

DÉBORA EVELYN CHRISTO DOS SANTOS

**REGRESSÃO E *DEEP LEARNING* APLICADOS EM MODELAGEM DE ÁRVORE
INDIVIDUAL PARA EUCALIPTO**

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

Orientador: Helio Garcia Leite

**VIÇOSA – MINAS GERAIS
2021**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

S237r
2021

Santos, Débora Evelyn Christo dos, 1995-
Regressão e *deep learning* aplicados a modelagem em nível
de árvore individual para plantios de eucalipto [recurso
eletrônico] / Débora Evelyn Christo dos Santos. – Viçosa, MG,
2021.

1 dissertação eletrônica (43 f.): il. (algumas color.).

Orientador: Helio Garcia Leite.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 39-43.

DOI: <https://doi.org/10.47328/ufvbbt.2021.033>

Modo de acesso: World Wide Web.

1. Florestas - Manejo. 2. Aprendizado profundo. 3. Redes
neurais (Computação). I. Universidade Federal de Viçosa.
Departamento de Engenharia Florestal. Programa de
Pós-Graduação em Ciência Florestal. II. Título.

CDO adapt. CDD 22. ed. 634.96242

DÉBORA EVELYN CHRISTO DOS SANTOS

**REGRESSÃO E *DEEP LEARNING* APLICADOS EM MODELAGEM DE ÁRVORE
INDIVIDUAL PARA EUCALIPTO**

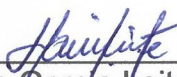
Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência Florestal, para obtenção do título de *Magister Scientiae*.

APROVADA: 30 de junho de 2021.

Assentimento:



Débora Evelyn Christo dos Santos
Autora



Helio Garcia Leite
Orientador

AGRADECIMENTOS

A Deus, pela vida e por todas as bênçãos.

À minha Família, em especial aos meus pais, Janete e Vanderci, pelo amor, incentivo e apoio incondicional.

Ao meu Orientador Helio Garcia Leite agradeço pela paciência, motivação, orientação e por todos ensinamentos valiosos.

Ao Professor Carlos Alberto Araújo Júnior por todo aprendizado transmitido e ajuda.

Ao Leonardo Fardin e Gianmarco por toda contribuição, incentivo, disposição e paciência. Aos amigos da pós que sempre se fizeram presentes nesta jornada.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos.

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

E por fim, fico muito grata a todos aqueles que de alguma forma contribuíram para construção desse trabalho e para conquista de mais uma etapa em minha vida profissional!

Muito obrigada!

RESUMO

SANTOS, Débora Evelyn Christo, M.Sc., Universidade Federal de Viçosa, junho de 2021. **Regressão e *Deep Learning* aplicados em modelagem de árvore individual para eucalipto.** Orientador: Helio Garcia Leite.

Neste estudo avaliamos a eficiência do emprego de redes neurais, convencionais e profundas (deep learning), em relação ao emprego de regressão, para projetar diâmetro (dap), altura total (Ht) e sobrevivência (N) em povoamentos de eucalipto. O estudo foi conduzido com dados de parcelas permanentes de inventários florestais contínuos de povoamentos de eucalipto localizados no estado de Minas Gerais. Para o processamento dos dados com uso de regressões e redes neurais convencionais utilizamos o software *Statistica 12* e para redes profundas o software *R*. Os dados foram preliminarmente separados para treinamento (50%) e validação (50%) utilizando um algoritmo genético implementado no *software Select2.0*. A eficiência relativa dos dois métodos de aprendizado de máquinas, em relação às regressões, foi feita com base nas estatísticas: coeficiente de correlação linear entre valores observados e correspondentes valores estimados ou preditos ($r_{\hat{y}y}$), viés (bias) e raiz quadrada do erro quadrático médio (RQEM). Foi utilizada, de modo complementar, a análise gráfica dos erros relativos (ER%), a análise da relação entre valores observados e estimados e a distribuição de frequência de ER%. Os resultados das modelagens foram eficientes, as Rna convencional e profunda apresentaram melhores desempenhos estatísticos comparada a regressão para estimativa de dap_2 , Ht_2 e N_2 . Concluiu-se que os métodos de aprendizagem de máquina utilizados são eficientes e aumentam a exatidão das estimativas de diâmetro e altura de árvores na modelagem de árvore individual.

Palavras-chave: Aprendizagem profunda. Manejo Florestal. Redes neurais artificiais.

ABSTRACT

SANTOS, Débora Evelyn Christo, M.Sc., Universidade Federal de Viçosa, June, 2021.
Regression and Deep Learning applied in single-tree modeling for eucalyptus.
Advisor: Helio Garcia Leite.

We calculate the efficiency of using conventional and deep learning neural networks in relation to the use of regression to project diameter (DBH), total height (TH), and survival (S) in eucalyptus stands. The data were from permanent plots of continuous forest inventories of eucalyptus stands located in Minas Gerais. We used the software *Statistica 12* for data processing using regressions and conventional neural networks and, R software for deep networks. The data were preliminarily separated for training (50%) and validation (50%) using a genetic algorithm executed in the software *Select2.0*. We found the relative efficiency of two machine learning methods, in relation to regressions, based on the statistics: linear correlation coefficient, bias, and mean square root error (MSRE). The graphical analysis of relative errors (RE%), the analysis of the comparing observed values with the expect values, and the frequency distribution of RE% were used in a complement way. The results of the modeling were efficient, conventional and deep ANN showed better statistical performances compared to regression for estimation of DBH_2 , TH_2 and S_2 . We concluded that the machine learning methods used could themselves be a potential alternative for estimating dendrometric variables presenting efficiency and increased accuracy in single-tree modeling.

Keywords: Deep learning. Forest management. Artificial neural networks.

LISTA DE ILUSTRAÇÕES

Figura 1 - Estrutura da RNA.....	16
Figura 2 - Diferenças de rede neural convencional e rede neural profunda de acordo com as camadas ocultas.....	17
Figura 3 - Mapa de localização dos municípios da área de estudo	18
Figura 4 - Curvas de índices de sítio para povoamento de clones híbridos de <i>Eucalyptus urophylla</i> X <i>Eucalyptus grandis</i> localizados no estado de Minas Gerais com idade-índice de 72 meses	25
Figura 5 - Análise gráfica dos resíduos dos modelos de regressão para variável dap_2	28
Figura 6 - Análise gráfica dos resíduos dos modelos de regressão para variável Ht_2	30
Figura 7 - Análise gráfica dos resíduos do modelo de regressão Pienaar e Shiver (1981) para variável N_2	31
Figura 8 - Análise gráfica das RNA convencional e profunda para variável dap_2	33
Figura 9 - Análise gráfica das RNA convencional e profunda para variável Ht_2	33
Figura 10 - Análise gráfica das RNA convencional e profunda para variável N_2	34
Figura 11 -Faixa de dados discrepantes na modelagem de árvore individual	35
Figura 12 - Erro relativo em função das idades (meses) para projeção de Dap_2 das árvores individuais	36
Figura 13 - Erro relativo em função das idades (meses) para projeção de Ht_2 das árvores individuais	36
Figura 14 - Erro relativo em função das idades (meses) para projeção de N_2	37

