistemas de recuperação de informação. 2006. **Ci. Inf.**, Brasília, v. 35, n. 1, p. 25-30, 2006

FREIMAN, J. P. **Utilização de Redes Neurais Artificiais na previsão de indicadores financeiros para avaliação econômica de negócios em situação de risco**. Dissertação (Mestrado em Engenharia de Produção) – Pós graduação em Engenharia de Produção. Universidade Federal de Itajubá (UNIFEI), Itajubá, 2004.

GARCÍA, O. Growth modeling: a (Re)view. **New Zealand Forestry**, v.33, n.3, p.14-17, 1988.

GORGENS, E. B. **Estimação do volume de árvores utilizando redes neurais artificiais**. 2006. 94p. Dissertação (Mestrado) – Programa de Pós Graduação em Ciência Florestal. Universidade Federal de Viçosa (UFV), Viçosa, 2006.

GRANATYR, J; PACHOLOK, E. Redes Neurais Artificiais em R. **Udemy**. 2018

GUYTON, A. C.; HALL, J. E. Fisiologia Médica. 11^a ed. Rio de Janeiro. **Elsevier**, 2006. 366p.

HAYKIN, S. **Redes neurais: princípios e prática**. Trad. de Paulo Martins Engel. 2. ed. Porto Alegre: Bookman, 2001. 900 p.

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

INDÚSTRIA BRASILEIRA DE ÁRVORES (IBÁ). **Relatório IBÁ 2020**. 2020. Disponível em < <https://iba.org/datafiles/publicacoes/relatorios/relatorio-iba-2020.pdf>>. Acesso em: 18 fev. 2021.

INSTITUTO DE PESQUISA E ESTUDOS FLORESTAIS (IPEF). **IPEF Notícias n° 242**. 2016. Disponível em: <<https://www.ipef.br/publicacoes/ipefnoticias/ipefnoticias242.pdf>>. Acesso em 15 jan. 2021.

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

LAFETÁ, B. O.; SANTANA, R. C.; NOGUEIRA, G. S.; NEVES, J. C. L.; PENIDO, T. M. A. Eficiência de utilização de macronutrientes em eucalipto por método não destrutivo estimados por redes neurais artificiais. **Ciência Florestal**, Santa Maria, v. 28, n. 2, p. 613-623, 2018.

LOPES, L. S. S.; RODE, R.; PAULETTO, D.; BALONEQUE, D. D.; SANTOS, F. G. dos.; SILVA, A. R.; BINOTI, D. H. B.; LEITE, H. G. Uso de regressão e redes neurais artificiais na estimativa do volume de *Khaya ivorensis*. **Ciência da Madeira**, v. 11, n. 2, p. 74-84, 2020.

MAGALHÃES, J. G. de S. **Prognose e regulação de florestas de eucalipto utilizando modelo de crescimento e inteligência computacional**. 2014. 77p. Dissertação (Mestrado) – Programa de Pós Graduação em Ciência Florestal. Universidade Federal de Viçosa (UFV), Viçosa, 2014.

MARTINS, E. dos. R.; BINOTI, M. L. M. da. S.; LEITE, H. G.; BINOTI, D. H. B.; DUTRA, G. C. Configuração de redes neurais artificiais para prognose da produção de povoamentos clonais de eucalipto. **Agrária - Revista Brasileira de Ciências Agrárias**, Recife, v. 10, n .4, p. 532-537, 2015.

MORAIS, E. C. **Reconhecimento de padrões e redes neurais artificiais em predição de estruturas secundárias de proteínas**. 2010. 148 p. Tese (Doutorado) – Pós-graduação em Engenharia de Sistemas e Computação. Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2010.

MINSKY, M.; PAPERT, S. **Perceptrons: An Introduction to Computational Geometry**. Cambridge, Mass: M.I.T. Press, 1969. 258 p.

NETTO, C. C. **Dinâmica da distribuição diamétrica de povoamentos de *Pinus taeda* L. em diferentes idades e espaçamentos**. 2008. 106 p. Dissertação (Mestrado) – Pós-Graduação em Engenharia Florestal. Universidade Federal de Santa Maria (UFSM), Santa Maria, 2008.

ÖZÇELİK, E. R.; DIAMANTOPOULOU, M. J.; CRECENTE-CAMPO, F.; ELER, U.; Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology and Management**, v.306, n.1, p. 52-60, 2013

ÖZÇELİK, R.; DIAMANTOPOULOU, M. J.; EKER, M.; GÜRLEVIK, N. Artificial neural network models: an alternative approach for reliable aboveground Pine tree biomass prediction. **Forest Science**, v. 63, p. 291–302, 2017.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations by backpropagating errors. **Nature**, v.323, p533-536, 1986.

SHIMIZU, J. Y. et al. **Pinus na Silvicultura Brasileira**. Colombo: Embrapa Florestas, 2008, 223 p.

SILVA, I. N. da.; SPATTI, D. H.; FLAUZINO, R. A. Redes Neurais Artificiais para engenharia e ciências aplicadas. **Artliber**, 2010. 397 p.

SILVA, J. P. M.; CABACINHA, C. D.; ASSIS, A. L.; MONTEIRO, T. C.; ARAÚJO JÚNIOR, C. A.; MAIA, R. D. Redes neurais artificiais para estimar a densidade básica de madeiras do cerrado. **Pesquisa Florestal Brasileira**, Colombo, v. 38, p. 01-10, 2018.

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

SOARES, C. P. B.; PAULA NETO, F. de.; SOUZA, A. L. de. **Dendrometria e Inventário Florestal**. Viçosa: Editora UFV, 2006. 272p.

VALENÇA, M. **Fundamentos das redes neurais: exemplos em Java**. Livro Rápido, Olinda, 2008. 382p.

VANCLAY, J. K. Growth models for tropical forests: a synthesis of models and methods. **Forest Science**, v.41, p. 7-42, 1995.