LISTA DE TABELAS

Tabela 1 - Modelos de regressão ajustados para diâmetro e altura total	22
Tabela 2 - Variáveis utilizadas no treinamento de redes neurais artificiais para predição de diâmetro, altura e sobrevivência de árvores	23
Tabela 3 - Curvas de índices de sítio para clones e híbridos de <i>Eucalyptus urophylla</i> X <i>Eucalyptus grandis</i>	26
Tabela 4 - Estimativas dos parâmetros dos modelos de regressão ajustados para estimar o diâmetro futuro (dap_2) de árvores de eucalipto na área de estudo e correspondentes estatísticas de precisão e de exatidão	27
Tabela 5 - Estimativas dos parâmetros dos modelos de regressão ajustados para estimar a altura total futura (Ht_2) de árvores de eucalipto na área de estudo e correspondentes estatísticas de precisão e de exatidão	29
Tabela 6 - Parâmetros do modelo de regressão para variável N_2	31
Tabela 7 - Parâmetros das RNA convencional e profunda para variáveis dap_2 e Ht_2	32
Tabela 8 - Parâmetros das RNA convencional e profunda para variáveis N_2	34

SUMÁRIO

1. INTRODUÇÃO	9
2. OBJETIVOS	11
3. REVISÃO BIBLIOGRÁFICA	12
4. MATERIAL E MÉTODOS	18
4.1. DADOS	18
4.2. CONSTRUÇÃO DE CURVAS DE ÍNDICE DE SÍTIO E PROJEÇÃO DA SOBREVIVÊNCIA ...	20
4.3. MODELAGEM DO CRESCIMENTO EM DIÂMETRO E ALTURA	21
4.4. TREINAMENTO DE REDES NEURAS ARTIFICIAIS.....	23
4.4.1. Redes neurais convencionais	23
4.4.2. Redes neurais profundas (deep learning)	23
4.5. VALIDAÇÃO	24
5. RESULTADOS E DISCUSSÃO	25
5.1. CLASSIFICAÇÃO DA CAPACIDADE PRODUTIVA	25
5.2. MODELOS DE REGRESSÃO	26
5.3. REDES NEURAS ARTIFICIAIS E REDES NEURAS PROFUNDAS	32
5.4. ANÁLISE DOS RESÍDUOS	35
6. CONCLUSÕES	38
REFERÊNCIAS	39

1. INTRODUÇÃO

Os plantios do gênero *Eucalyptus* ocupam uma área de aproximadamente 7,6 milhões de hectares no Brasil, sendo o gênero florestal mais cultivado no país. A produção madeireira é destinada à produção de celulose, papel, madeira serrada, painéis de madeira, pisos laminados, carvão vegetal e biomassa (IBGE, 2019).

A administração e o manejo das florestas equiâneas para o atendimento das metas anuais de produção levam consigo o aporte constante das atividades de pesquisa e planejamento, buscando eficiência e adaptabilidade às condições do solo e do clima para atingir maior produtividade. O planejamento florestal permite conduzir ações durante o ciclo de corte do povoamento, visando maximizar o retorno econômico do investimento florestal, com garantias de sustentabilidade (VEGA, 2018).

Para o manejo adequado da floresta é necessário o conhecimento de seu estoque em crescimento e da colheita. Dentre as técnicas relacionadas ao manejo florestal, a modelagem do crescimento e da produção de plantios comerciais de eucalipto no Brasil apresenta-se bem aplicada, e com avanços significativos nesta área (MARTINS, 2011). A modelagem pode ser de crescimento e produção em nível de povoamento, árvore individual e modelagem de distribuição de diâmetros (CASTRO, 2011; MARTINS, 2011).

A modelagem de crescimento e produção em nível de árvore individual, gera resultados detalhados das estruturas horizontal e vertical e pode ser utilizada para simular tratamentos silviculturais como desbaste e a desrama artificial, além disso, ao somar os volumes de árvores individuais é obtido o volume por hectare (CAMPOS; LEITE, 2017). Um modelo de árvore individual (MAI) é constituído por sistema de equações e procedimentos para projeção (ou predição) de diâmetro e altura de árvores individuais, além de um componente de mortalidade, ou sobrevivência (MARTINS, 2011).

Buscando alternativas ao procedimento tradicional de modelagem em nível de árvore individual, realizado por meio dos sistemas de equações, mais recentemente tem sido empregado modelos de inteligência artificial, como redes neurais artificiais e máquinas vetor de suporte (SILVA et al., 2009; CASTRO, 2011; MAGALHÃES, 2014; VEGA, 2018). Esses métodos de aprendizado de máquinas podem ser utilizados para prognose de produção volumétrica em nível de povoamento, classe de diâmetro (BINOTI et al., 2013; BINOTI, 2010; LEDUC, et al., 2001; VEGA, 2018) e árvores

individuais (CASTRO, 2011). Por serem métodos relativamente novos é sempre importante verificar a eficiência do uso dos mesmos em diferentes condições e tipos de povoamentos.

No presente estudo avaliamos a eficiência do uso de redes neurais artificiais convencionais e redes profundas (deep learning), em relação ao emprego de regressão, para projetar diâmetro e altura de árvores em povoamentos de eucalipto. Foi objetivo, também, avaliar a exatidão de estimativas de sobrevivência obtidas com as duas técnicas de aprendizado de máquinas.

2. OBJETIVOS

Este estudo teve por objetivos:

- a) Análise de alternativas de modelagem em nível de árvore individual para povoamentos de eucalipto utilizando regressão;
- b) Modelagem do crescimento em nível de árvore individual empregando regressão, redes neurais artificiais e redes neurais profundas;
- c) Avaliação da eficiência do uso de dois métodos de aprendizado de máquinas na predição de diâmetro e altura de árvores individuais, em relação ao emprego de regressão;
- d) Avaliar a exatidão de sobrevivências preditas por redes neurais em povoamentos de eucalipto.

3. REVISÃO BIBLIOGRÁFICA

Os plantios do gênero eucalipto ocupam 7,6 milhões de hectares do total de áreas florestadas, enquanto as áreas com pinus somam dois milhões de hectares, e outras espécies, entre elas seringueira, acácia, teca e paricá, representam cerca de 387 mil hectares (IBGE, 2019). Os plantios estão localizados principalmente nos Estados de Minas Gerais, São Paulo e Mato Grosso do Sul e a madeira é destinada principalmente para produção de celulose carvão vegetal (IBÁ, 2020).

O inventário florestal permite conhecer informações quantitativas e qualitativas de povoamentos florestais gerando informações para uso direto ou indireto em planejamento e manejo florestal para que assim se possa decidir sobre aproveitamento, densidade de desbaste, produtividade anual e intervenções nos povoamentos (CAMPOS; LEITE, 2017).

A estimação de crescimento e produção é feita usualmente com emprego de modelos estatísticos. Estes modelos são agrupados em modelos em nível de povoamento, de distribuição de diâmetros e de árvores individuais (DAVIS; JOHSON, 1987; VEGA, 2018).

Os modelos em nível de árvore individual são sistemas de equações e procedimentos utilizados para estimar crescimento e produção de povoamentos tendo como unidades de modelagem as árvores, em vez de parcela (CAMPOS; LEITE, 2017; AVERY; BURKHART, 1994). Esse tipo de modelagem é importante no planejamento florestal, pois fornecem informações detalhadas sobre a dinâmica e a estrutura horizontal e vertical do povoamento, incluindo a distribuição de volume por classe de tamanho (CLUTTER et al., 1983), além de diferentes enfoques para expressar a competição (CAMPOS; LEITE, 2017).

Esses sistemas constituídos permitem gerar equações para projeção de diâmetro, de uma função de mortalidade e de uma função de crescimento para altura, ou de uma relação hipsométrica; o volume por hectare é obtido ao somar volumes de árvores individuais (HASENAUER, 2000; CAMPOS; LEITE, 2017). Diferente dos modelos em nível de povoamento, para a construção dos modelos de árvore individual há necessidade de dados específicos do tamanho de cada árvore dentro do povoamento e, em alguns casos, da distribuição espacial de cada árvore e das classes de altura e copa em que a árvore se encontra (CLUTTER et al., 1983; VANCLAY, 1995; CAMPOS; LEITE, 2017).

Um sistema de equação usado para estimar crescimento em nível de árvore individual é composto por três submodelos, sendo eles: mortalidade, regeneração ou ingresso e crescimento dimensional (GUAN; GERTNER, 1991; HAUSENAUER, 2000).

Dentre os submodelos do sistema, a mortalidade é o mais difícil de se modelar devido a dificuldade de prever (ROSSI et al., 2007; MARTINS, 2011; CASTRO et al., 2013). Existem dois tipos de mortalidade: a regular e a irregular. A ocorrência do tipo de mortalidade irregular é devido ações antrópicas, isto é, causadas por pragas, incêndios, danos mecânicos, catástrofes ambientais entre outras causas que são difíceis de prever. O tipo regular é em decorrência de fatores não catastróficos como competição, fatores genéticos e senescência, a qual pode ser mensurada (LEE, 1971; PEET; CHRISTENSEN, 1987; CAMPOS; LEITE, 2017; ROSSI et al., 2007; CASTRO, 2011).

Para estimar a mortalidade do tipo regular, são utilizados os índices de competição. Os modelos utilizados podem ser classificados de acordo com o índice de competição os quais são denominados: modelo de distância independente e modelo de distância dependente (CAMPOS; LEITE, 2017; GLOVER; HOOL, 1979).

O modelo de distância independente não considera a distância e o porte das árvores vizinhas, este modelo é composto de uma função de crescimento em diâmetro, uma função para estimar crescimento em altura e um componente de mortalidade. No modelo de distância dependente além dos fatores considerados no modelo anterior é considerada a localização das árvores vizinhas nos índices de competição (CAMPOS; LEITE, 2017).

Ainda sobre o submodelo de mortalidade, de acordo com Campo e Leite (2017) mortalidade e sobrevivência de árvores são variáveis complementares, ou seja, obter o conhecimento de uma induz ao conhecimento da outra. Sendo assim, a modelagem para estimar a sobrevivência futura apresenta-se como alternativa para explicar a mortalidade regular das árvores dentro de um povoamento.

O submodelo regeneração ou ingresso refere-se ao recrutamento ou ingresso, ou seja, os indivíduos medidos em uma idade qualquer que não foram medidos em idades anteriores, por não alcançarem um tamanho mínimo predeterminado (HAUSENAUER, 2000; CAMPOS; LEITE, 2017; MARTINS, 2011). Em florestas

plantadas, normalmente desconsidera-se este submodelo tendo em vista o diâmetro mínimo de medição e a idade inicial de medição (MARTINS, 2011).

E o último submodelo é o que modela crescimento e produção. Esse submodelo pode variar no modo de simular, mas geralmente contém três componentes primários de crescimento: crescimento em diâmetro, crescimento em altura e crescimento da copa. Quando incorporado no modelo crescimento e produção uma função de mortalidade podem ser simuladas produções e estruturas futuras do povoamento florestal (CAMPOS; LEITE, 2017).

De forma individual as árvores de um povoamento florestal são modeladas para estimar crescimento em diâmetro e altura e o tamanho da copa, além de sobrevivência (DAVIS; JOHNSON, 1987 citado por CAMPOS; LEITE, 2017). São exemplos de sistemas de árvore individual: Ptaeda (DANIELS; BURHART, 1975), Prognosis (STAGE, 1973; WYKOFF, 1990) e Stems (BELCHER et al., 1982). Uma das aplicações desses sistemas é na simulação de desbastes, conforme mostrado em Wykoff et al. (1982), para o Prognosis.

Além dos sistemas de equações, mais recentemente tem sido empregado modelos de inteligência artificial, como redes neurais artificiais e máquinas vetor de suporte (SILVA et al., 2009; CASTRO, 2011; MAGALHÃES, 2014; VEGA, 2018). As redes neurais artificiais são capazes de realizar prognose de produção volumétrica em nível de povoamento, classe de diâmetro (BINOTI et al., 2013; LEDUC, et al., 2001; VEGA, 2018) e árvores individuais (CASTRO, 2011) sendo uma técnica eficiente para modelagem, prognose e outras aplicações em manejo florestal (BINOTI, 2010).

Redes neurais artificiais (RNA) são sistemas computacionais paralelos constituídos de várias unidades de processamento simples, conectadas entre si, para calcular determinadas funções matemáticas, sendo este modelo computacional inspirado no sistema nervoso de seres vivos (HAYKIN, 2001; SILVA et al. 2010).

De acordo com Silva et al. (2009) na área florestal estudos tem indicado a eficiência da aplicação das redes neurais artificiais para descrever comportamento de variáveis dendrométricas em comparação aos modelos de regressão, possibilitando redução de custos e aumento da exatidão.

Conforme Binoti (2012), os estudos na área de RNA se justificam por diversas propriedades úteis e capacidade como: tolerância a falhas e ruídos, aprendizagem e

generalização a partir de dados de treinamento, paralelismo massivo que as tornam muito eficientes, uniformidade de análise e projetos, analogia neurobiológica, mapeamento entrada-saída e não linearidade.

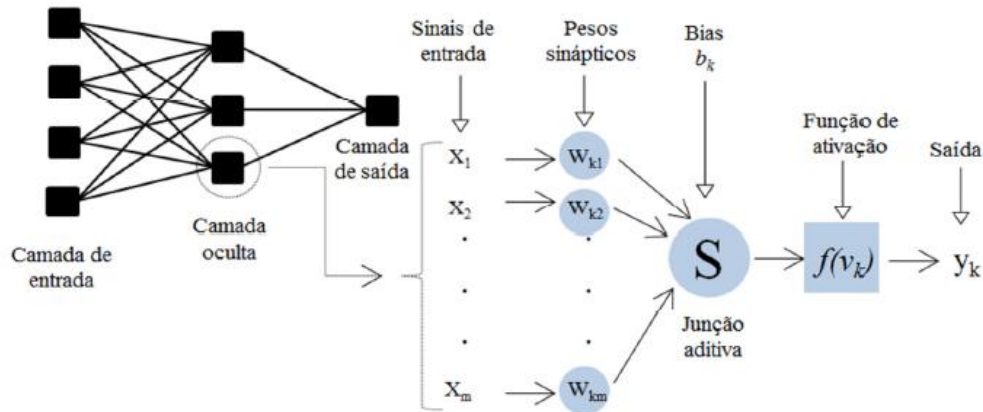
A RNA também trata problemas em sua aprendizagem como: classificação de padrões, aproximação de função, agrupamento de dados, otimização, recuperação de conteúdo, predição (séries temporais) e controle (BINOTI, 2012), sendo uma área de pesquisa multidisciplinar (VENDRUSCOLO et al, 2015).