VENDRUSCOLO, D. G. S.; DRESCHER, R.; SOUZA, H. S.; MOURA, J. P. V. M.; MAMORÉ, F. M. D.; SIQUEIRA, T. A. da. S.; Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. **Rev. Bras. Biom.**, São Paulo, v. 33, n. 4, p. 556-569, 2015.

SEGUNDA PARTE – ARTIGO

REDES NEURAS ARTIFICIAIS PARA PROGNÓSE DA PRODUÇÃO PARA POVOAMENTOS FLORESTAIS DE PINUS

Lorena Ramalho Leite Reis^{1a}, Daniel Dantas^{1b}, Natalino Calegari^{1c}

¹*Forest Science Department, Federal University of Lavras (UFLA), Lavras, Brazil, 3037, lorenaramalhor@gmail.com – ORCID: 0000-0001-7292-6978^a; dantasdaniel12@yahoo.com.br - ORCID: 0000-0002-7928-9155^b; calegari@ufla.br – ORCID: 0000-0001-8323- 1223^c.*

RESUMO

O objetivo geral deste trabalho foi aplicar e avaliar redes neurais artificiais (RNA) na projeção volumétrica de povoamentos de *Pinus taeda* localizados no estado de Santa Catarina, assim como comparar com os resultados obtidos pelo Modelo de crescimento e produção de Clutter (1993), comumente utilizado em empresas de base florestal. O treinamento e aplicação das RNA do tipo Multilayer Perceptron (MLP) foi realizado no *software* Rstudio versão 1.3.1093. Para todos os cenários, foram aplicados o algoritmo de treinamento *Resilient propagation* (RPROP+) e a função de ativação logística (ou sigmoideal), com intervalo de 0 a 1, sendo necessária a normalização das variáveis de entrada e saída, além da transformação das variáveis qualitativas em sistema binário para enquadrarem na amplitude desejada. Para a obtenção do volume futuro a partir das RNA, foram utilizadas parcelas permanentes coletadas entre os anos de 2010 e 2018, tendo como variáveis de entrada quantitativas, idade, área basal, índice de sítio, volume, declividade, temperatura, e como variáveis qualitativas (categórica), material genético e tipos de solo. As estimativas do volume futuro obtidas com o emprego das RNA e pelo modelo de regressão foram avaliadas com base no coeficiente de correlação, na raiz do erro quadrado médio percentual (*Root Mean Square Error – RMSE %*), *Bias* e no gráfico de dispersão dos erros percentuais. O treinamento da rede MPL foi realizado com 3 conjuntos diferentes de entrada, além de 5 cenários com proporções variadas de amostra para treinamento da rede e para validação, totalizando 15 redes de prognose. Para todos os testes em análise, as RNA apresentaram resultados superiores quando comparada a equação de Clutter. Estes resultados confirmam a aplicabilidade prática de redes neurais artificiais para projeção do volume futuro para povoamentos de *Pinus taeda*.

Palavras chave: Volume. Estimativa. Inteligência artificial.

1. INTRODUÇÃO

Os modelos de prognose de produção em povoamentos equiâneos devem ser precisos, realistas e eficientes, pois o estoque futuro é um dos parâmetros que norteia o planejamento estratégico de empresas florestais e nas tomadas de decisões de acordo com a disponibilidade de recursos (ZONETE et al, 2010; PENIDO, 2020), como planejamento de colheita, prospecção de novas áreas, comércio de madeira, logística da colheita (BERGER et al., 2002). Os modelos convencionais são baseados nas medições dendrométricas retiradas do inventário florestal contínuo, como idade, área basal, índice de sítio e volume atual, sendo o modelo de Clutter (1963) o mais utilizado no Brasil, por ser explícito e consistente (CAMPOS; LEITE, 2002). De acordo com Maestri et al (2013), o melhor modelo é o que tem a capacidade de ser genérico, realístico e preciso.

Com o avanço da Inteligência Artificial, novas metodologias de modelagem de produção florestal surgiram, como as Redes Neurais Artificiais (RNA), e têm apresentado bons resultados e superioridade na precisão de estimativas, no geral (GORGENS, 2006). Em geral, redes se destacam por sua forma de treinamento, (processo de aprendizagem e detecção de relações não lineares), tolerância à *outliers* e principalmente por permitir a inclusão de variáveis não dendrométricas, do tipo contínua ou categórica, como variáveis climáticas e classes de solo, possibilitando maior condição de generalidade (OÇZERIC et al., 2013; BULLINARIA, 2014). As aplicações de RNA para prognose têm apresentado resultados relevantes para o setor florestal, sendo aplicados majoritariamente em povoamentos de eucalipto (MAGALHÃES, 2014; MARTINS et al., 2015; CORDEIRO,2020; CASAS, 2021) e deficiente para populações de pinus.

O processo de inventário florestal para as medições dendrométricas tem alto custo, e modelos que dependem apenas destas variáveis necessitam de alta intensidade amostral para que o modelo reflita a realidade da melhor forma. Desta forma, objetivou-se o desenvolvimento de modelos de prognose de produção em povoamentos de pinus, utilizando redes neurais artificiais, com a utilização de diferentes combinações de variáveis de entrada, sendo dendrométricas e ambientais, para escolha do melhor modelo, e comparar a exatidão com o modelo convencional de Clutter.

2. MATERIAIS E MÉTODOS

2.1. Dados

2.1.1. Inventário Florestal Contínuo (IFC)

Os dados dendrométricos utilizados neste estudo são provenientes de 2.190 parcelas permanentes de inventários florestais contínuos, realizados entre os anos de 2010 e 2018, em povoamentos de *Pinus taeda*, distribuídos em 28 municípios do estado de Santa Catarina. Foram disponibilizadas as seguintes informações: idade (I) em anos, média das alturas das árvores dominantes (Hdom) em m, área basal (B) em $\text{m}^2 \text{ha}^{-1}$, número de fustes (n) por unidade de área, volume comercial (V) em $\text{m}^3 \text{ha}^{-1}$ ha e material genético. A partir de Hdom, foi calculado o Índice de Sítio (S) em metros, a fim de classificar a capacidade produtiva local (KLIMA, 2013).

2.1.2. Edafoclimáticos

Além dos dados de IFC, foram inseridas informações edafoclimáticas da área em estudo. A declividade média por parcela foi obtida através do Modelo Digital de Terreno do SIGSC – Sistema de Informações de Santa Catarina (2010), o tipo de solo pelo Instituto Brasileiro de Geografia e Estatística (2017) e dados climáticos do banco de dados WorldClim (FICK, 2017). As biovariáveis temperatura mínima no mês mais frio (BIO 6) e temperatura média no período mais frio (BIO 11) foram escolhidas por favorecerem a diminuição de crescimento e/ou ocasionar a mortalidade de indivíduos pela ocorrência de geadas, afetando o estoque de madeira (EMBRAPA).

2.2. Organização dos dados

Foram associados à base de IFC as informações edafoclimáticas, por localização dos centroides das parcelas, a partir do *software* ArcGIS.

O banco de dados foi organizado no formato utilizado em modelos de crescimento e produção, de forma que os dados por parcela sejam pareados por intervalo, sem sobreposição, como I_0 - I_1 , I_1 - I_2 ,... , I_{n-1} , I_n (CASAS, 2021). A estruturação do banco de dados apresenta o seguinte arranjo por observação (colunas): I_0 e I_1 , sendo a idade atual e a futura, respectivamente; B_0 , B_1 , que são as áreas basais relacionadas às idades; S representando o índice de sítio da parcela; V_0 e V_1 , em que V_0 é o volume na idade I_0 e

V_1 é o volume na idade I_1 ; por fim as variáveis climáticas (BIO6 e BIO11), edáficas (tipos de solo) e o material genético, como apresentado na tabela 1.

Tabela 1: Valores máximos, mínimos, médios e desvio padrão para as variáveis contínuas e as classes por variável categórica.

Variáveis	Máxima	Mínima	Média	Desvio padrão
Idade (anos)	17,4	4,1	9,2	3,3
Índice de Sítio (m)	30,3	13,0	21,8	2,1
Área basal ($m^2 ha^{-1}$)	100,0	3,8	47,4	17,4
Volume ($m^3 ha^{-1}$)	955,0	4,0	308,5	193,0
Temperatura mínima no mês mais frio ($^{\circ}C$)	9,1	3,8	6,0	10,3
Temperatura média no trimestre mais frio ($^{\circ}C$)	14,4	10,3	12,0	6,9
Declividade (%)	69	1	14	8
Material Genético	M_1, M_2, M_3, M_4			
Tipo de solo	S_1, S_2, S_3, S_4			

2.3. Definição dos testes

Devido ao grande número de combinações possíveis de variáveis de entrada, foram definidos três estudos para projeção de volume, conforme apresentado na tabela 2. No estudo 1, as RNA foram treinadas considerando apenas os dados de inventário como variáveis de entrada (idades, índice de sítio, áreas basais, volumes e material genético), no estudo 2 foram utilizados os dados contínuos de inventário e do ambiente (idades, índice de sítio, áreas basais, volumes, declividade e temperaturas) e por fim, o estudo 3 apresentou todas as variáveis disponíveis como entrada do modelo (idades, índice de sítio, áreas basais, volumes, declividade, temperaturas, material genético e tipo de solo).

Tabela 2: Relação de camadas de entrada e saída por estudo

Estudos	Entrada categórica	Entrada contínua	Saída
1	Material genético	Idade atual e futura Índice de sítio Área basal atual e futura Volume atual	Volume futuro (m^3)
2		Idade atual e futura Índice de sítio Área basal atual e futura Volume atual Declividade Temperaturas WorldClim	Volume futuro (m^3)

3	Material genético Tipo de solo	Idade atual e futura Índice de sítio Área basal atual e futura Volume atual Declividade Temperaturas WorldClim	Volume futuro (m ³)
---	-----------------------------------	-------------------------------------------------------------------------------------------------------------------------------	------------------------------------

Para fins de comparação foi ajustado o modelo de Clutter, utilizando o método de mínimo quadrados ordinários em dois estágios.

2.4. Treinamento das Redes Neurais Artificiais

O treinamento e aplicação das RNA do tipo *Multilayer Perceptron* (MLP) foi feito no *software* RStudio, com a utilização do pacote *neuralnet* (FRITSCH; GÜNTHER, 2008). Para todos os cenários, foram aplicados o algoritmo de treinamento *Resilient propagation* (RPROP+) (Ridmiller and Brawn, 1993) e a função de ativação logística (ou sigmoideal), com intervalo de 0 a 1, sendo necessária a normalização das variáveis de entrada e saída, além da transformação das variáveis qualitativas em sistema binário para enquadrarem na amplitude desejada. A normalização linear foi obtida a partir da equação (1) (SOARES et al., 2011; MARTINS et al., 2015):

$$x' = \frac{(x - x_{min}) * (b - a)}{(x_{max} - x_{min})} + a \quad (1)$$

Em que: x' : valor normalizado; x : valor original; x_{min} : valor mínimo do conjunto; x_{max} : valor máximo do conjunto; a e b : limite inferior e superior do intervalo de normalização.

A estrutura das redes, composta por três camadas – sendo uma de entrada, uma intermediária e uma de saída – teve alteração no número de neurônios de entrada e saída, de acordo com a quantidade de variáveis de entrada nos estudos. A camada de saída não variou porque é o volume futuro desejado. A quantidade de neurônios na camada de entrada é o número de variáveis independentes do banco. Como evidenciado na tabela 3, as variáveis categóricas, ao serem convertidas para sistema binário, somam oito entradas (quatro materiais genéticos e quatro tipos de solo).