A utilização de RNA na ciência florestal brasileira é recente, no entanto, tem sido cada vez mais frequente a utilização dessa técnica para estimação de diversos parâmetros florestais. Muitas vezes os resultados tem sido melhores do que aqueles obtidos pelo emprego de regressão (VENDRUSCOLO et al, 2015).

Diversos estudos tem reportado a eficiência de RNA para situações como: estimação de volume, altura, afilamento, crescimento e produção (VENDRUSCOLO et al, 2015). Uma das grandes vantagens do uso de RNA, é que elas possibilitam a utilização de variáveis dendrométricas, edáficas, climáticas, regiões, entre outras (BINOTI, 2010; VENDRUSCOLO et al, 2015).

Segundo Haykin (2001), uma RNA pode ser constituída por uma ou mais camadas. Cada camada pode conter um ou mais neurônios (unidades de processamento simples). Silva et al. (2010) relatam que a RNA pode ser dividida em 3 camadas: camada de entrada, camada oculta e camada de saída. A camada de entrada recebe as informações (dados) advindas do meio externo, nesta mesma camada os dados geralmente são normalizados, para uma melhor precisão numérica frente as operações matemáticas realizadas pela rede. A camada oculta, é constituída de neurônios que são responsáveis por extrair as características associadas ao sistema a ser inferido. E por fim, a camada de saída, a qual também é composta por neurônios que possuem a função de produzir e apresentar os resultados finais da rede (Figura 1).

Figura 1 - Estrutura da RNA



Fonte: HAYKIN, 2001.

Na ciência florestal os principais tipos de redes utilizadas são as Multilayer Perceptron (MLP) e Radial Basis Function (RBF) (BINOTI, 2012; CASTRO; 2011). Redes MLP são redes de multicamadas alimentadas adiante, que possuem uma ou mais camadas de neurônios na camada oculta, que conseguem extrair padrões não-lineares dos dados, por outro lado as RBF apresentam somente uma camada oculta extraíndo apenas padrões lineares na camada de saída (LIPPNANN, 1987; BRAGA et al., 2000; HAYKIN, 2001).

De acordo com Silva et al. (2010) o treinamento de uma RNA inicia pela apresentação dos dados (variáveis de entrada e saída) a uma arquitetura pré-estabelecida ou não. A arquitetura de uma RNA consiste na maneira como os neurônios estão estruturados e suas conexões, ou seja, número de camadas da rede, número de neurônios em cada camada e tipo de conexão entre os neurônios (BRAGA et al., 2000).

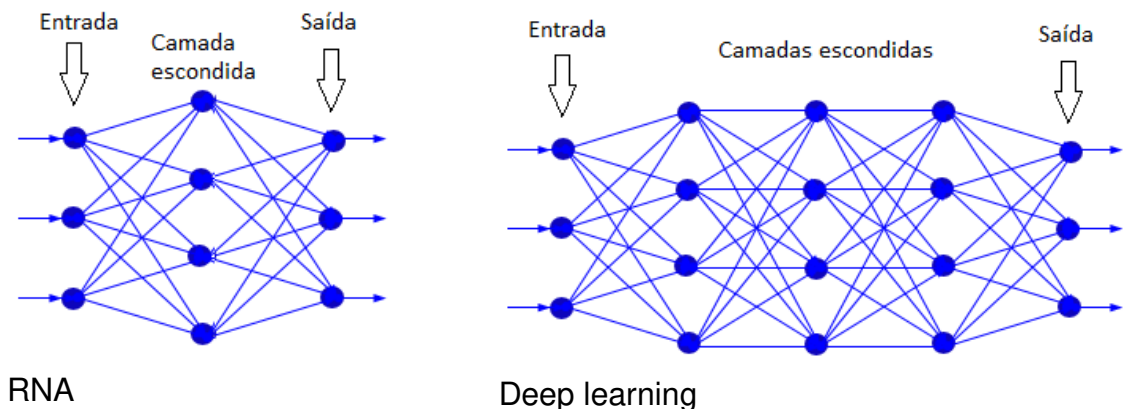
O processo de treinamento inicia com valores aleatórios dos pesos e esses pesos são atualizados de acordo com o algoritmo utilizado. O principal algoritmo utilizado é o back-propagation que, de acordo com Miguez et al. (2012), é algoritmo computacionalmente eficiente para o treinamento de redes MLP (Multilayer Perceptron) e que resolve o problema de realizar a propagação reversa do erro em RNAs com múltiplas camadas. Com base nestes valores obtidos, a primeira saída é comparada com o respectivo valor real da primeira observação. A diferença entre a saída estimada pela rede e o valor real gera um sinal de erro que calibra o ajuste de pesos, iniciando assim um novo ciclo, a fim de aproximar a saída do resultado desejado, ou seja, de minimizar o erro (HAYKIN, 2001).

De acordo com Silva et al. (2010) as funções de ativação podem ser do tipo: linear, sigmoide ou tangente hiperbólica. Basicamente as funções de ativação calculam o estado de atividade do neurônio decidindo se o neurônio deve ser ativado ou não, isto é, se a informação fornecida por ele é relevante ou não para desempenhar a tarefa.

O uso da técnica de RNA na modelagem de árvores individuais é encontrado em Guan e Gertner (1991), que construíram um modelo utilizando redes RBF para estimativas de mortalidade para árvores de Pinus, Hasenauer e Merkl (1997), que estimaram mortalidade de Picea abies na Áustria, Castro (2011 e 2012), que utilizou redes MLP para estimar a mortalidade de árvores de Eucalyptus no Brasil.

A diferença entre uma rede neural profunda (deep learning) e uma rede convencional é que a deep learning possui múltiplas camadas ocultas (Figura 2) de transformações não lineares, sendo geralmente utilizada para problemas relativamente mais complexos (DATA SCIENCE ACADEMY, 2019; DATA CAMP, 2018).