Tabela 3: Códigos dos neurônios de entrada utilizados no script por estudo

Estudos	Entradas (código R)	Saída
1	i1, i2, b1, b2, v1, is, mg1, mg2, mg3, mg4	y
2	i1, i2, b1, b2, v1, is, bio6, bio11, d	y
3	i1, i2, b1, b2, v1, is, bio6, bio11, d, mg1, mg2, mg3, mg4	y

O número de neurônios na camada oculta foi calculado pela equação (2):

$$2 \cdot \sqrt{n} + n_2 \leq n_1 \leq 2 * n + 1 \quad (2)$$

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 limite máximo evita *overfitting*, que é a memorização dos dados, e o limite mínimo evita o *underfitting*, número insuficiente de interações para treinar a rede (HAYKIN, 2001; BENTES, 2018). Desta forma, foram definidos 10, 10 e 20 números de neurônios nas únicas camadas ocultas dos estudos 1, 2 e 3, respectivamente, estruturando as redes nos formatos: 10:10:1, 11:10:1 e 17:20:1.

O critério de parada do processo de treinamento utilizado foi o número máximo de 1000 ciclos, ou o erro quadrático médio inferior a 0,01.

A base de dados foi dividida aleatoriamente em dois conjuntos, treinamento e validação, sendo que nesse último conjunto, as redes geradas no treinamento foram aplicadas com a finalidade de avaliar a capacidade da rede em produzir saídas adequadas para entradas desconhecidas (HAYKIN, 2010). Para o presente estudo, foram definidas cinco combinações de proporção de dados, como representado na tabela 4.

Tabela 4: Percentual de treinamento e validação para avaliação das redes neurais artificiais por cenário.

Cenário	Proporção dos dados	
	Treinamento (%)	Teste (%)
1	80	20
2	70	30
3	50	50
4	30	70
5	20	80

Ao final do processamento e aplicação das redes, foi realizada a análise das variações do volume observado e o estimado a partir do gráfico do percentual (E%) (equação 3), o grau de associação entre elas pelo gráfico de correlação linear (r) (equação 4), a análise de exatidão das RNAs através da raiz quadrada do erro quadrático médio (RQEM) (equação 5) e o desvio sistemático do valor real (bias%) (equação 6), descritos nas equações a seguir:

$$E (\%) = \left(\frac{V_{est} - V_{obs}}{V_{obs}} \right) * 100 \quad (3)$$

$$r (\%) = \frac{cov(V_{obs}, V_{est})}{\sqrt{S_{obs}^2 * S_{est}^2}} * 100 \quad (4)$$

$$RQEM (\%) = \sqrt{\sum_{i=1}^n \frac{(V_{obs} - V_{est})^2}{n} * 100} * 100 \quad (5)$$

$$bias (\%) = \frac{100}{n} * \sum_{i=1}^n \left[\frac{V_{obs} - V_{est}}{V_{obs}} \right] \quad (6)$$

Em que: n é o número de observações; V_{obs} é o volume observado em campo; V_{est} é o estimado pela rede; V_{med} é o valor médio do volume observado em campo; cov é a covariância; S^2 é variância.

2.5. Ajuste do modelo de Clutter

A base de dados utilizada na modelagem por RNA também foi utilizada para projeção de volume a partir do modelo de Clutter 1993 (Equações 7 e 8):

$$\ln(V_2) = \beta_0 + \beta_1 * \frac{1}{I_2} + \beta_2 S + \beta_3 \ln(G_2) + \varepsilon \quad (7)$$

$$\ln(G_2) = \ln(G_1) * \left(\frac{1}{I_2} \right) + \alpha_0 * \left(1 - \frac{I_1}{I_2} \right) + \alpha_1 * \left(1 - \frac{I_1}{I_2} \right) * S + \varepsilon \quad (8)$$

Em que: G_1 é área basal na idade I_1 em $m^3 ha^{-1}$; G_2 é área basal na idade I_2 , em $m^3 ha^{-1}$; V_2 é o volume estimado ($m^3 ha^{-1}$); S é o índice de sítio (m); β_i parâmetros de regressão; e ε é erro aleatório.

Foi utilizado o método de mínimos quadrados ordinários em dois estágios para ajuste do modelo.

3. RESULTADOS E DISCUSSÃO

Os resultados das RNAs são apresentados na Tabela 5, com a descrição das RNA e as estatísticas obtidas durante o treinamento e a validação.

Tabela 5: Raiz quadrada do erro quadrático médio, *Bias* e coeficiente linear por cenário de estudo, para os conjuntos de treinamento e validação.

Estudo	Cenário	T (%)	G (%)	Análise	RQEM (%)	Bias (%)	R (%)	
1	1a	80	20	Treinamento	4,79	3,7325	99,5	
	1a			Validação	5,21	4,0606	99,3	
	1b	70	30	Treinamento	5,00	3,9887	99,3	
	1b			Validação	5,08	3,9742	99,3	
	1c	50	50	Treinamento	5,21	3,9212	99,3	
	1c			Validação	4,95	3,7859	99,4	
	1d	30	70	Treinamento	4,94	3,8844	99,4	
	1d			Validação	5,04	3,8371	99,3	
	1e	20	80	Treinamento	4,942	3,884	99,4	
	1e			Validação	40,770	8,463	74,0	
2	2a	80	20	Treinamento	4,997	3,821	99,3	
	2a			Validação	4,600	3,789	99,4	
	2b	70	30	Treinamento	5,112	3,803	99,3	
	2b			Validação	4,872	3,944	99,4	
	2c	50	50	Treinamento	4,691	3,604	99,4	
	2c			Validação	4,769	3,706	99,4	
	2d	30	70	Treinamento	5,080	3,940	99,3	
	2d			Validação	5,096	4,032	99,3	
		Cenário	T (%)	G (%)	Análise	RQEM (%)	Bias (%)	R (%)
		2e	20	80	Treinamento	5,061	4,404	99,3
		2e			Validação	5,402	4,387	99,2
	3	3a	80	20	Treinamento	4,472	3,466	99,5
3a		Validação			5,070	3,887	99,3	
3b		70	30	Treinamento	4,511	3,538	99,5	
3b				Validação	5,034	4,043	99,3	
3c		50	50	Treinamento	4,610	3,746	99,4	
3c				Validação	5,259	4,318	99,3	
3d		30	70	Treinamento	4,288	3,443	99,5	
3d				Validação	5,368	4,216	99,2	