Figura 2 - Diferenças de rede neural convencional e rede neural profunda de acordo com as camadas ocultas



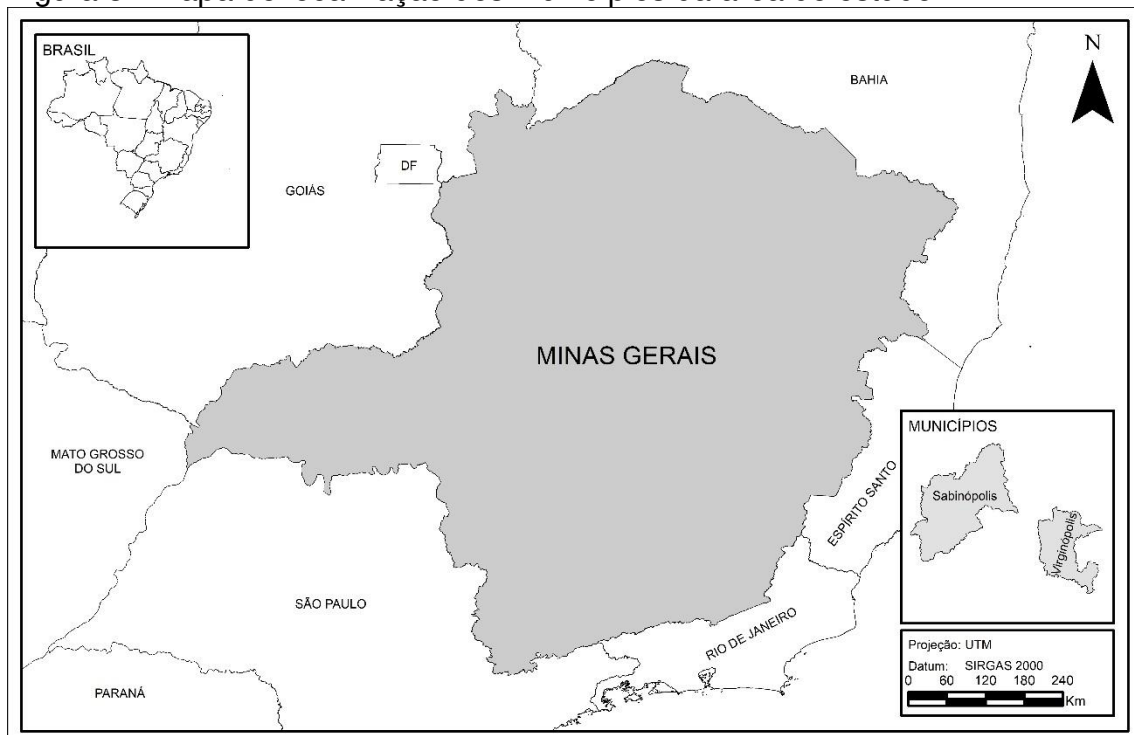
Redes neurais profundas tem sido utilizadas para diagnóstico de doenças por meio de classificação de imagens (SOUZA et al., 2019), reconhecimento de imagens de trânsito (SOBRINHO et al., 2016), algoritmos de recomendação de conteúdos pela Netflix, Instagram, Amazon e Spotify (MIRANDA, 2017; SILVA, 2016; CARRANZA, 2018) e, mais recentemente, identificação de áreas de desmatamento (TAQUARY, 2019), identificação de queimadas (OLIVEIRA, 2019), reconhecimento de espécies utilizando imagens digitais da madeira (OLIVEIRA, 2018) e reconhecimento de espécies a partir de imagens digitais dos troncos (COTRIM; PEDRINI, 2019).

4. MATERIAL E MÉTODOS

4.1. Dados

Para este estudo foram utilizados dados provenientes de 1061 parcelas permanentes de inventário florestal contínuo (IFC) realizado em povoamentos de clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* entre os anos de 2012 e 2019. Os mesmos se encontram em fazendas localizadas nos municípios de Virgíópolis e Sabinópolis, ambos no Estado de Minas Gerais (Figura 3), sob regime de alto fuste.

Figura 3 - Mapa de localização dos municípios da área de estudo



As medições foram realizadas nas idades de 16 a 103 meses, sendo registrados os seguintes dados: parcela, área da parcela, medição, cova, fuste, idade, diâmetro à altura de 1,30 m (dap) e altura total (Tabela 1).

Em cada parcela foram mensurados os diâmetros à 1,30 m do solo (dap) maiores que 5,0 cm e as alturas totais das 15 primeiras árvores. A altura dominante (Hd) de cada parcela em cada medição foi definida pelo critério de Assmann/Hummel (ASSMANN, 1970), calculando-se a média das alturas das 100 árvores de maior diâmetro por hectare. A altura total das demais árvores foi estimada a partir da equação (1), ajustada considerando a base de dados observados.

$$\ln Ht = 2,1406 + \left(\frac{-5,7469}{dap} \right) + 0,3470 * \ln Idade \quad R^2 = 0,7968, RQME = 11,8\% \quad (1)$$

em que: dap é o diâmetro com casca medido a 1,30 m de altura (cm), Ht é a altura total (m), idade é a idade (meses), Ln é o logaritmo neperiano, R² é o coeficiente de determinação e RQME é a raiz quadrada do erro médio, em porcentagem da média.

Tabela 1 - Algumas estatísticas descritivas dos dados utilizados no estudo

Idade (meses)	Variáveis	Média	Mínimo	Máximo
24	Dap (cm)	10,2	5,0	19,9
	Ht (m)	13,7	5,8	24,9
	B	0,2381	0,0194	0,4641
	Vol/há	55,9805	2,4150	124,5768
	N/há	949	445	1474
36	Dap (cm)	12,0	5,0	25,9
	Ht (m)	16,8	7,3	31,7
	B	0,3299	0,0963	0,6272
	Vol/há	92,5283	13,3950	262,5006
	N/há	956	405	1474
48	Dap (cm)	14,2	5,0	25,0
	Ht (m)	21,1	9,9	31,1
	B	0,4638	0,1427	0,7981
	Vol/há	160,3168	32,9369	298,6483
	N/há	953	461	1340
60	Dap (cm)	15,2	5,0	29,5
	Ht (m)	23,2	9,4	34,0
	B	0,5319	0,0511	0,9064
	Vol/há	202,7452	20,9376	372,8622
	N/há	945	464	1360
72	Dap (cm)	15,8	5,0	33,2
	Ht (m)	24,6	9,1	36,7
	B	0,5866	0,0287	1,3370
	Vol/há	236,4571	11,1799	517,1700

	N/há	961	707	1247
	Dap (cm)	16,7	5,1	29,5
	Ht (m)	26,4	11,1	40,0
84	B	0,6589	0,2663	1,2292
	Vol/há	282,1165	90,9709	473,6610
	N/há	984	601	1221
	Dap (cm)	17,8	5,4	30,2
	Ht (m)	28,8	10,9	41,5
96	B	0,7338	0,0366	1,0238
	Vol/há	334,4538	17,4900	562,4438
	N/há	1004	614	1217

O volume individual com casca (V_{cc}) de cada árvore foi estimado com uma equação de volume já disponível para os povoamentos:

$$V_{cc} = \text{Exp}(-10,0357 + 1,9002\text{LnDap} + 1,0248\text{LnHt}) \quad R^2 = 0,9934 \quad \text{RQME} = 2,40\% \quad (2),$$

em que: dap é o diâmetro com casca medido a 1,30 m de altura (cm), Ht é a altura total (m), Ln é o logaritmo neperiano, R^2 é o coeficiente de determinação e RQME é a raiz quadrada do erro médio, em porcentagem da média.