	3e	20	80	Treinamento	4,894	3,911	99,3
	3e			Validação	15,153	6,816	87,9

Inicialmente, dentre as 15 redes treinadas e validadas, 13 apresentaram excelentes resultados de treinamento e validação. Ou seja, valores de RQEM, em geral, abaixo de 10%, variando de 4,23 a 5,40; viés abaixo de 10%, entre 3,44 a 4,40; e correlação linear acima de 99%. As duas redes com pior ajuste são 1e e 3e, que tiveram bons ajustes na fase de treinamento, porém na validação das estatísticas de exatidão foram acima de 20%. Ambos os cenários podem indicar o fenômeno de *underfitting*. Estas redes de pior desempenho evidenciam a importância da validação, nas estimativas volumétricas.

De forma a avaliar a presença de dados discrepantes e possível tendenciosidade nas estimativas, foram gerados gráficos de dispersão dos erros percentuais, os resultados são compilados nas figuras a seguir. A figura 1 é referente ao domínio 1.

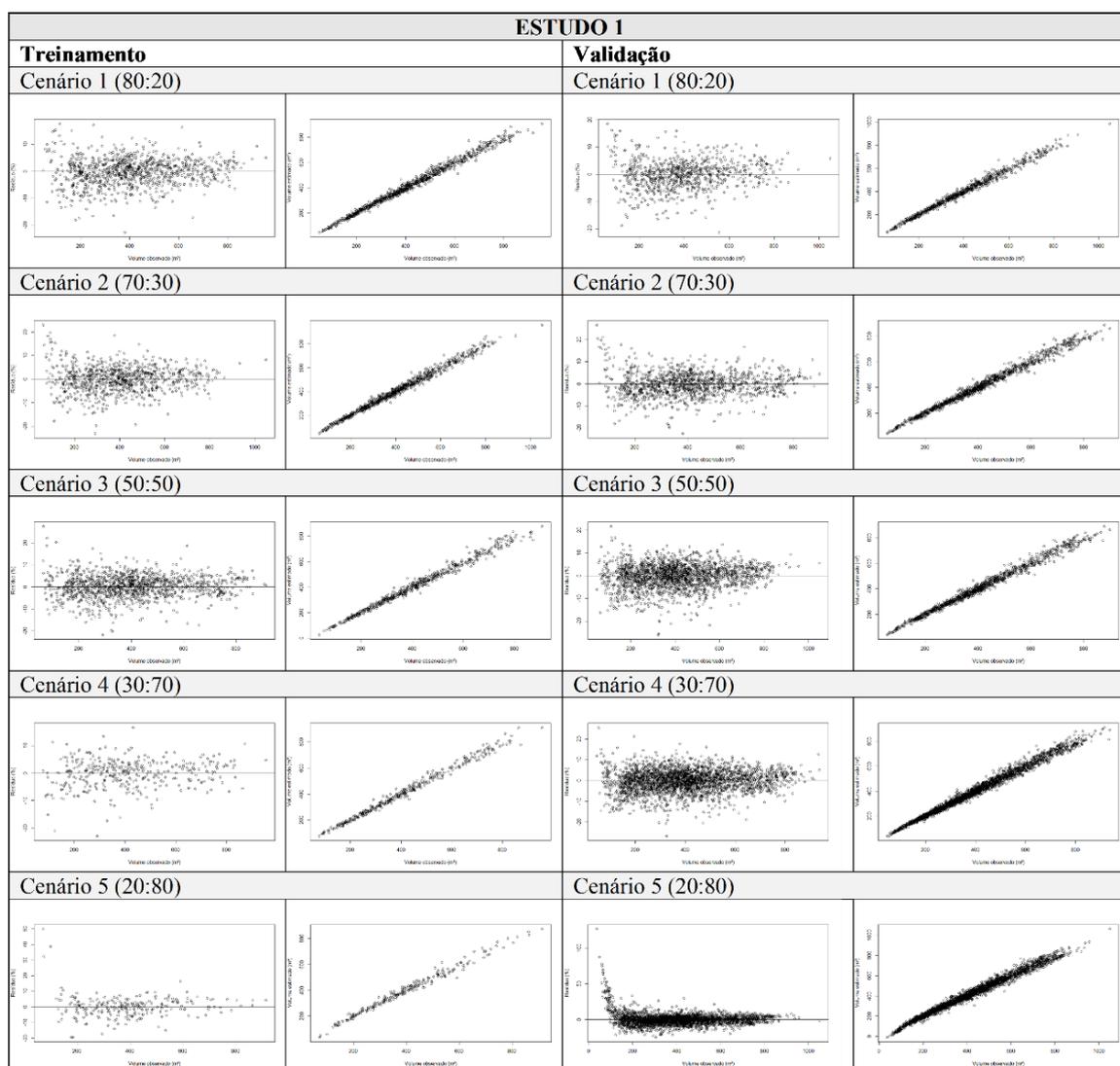


Figura 1: gráficos de valores observados em campo *versus* estimado pelas redes e gráfico de dispersão do erro percentual para os cenários do estudo 1.

Pela análise gráfica do estudo 1, é possível verificar que os cenários 1, 2, 3 e 4 não apresentam tendenciosidade nas estimativas. Porém, o cenário 5 apresentou tendência, devido à baixa amplitude da base de entrada para o treinamento da rede. Todos os cenários apresentaram correlação alta.

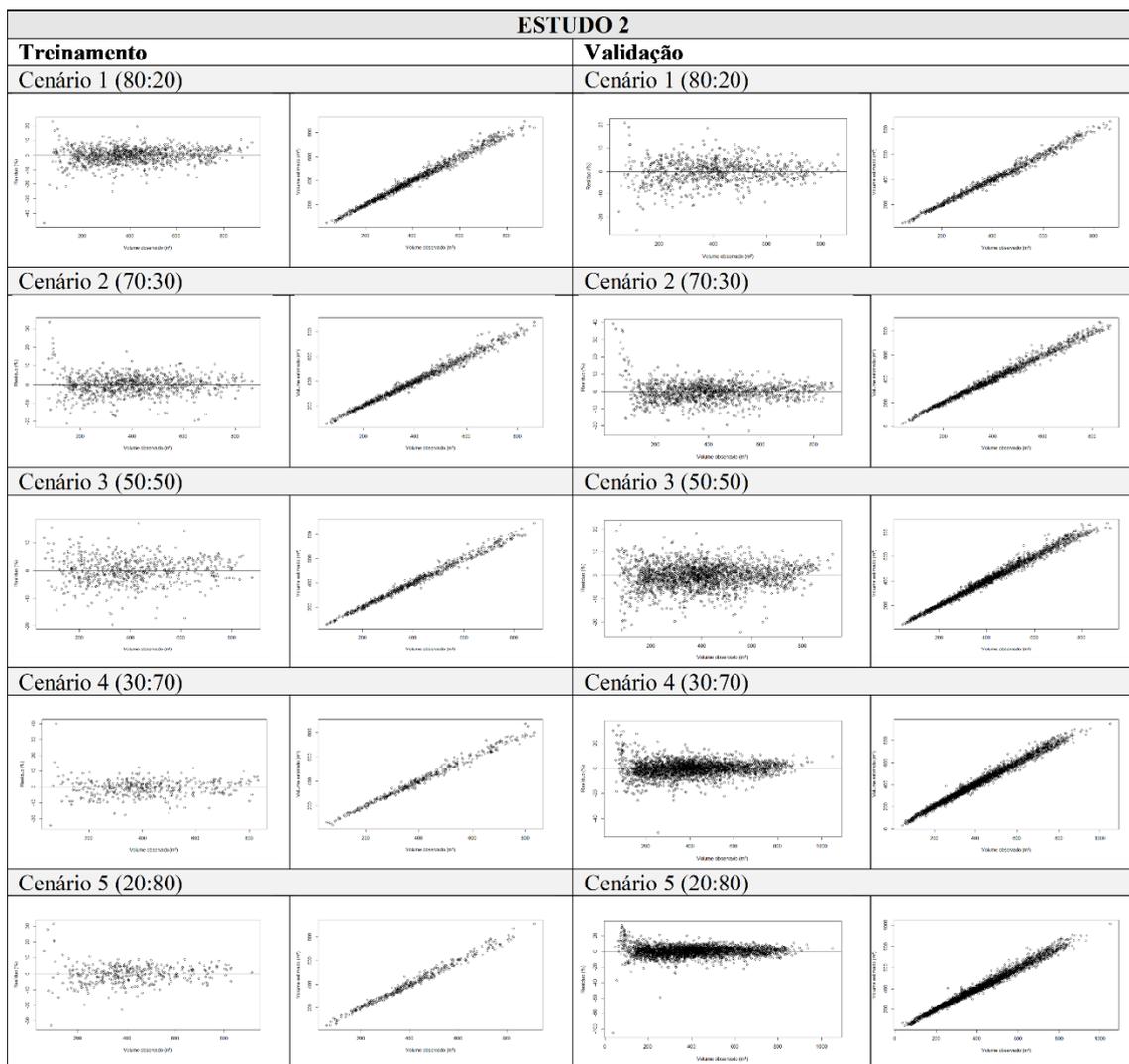


Figura 2: gráficos de valores observados em campo *versus* estimado pelas redes e gráfico de dispersão do erro percentual para os cenários do estudo 2.

Assim como no primeiro cenário, o estudo 2 não apresentou tendenciosidade nas estimativas volumétricas e apresentou alta correlação para todos os testes realizados.