4.2. Construção de curvas de índice de sítio e projeção da sobrevivência

Antecedendo o ajuste de modelos de crescimento e de sobrevivência para árvore individual, foi feita a classificação da capacidade produtiva das áreas de estudo. O método da curva-guia foi aplicado para uma idade-índice de 72 meses, sendo construídas curvas anamórficas de índices de sítio, com amplitude de 2 m na

idade-índice. O modelo utilizado foi o Gompertz, $Hd = \beta_0 e^{-e^{(\beta_1 - \beta_2 I)}}$ (3), em que: Hd é a altura dominante (m), I a idade (meses), e $\beta_0; \beta_1; \beta_2 =$ parâmetros a serem estimados. Os índices gerados (S) foram utilizados como entrada para o ajuste dos modelos de crescimento e de sobrevivência. Os índices de sítio foram obtidos pela

transformação $Hd_f = Hd_a f(I_f) f(I_a)^{-1}$, para $I_f =$ idade-índice e $I_a =$ uma idade qualquer (método da curva-guia), resultando em:

$$S = Hd \frac{\beta_0 e^{-e(\beta_1 - \beta_2 I_i)}}{\beta_0 e^{-e(\beta_1 - \beta_2 I)}} = Hd \frac{e^{-e(\beta_1 - \beta_2 I_i)}}{e^{-e(\beta_1 - \beta_2 I)}} \quad (4),$$

em que S é a altura dominante (m) na idade-índice I_i (índice de sítio) e Hd é a altura dominante observada na idade I, sendo as idades em meses.

Para estimar a sobrevivência futura (N_2) foi seguida a metodologia descrita por Campos e Leite (2017). Inicialmente foi realizada a contagem do número de árvores em todas as parcelas e medições. Posteriormente os dados foram organizados em estrutura de projeção, com as informações pareadas ano a ano. A base de dados completa foi então separada em 50% dos dados para treinamento e 50% para validação dos modelos e RNA convencional e profunda.

O modelo ajustado para estimar N_2 foi $N_2 = N_1 + e^{\beta_0(I_2^{-\beta_1} - I_1^{-\beta_1})} + \varepsilon$ (5) (Pienaar e Shiver, 1981), em que: N_1 = número de árvores por hectare em uma idade atual, N_2 = número de árvores por hectare em uma idade futura, I_1 e I_2 = idade quaisquer, anterior e posterior, respectivamente, e β_0 ; β_1 = parâmetros a serem estimados. Esta relação funcional é obtida ao transformar um modelo de predição $Y = \phi_0 e^{\phi_1 X^{-\phi_2}} + \varepsilon$ em um modelo de projeção, pela mesma transformação simples utilizada no caso do método da curva-guia, $Y_f = Y_a f(x_f) f(x_a)^{-1}$, resultando em:

$$N_2 = N_1 \frac{\phi_0 e^{\phi_1 I_2^{-\phi_2}}}{\phi_0 e^{-\phi_1 I_1^{-\phi_2}}} = e^{\phi_1 I_2^{-\phi_2}} e^{-\phi_1 I_1^{-\phi_2}}$$

$$N_2 = N_1 e^{\phi_1 (I_2^{-\phi_2} - I_1^{-\phi_2})}$$

4.3. Modelagem do crescimento em diâmetro e altura

Inicialmente utilizamos um algoritmo genético implementado no software Select® (BINOTI, 2012) para separar os dados em duas sub-amostras, sendo uma para ajuste de regressão (e treino de redes neurais) e outra para validação. A separação resultou em cerca de 50% dos dados para ajuste dos modelos de regressão e treinamento das redes e os outros 50% dos dados para validação.

Com a amostra separada para ajuste (ou treino) foram ajustados três modelos para diâmetro (dap) e para altura total (Ht) (Tabela 1). O modelo de Pienaar e Shiver (1981) foi ajustado na forma apresentada pelos autores, ou seja, numa estrutura de

projeção explícita. Para que fosse possível utilizar todos os dados da sub-amostra de treino os modelos exponencial e Gompertz (transformados ela inclusão do índice de sítio) foram ajustados como modelos de predição, sendo as equações transformadas por $Y_f = Y_a f(x_f) f(x_a)^{-1}$. No caso do modelo de Pienaar e Shiver optamos pelo ajuste já na forma de projeção para se ter uma base comparativa com outros estudos. Os ajustes dos modelos foram feitos com o *software Statistica 12.0* (STATSOFT, INC, 2018), utilizando o método de estimação *Livenberg-Marquardt*.

Tabela 1 - Modelos de regressão ajustados para diâmetro e altura total

Modelos	
Pienaar e Shiver (1981)	$Y_2 = Y_1 + \exp\left(-\beta_0 \cdot \left(I_2^{\beta_1} - I_1^{\beta_1}\right)\right) + \varepsilon$
Exponencial Modificado	$Y = e^{\beta_0 + \beta_2 I^{-1} S^{-1}} + \varepsilon$
Gompertz Modificado	$Y = \beta_0 e^{-e(\beta_1 - \beta_2 I S)} + \varepsilon$

Y_1 e Y_2 = Diâmetro no tempo atual e futuro (cm) ou altura no tempo atual e futuro (m) ou número de árvores por hectare atual e futuro; I_1 e I_2 = Idade atual e futura (meses); S = índice de sítio (m); Ln = logaritmo neperiano; β_0 ; β_1 ; β_2 = parâmetros a serem estimados; ε = erro aleatório.

Para o modelo de Pienaar e Shiver (1981), foram avaliadas duas situações: (a) uma única equação de crescimento para todos os dados; (b) uma equação de crescimento para cada classe de índice de sítio (I, II e III).

Os modelos Exponencial e Gompertz, ajustados como modelos de predição, foram convertidos em modelos de projeção, conforme Campos e Leite (2017), com a aplicação da transformação já mencionada, $Y_f = Y_a f(x_f) f(x_a)^{-1}$, resultando nos seguintes estimadores:

$$Y_2 = Y_1 \frac{e^{\beta_0 + \beta_2 I_2^{-1} S^{-1}}}{e^{\beta_0 + \beta_2 I_1^{-1} S^{-1}}} \quad \text{e} \quad Y_2 = Y_1 \frac{\beta_0 e^{-e(\beta_1 - \beta_2 I_2 S)}}{\beta_0 e^{-e(\beta_1 - \beta_2 I_1 S)}} = Y_1 \frac{e^{-e(\beta_1 - \beta_2 I_2 S)}}{e^{-e(\beta_1 - \beta_2 I_1 S)}}, \text{ em que: } Y \text{ é o dap ou a}$$

altura total e as demais variáveis conforme já definido. De acordo com Campos e Leite (2017) a aplicação dessas relações funcionais resulta em consistência nas projeções, de modo que o crescimento estimado para uma idade atual é igual ao crescimento observado nessa idade, ou seja, os modelos passam a ser de projeção em lugar de predição.

4.4. Treinamento de redes neurais artificiais

4.4.1. Redes neurais convencionais

Foram testadas diferentes arquiteturas de redes Multilayer Perceptrons (MLP), sendo avaliadas combinações de diferentes funções de ativação nas camadas oculta e de saída (tangente hiperbólica e logística). O número de neurônios na camada oculta foi dois, definido em função do número de variáveis de entrada e de saída (Tabela 2).

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

Para a normalização dos dados utilizou-se a transformação (6), em que X' é o valor normalizado e X_{\min} e X_{\max} são os valores mínimo e máximo da variável.

Tabela 2 - Variáveis utilizadas no treinamento de redes neurais artificiais para predição de diâmetro, altura e sobrevivência de árvores

Estimativa	Saída	Entradas Numéricas
Crescimento	dap ₂	dap ₁ , l ₁ , l ₂
	dap ₂	dap ₁ , l ₁ , l ₂ , s
	Ht ₂	Ht ₁ , l ₁ , l ₂
	Ht ₂	Ht ₁ , l ₁ , l ₂ , s
Sobrevivência	N ₂	N ₁ , l ₁ , l ₂

dap_1 e dap_2 = Diâmetro no tempo atual e futuro (cm); ht_1 e ht_2 = altura no tempo atual e futuro (m); N_1 e N_2 = número de árvores por hectare atual e futuro; l_1 e l_2 = Idade atual e futura (meses); S = índice de sítio (m); e Hd = altura dominante (m).

O algoritmo de treinamento empregado foi o *Resilient Propagation* (RPROP +) utilizando-se o *software NeuroForest 4.0*. O treinamento foi encerrado quando a execução do algoritmo alcançou 3.000 ciclos.

4.4.2. Redes neurais profundas (*deep learning*)

Para treinamento da rede neural profunda foi utilizado o pacote H2O (LEDELL et al., 2020), que é uma plataforma de *machine learning* e preditiva, de software livre, com execução em *R Studio* (R CORE TEAM, 2020). As variáveis de entrada foram as mesmas utilizadas no treinamento das redes convencionais (Tabela 2).

Foram testadas diferentes arquiteturas de redes neurais profundas, as quais resultaram em diferentes quantidades de camadas ocultas e neurônios. Nas camadas

ocultas foram testadas as funções de ativação dos neurônios: *Tanh*, *TanhDropout*, *Rectifier Maxe RectifierDropout* e *Maxout*. Para camada de saída foi utilizada a função de ativação linear em todas as redes neurais profundas treinadas. Já a função de erro, que tem por finalidade diminuir o erro durante o treinamento da rede foi a *Quadratic* (erro quadrático) e a distribuição Gaussiana. O algoritmo de treinamento empregado foi o Gradiente descendente. As demais configurações seguiram o padrão do pacote *H2O R*.

4.5. Validação

Os modelos de regressão e as RNA convencional e profunda foram avaliadas, considerando os seguintes parâmetros estatísticos: Coeficiente de correlação ($r_{\hat{y}y}$), equação (7); viés (bias), equação (8); Raiz quadrada do erro quadrático médio (RQEM), equação (9); Erros relativos (ER%), equação (10); e análises gráficas. Quanto maior o valor obtido de coeficiente de correlação e menor RQEM e *bias*, melhor a previsão da estimativa.

$$r_{\hat{y}y} = \frac{n^{-1} \sum_{i=1}^n (Y_{pi} - \hat{Y}_m)(Y_i - \bar{Y})}{\sqrt{n^{-1} \sum_{i=1}^n (Y_{pi} - \hat{Y}_m)^2 n^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (7)$$

$$Bias = \sum_{i=1}^n \frac{(Y_{pi} - Y_i)}{n} \quad (8)$$

$$RQEM \% = 100 \bar{Y}_i^{-1} \sqrt{n^{-1} \sum_{i=1}^n (Y_i - Y_{pi})^2} \quad (9)$$

$$ER \% = 100 \left(\frac{Y_{pi} - Y_i}{Y_i} \right) \quad (10)$$

Em que: n : número de observações, Y_{pi} : produção predita, Y_i : produção observada, \bar{Y} : média das produções observadas.

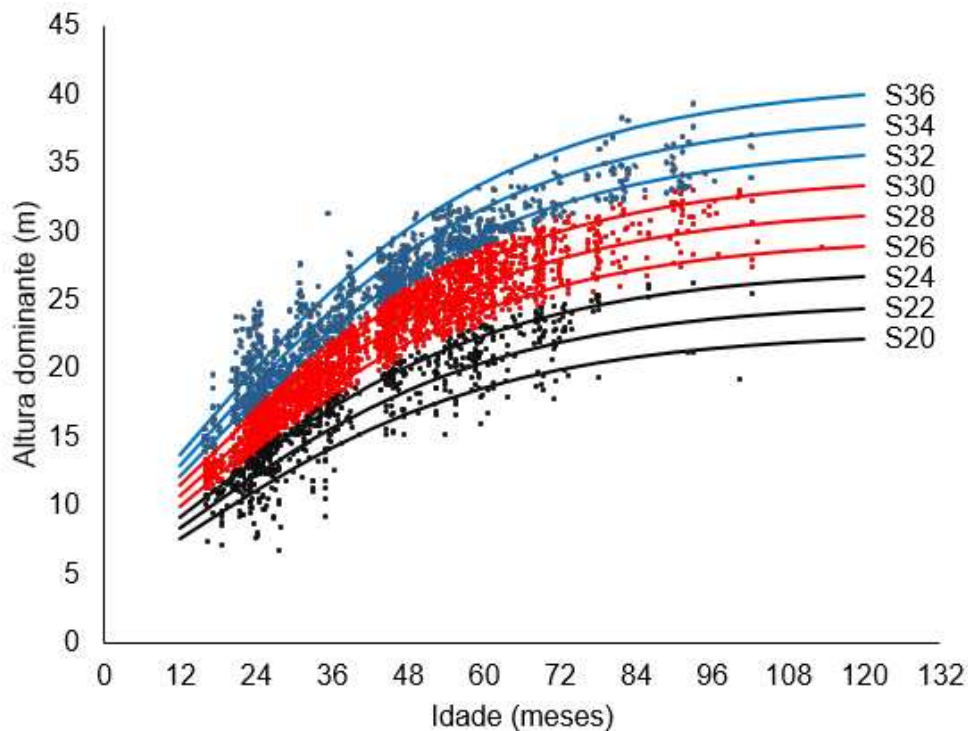
5. RESULTADOS E DISCUSSÃO

5.1. Classificação da capacidade produtiva

Para classificação da capacidade produtiva os dados de Hd variaram de 17 a 36 metros na idade índice (I_i) de 72 meses. O método da curva-guia foi aplicado com o uso do modelo de Gompertz, gerando assim a equação utilizada para estimar Hd numa idade qualquer. O ajuste resultou em um coeficiente de determinação (R^2) igual a 0,7170 e RQEM igual a 13,4236%.

Aplicado o modelo de Gompertz e considerando a idade índice (I_i) igual a 72 meses, foram geradas 6 curvas, cada uma representando valores de sítio (S) de 18 a 38 metros, abrangendo idades entre 12 e 120 meses, as quais foram agrupadas em três grupos denominados Classe I, Classe II e Classe III (Tabela 3), representando capacidade produtivas do local (Figura 4).