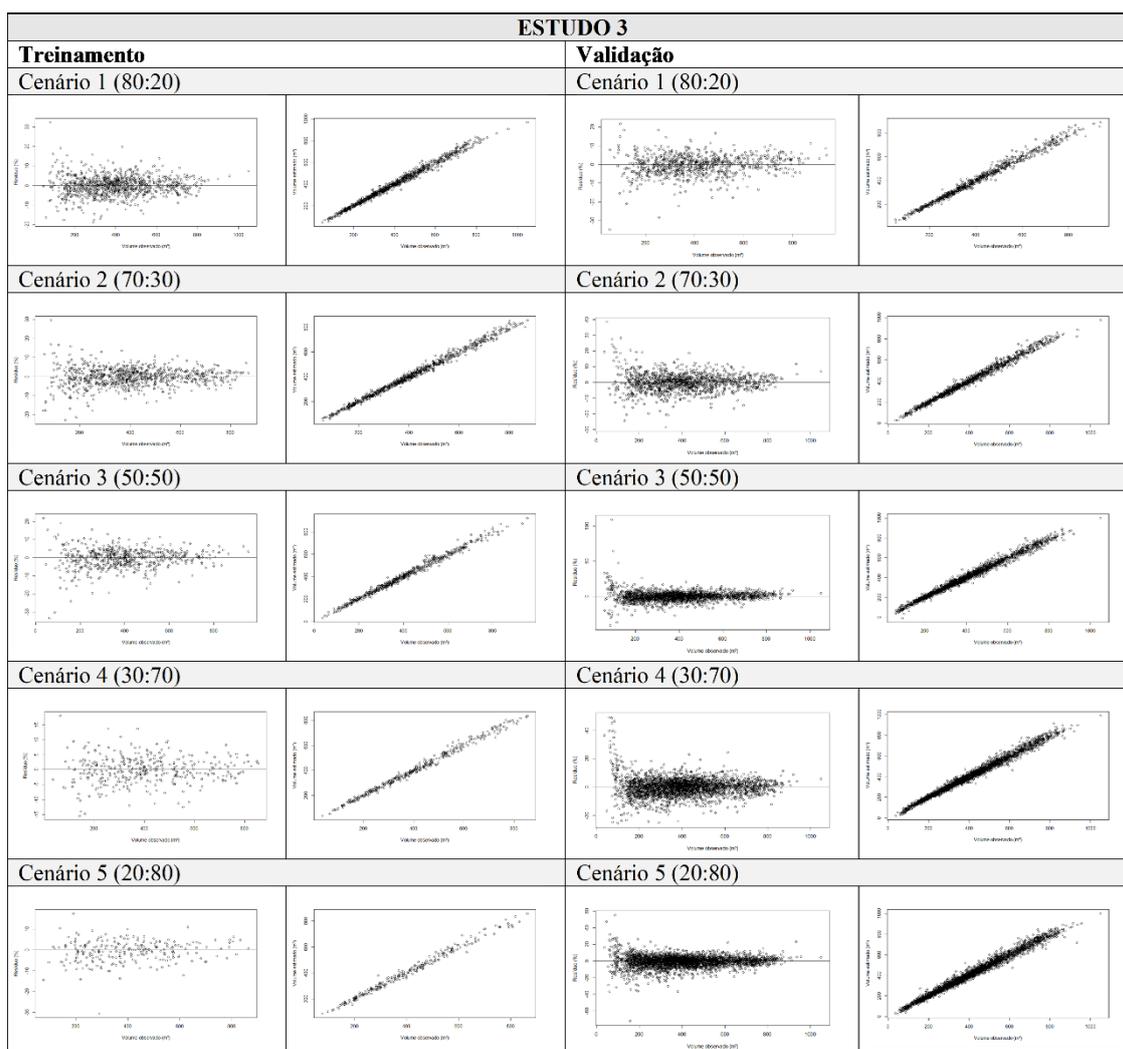


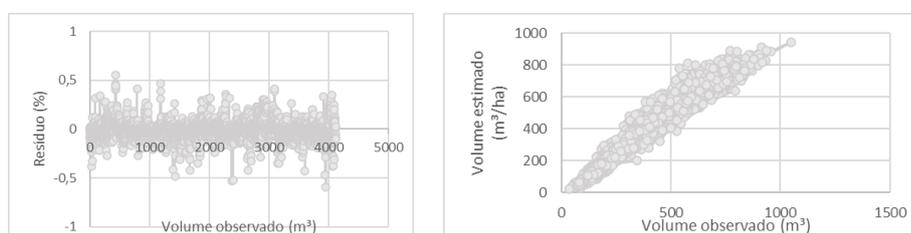
Figura 3: gráficos de valores observados em campo *versus* estimados pelas redes e gráfico de dispersão do erro percentual para os cenários do estudo 3.

Pela análise gráfica do estudo 3, é possível verificar que os cenários 1, 2, 3 e 5 estão sem tendenciosidades no gráfico de dispersão do erro, com os valores bem distribuídos ao redor do eixo x. O cenário 4 apresenta um pouco de tendência na validação. Todos os cenários apresentaram correlação alta.

O modelo escolhido como o mais preciso possui a melhor combinação de desempenho no treinamento e na validação, é a rede 2c, com RQEM de 4,69% para treinamento e 4,77% para validação. Ela possui como entrada, as variáveis contínuas do IFC, temperaturas e declividade. A partir desta análise, é possível afirmar que a inserção de variáveis ambientais na RNA pode reduzir a base de dados necessária de modo significativo. Por exemplo, o cenário 1a, com entradas dendrométricas e material genético, possui 80% dos dados de entrada para treinamento, e mesmo assim não

consegue alcançar os mesmos erros da rede 2c, mesmo que esta possua apenas 50% de dados para treinamento.

O método de Clutter utilizado para estimar o volume futuro, foi ajustado utilizando a base de dados contínuos a nível de povoamento, conforme equação 8. Foram definidos os coeficientes e logo em seguida foi realizada a análise estatística: RQEM foi de 10,22%; bias de 6,068% e coeficiente de correlação de 97%. Também foram gerados os gráficos de resíduo (%) e correlação (figuras 4 e 5). Mesmo o modelo alcançando bons resultados e apresentando erro percentual sem tendência, a rede neural superou este método tradicional, por apresentar menor erro e maior precisão para prognose de volume em povoamentos de *Pinus taeda*.



Figuras 4 e 5: dispersão do erro percentual (%) e a correlação linear entre o volume observado e estimado, a partir do método de Clutter.

4. CONCLUSÃO

A utilização de redes neurais artificiais para prognose de produção em povoamentos de *Pinus taeda* neste trabalho confirmam a aplicabilidade prática das RNAs em comparação ao modelo tradicional de regressão. Das 15 redes analisadas, 13 apresentaram maior exatidão de estimativa em relação ao modelo de Clutter. Houve relevância positiva nos treinamentos e na validação das redes com a inserção de variáveis edafoclimáticas. O investimento de empresas florestais em ferramentas de Inteligência Artificial pode melhorar as estimativas do estoque futuro e, conseqüentemente, ter impacto positivo na tomada de decisões e no planejamento estratégico. É possível reduzir a intensidade amostral utilizada nos Inventários Florestais Contínuos, com a inserção de variáveis categóricas, sem a perda de exatidão, o que impacta positivamente nos custos associados à essa atividade, essencial no manejo dos povoamentos de *Pinus Taeda*.

REFERÊNCIAS

- ALCÂNTARA, A. E. M. de. **Redes neurais artificiais para prognose do crescimento e da produção de povoamentos de eucalipto em Minas Gerais**. 2015. 58 p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal. Universidade Federal de Viçosa (UFV), Viçosa, 2015.
- BENTES, B. A. H. F. **Aplicação de redes neurais artificiais para estimativa de volume de árvores em pé em uma floresta ombrófila densa na Amazônia Oriental**. 2018. 59 p. Dissertação (mestrado) – Programa de Pós-Graduação em Ciência Florestal. Universidade Federal Rural da Amazônia (UFRA), Pará, 2018.
- BULLINARIA, J. A. **Introduction to neural computation**: notas de aula. 2014. Disponível em: <<http://www.cs.bham.ac.uk/~jxb/inc.html>>. 2014.
- BERGER, R.; SCHNEIDER, P. R.; FINGER, C. A. G.; HASELEIN, C. R. EFEITO DO ESPAÇAMENTO E DA ADUBAÇÃO NO CRESCIMENTO DE UM CLONE DE *Eucalyptus saligna* Smith. *Ciência Florestal*, Santa Maria, v. 12, n. 2, p. 75-87, 2002.
- CASAS, G. G. **Modelagem do crescimento e da produção de povoamentos de eucalipto utilizando regressão, redes neurais artificiais e redes profundas**. 2021. 101 p. Dissertação (Mestrado) - Programa de Pós-Graduação em Ciência Florestal (UFV), Viçosa, Brasil.
- CLUTTER, J.L. Compatible growth and yield models for loblolly pine. *Forest Science*, v.9, p.354-371, 1963
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal**: perguntas e respostas. Viçosa: Editora UFV, 2002. 408 p.
- CORDEIRO, M. A. **Modelagem do crescimento e produção em povoamento de eucalipto utilizando máquina de vetor de suporte e redes neurais artificiais**. 2020. 123p. Dissertação (Mestrado) – Programa de Pós Graduação em Ciência Florestal, Universidade Estadual do Centro-Oeste, IRATI, PR, 2020.
- FICK, S.E.; HIJMANS, R. J. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. [International Journal of Climatology, 2017.](#)
- GORGENS, E. B. **Estimação do volume de árvores utilizando redes neurais artificiais**. 2006. 94p. Dissertação (Mestrado) – Programa de Pós Graduação em Ciência Florestal. Universidade Federal de Viçosa (UFV), Viçosa, 2006.
- GÜNTHER, F.; FRITSCH, F. NeuralNet: Training of Neural Networks. *The R Journal*, 2, 30-38. 2010.

HAYKIN, S. **Redes neurais: princípios e prática**. Trad. de Paulo Martins Engel. 2. ed. Porto Alegre: Bookman, 2001. 900 p.

KLINA, L. **Construção de curvas de índice de sítio e determinação do volume populacional para *Eucalyptus grandis* W Hill ex Maiden no sudoeste do Pará**. 49p. TCC (Graduação), Universidade Tecnológica Federal do Paraná-UTFPR, 2015.

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.

MARTINS, E. dos. R.; BINOTI, M. L. M. da. S.; LEITE, H. G.; BINOTI, D. H. B.; DUTRA, G. C. Configuração de redes neurais artificiais para prognose da produção de povoamentos clonais de eucalipto. **Agrária - Revista Brasileira de Ciências Agrárias**, Recife, v. 10, n .4, p. 532-537, 2015.

OLIVEIRA, L. R. de.; LEITE, H. G.; NOGUEIRA, S. G.; CAMPOS, J. C. C. C. Modelagem e prognose em povoamentos não desbastados de clones de eucalipto. **R. Árvore**, Viçosa-MG, v.33, n.5, p.841-852, 2009.

ÖZÇELİK, R.; DIAMANTOPOULOU, M. J.; CRECENTE-CAMPO, F.; ELER, U. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology and Management**, v. 306, p. 52–60, 2013.

PENIDO, T. M. A.; LAFETÁ, B. O.; NOGUEIRA, G. S.; ALVES, P. H, Gorgens, E. B., & Oliveira, M. L. R. (2020). Modelos de crescimento e produção para a estimativa volumétrica em povoamentos comerciais de eucalipto. **Scientia Forestalis**, 48(128), 2020.

SDS - Secretaria de Estado do Desenvolvimento Econômico Sustentável de Santa Catarina. Levantamento Aerofotogramétrico 2010 - **Modelo Digital de Terreno**. Disponível em: <http://sigsc.sc.gov.br/>. 2010.

ZONETE, M. F.; RODRIGUEZ, L. C. E.; PACKALÉN, P. Estimación de parâmetros biométricos de plantios clonais de eucalipto no sul da Bahia: uma aplicação da tecnologia laser aerotransportada. **Sci. For.**, Piracicaba, v. 38, n. 86, p. 225-235, jun. 2010.

TERCEIRA PARTE

CONSIDERAÇÕES FINAIS DESTE TRABALHO DE CONCLUSÃO DE CURSO E SUGESTÕES PARA FUTUROS TRABALHOS

- Criar uma segunda RNA utilizando apenas variáveis edafoclimáticas com a mesma base de dados e comparar com a rede criada neste trabalho, para avaliar a possibilidade de prognose em locais sem base de dados de inventário.
- Gerar um mapa de produtividade para o estado de Santa Catarina, como o estudo de ALCÂNTARA (2015) para o estado de Minas Gerais. Porém, serão necessários mais dados de inventário de outras empresas para treinar a rede para outras variáveis categóricas que não foram contempladas na base de dados de apenas uma empresa florestal.