Figura 4 - Curvas de índices de sítio para povoamento de clones híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* localizados no estado de Minas Gerais com idade-índice de 72 meses



A delimitação dos intervalos e conseqüentemente os centros de classe de cada sítio é uma variável importante, pois a capacidade produtiva do local influencia diretamente no desenvolvimento do povoamento.

Tabela 3 - Curvas de índices de sítio para clones e híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis*

Sítio	Produtividade	Amplitude da Classe (Hd)	Centro de Classe
Classe I	Alta	$H_d > 31$	34
Classe II	Média	$25 < H_d \leq 31$	28
Classe III	Baixa	$H_d \leq 25$	22

Analisando os resultados de capacidade produtiva, pode-se observar que a frequência dos índices atingidos pela maioria das árvores mensuradas encontra-se na Classe II. Isto é, a maioria dos indivíduos podem atingir alturas de 25 a 31 m na idade de colheita aos 72 meses.

5.2. Modelos de regressão

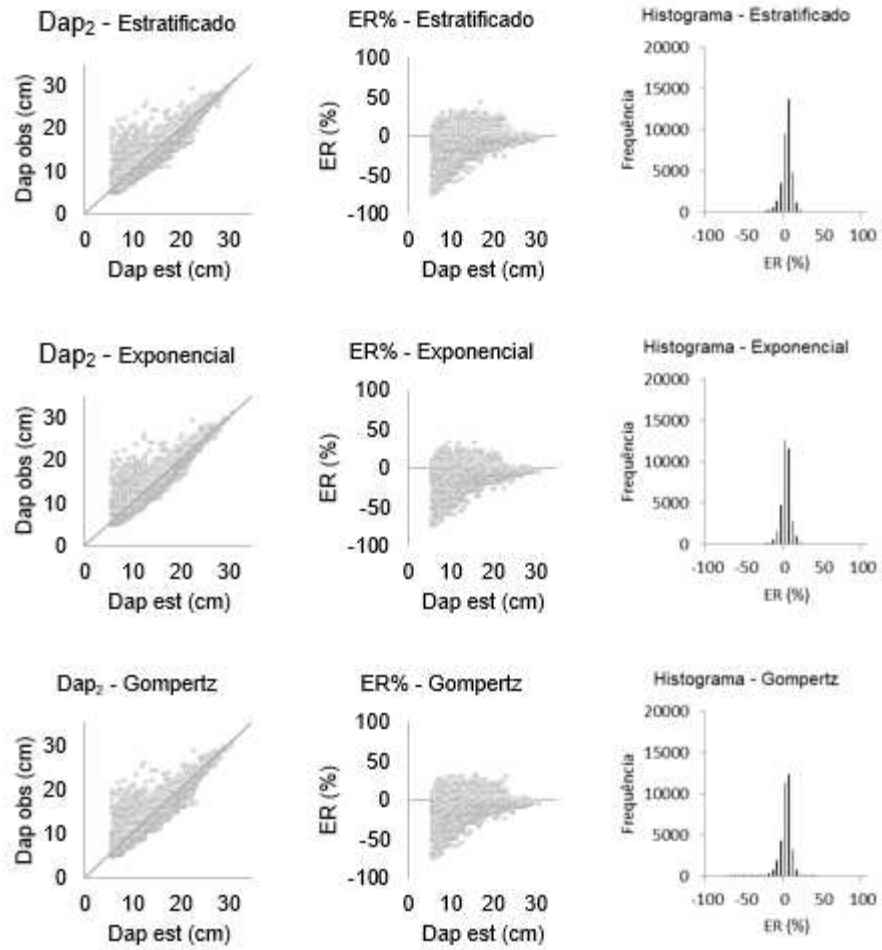
Para as estimativas da variável dap_2 os valores obtidos de coeficiente de correlação variaram de 0,9391 a 0,9417 e os valores de RQEM entre 7,6840 e 7,9831 (Tabela 4). O modelo de Pienaar e Shiver (1981) apresentou maior valor de r_{yy} e redução no RQEM e *bias* no ajuste que considerou a estratificação da área, os parâmetros estatísticos considerados demonstraram que o ajuste por classe de sítio tornou-o modelo mais eficaz e preciso. Em estudos Taquetti (2014) descreveu a eficácia da estratificação da área na estimativa de produção de madeira em povoamentos de *Eucalyptus sp.* E ainda, de acordo com Campos e Leite (2017) $Y = f(I, S)$ pode resultar em estimativas consistentes, permitindo avaliar o efeito da capacidade produtiva do local.

Para os modelos de projeção o modelo Exponencial apresentou melhor desempenho na avaliação estatística apresentando maiores valores r_{yy} e redução nos valores de RQEM e *bias* comparado aos demais modelos para estimar a variável dap_2 .

Tabela 4 - Estimativas dos parâmetros dos modelos de regressão ajustados para estimar o diâmetro futuro (dap_2) de árvores de eucalipto na área de estudo e correspondentes estatísticas de precisão e de exatidão

Modelo (dap_2)	Estratificação	β_0	β_1	β_2	Bias	RQEM	r_{yy}
Pienaar e Shiver (1981)		5,3117	-0,5332		0,0962	7,8888	0,9391
Pienaar e Shiver (1981)	Classe I	3,7289	-0,4037				
	Classe II	5,8748	-0,5675		0,0747	7,7673	0,9399
	Classe III	7,9285	-0,6892				
Exponencial Modificado		2,9505	-389,75	7,4399	0,2592	7,6840	0,9417
Gompertz Modificado		0,3483	0,2090	0,0411	0,2545	7,9831	0,9394

Analisando os gráficos, verificou-se que há alta correlação entre os dap_2 estimados e observados indicando a eficácia na estimativa da variável desejada (Figura 5). O modelo Exponencial apresenta uma sutil propensão ao agrupamento dos pontos no eixo de 45° apontando um melhor desempenho em estimar a variável dap_2 comparado ao demais modelos. Verificando os gráficos de erro relativo é possível observar uma tendência de subestimação da variável entre as classes de 10 a 20 cm de dap . Os histogramas de frequência dos resíduos apresentam uma maior distribuição dos erros entre as classes -30% a 30%, com exceção do modelo Pienaar e Shiver (1981) no ajuste geral dos dados e o modelo de Gompertz que, mesmo tendo subestimação, a frequência em que acontece é de menor grau, tendendo a normalidade.

Figura 5 - Análise gráfica dos resíduos dos modelos de regressão para variável dap_2 

Na estimativa de H_{t2} os modelos apresentaram desempenho semelhantes aos observados para estimativa do dap_2 , o modelo de Pienaar e Shiver (1981) com ajuste geral dos dados apresentou valor de r_{yy} sutilmente maior do que o ajuste com os dados estratificados por classe de sítio, entretanto, no ajuste do modelo estratificado houve redução no RQEM e *bias* (Tabela 5). O modelo Exponencial modificado novamente apresentou maior valor de r_{yy} e redução de RQEM (Tabela 5), enquanto, que o valor de *bias* foi de 0,2329 menor somente que o valor do *bias* 0,2900 do modelo de Gompertz.

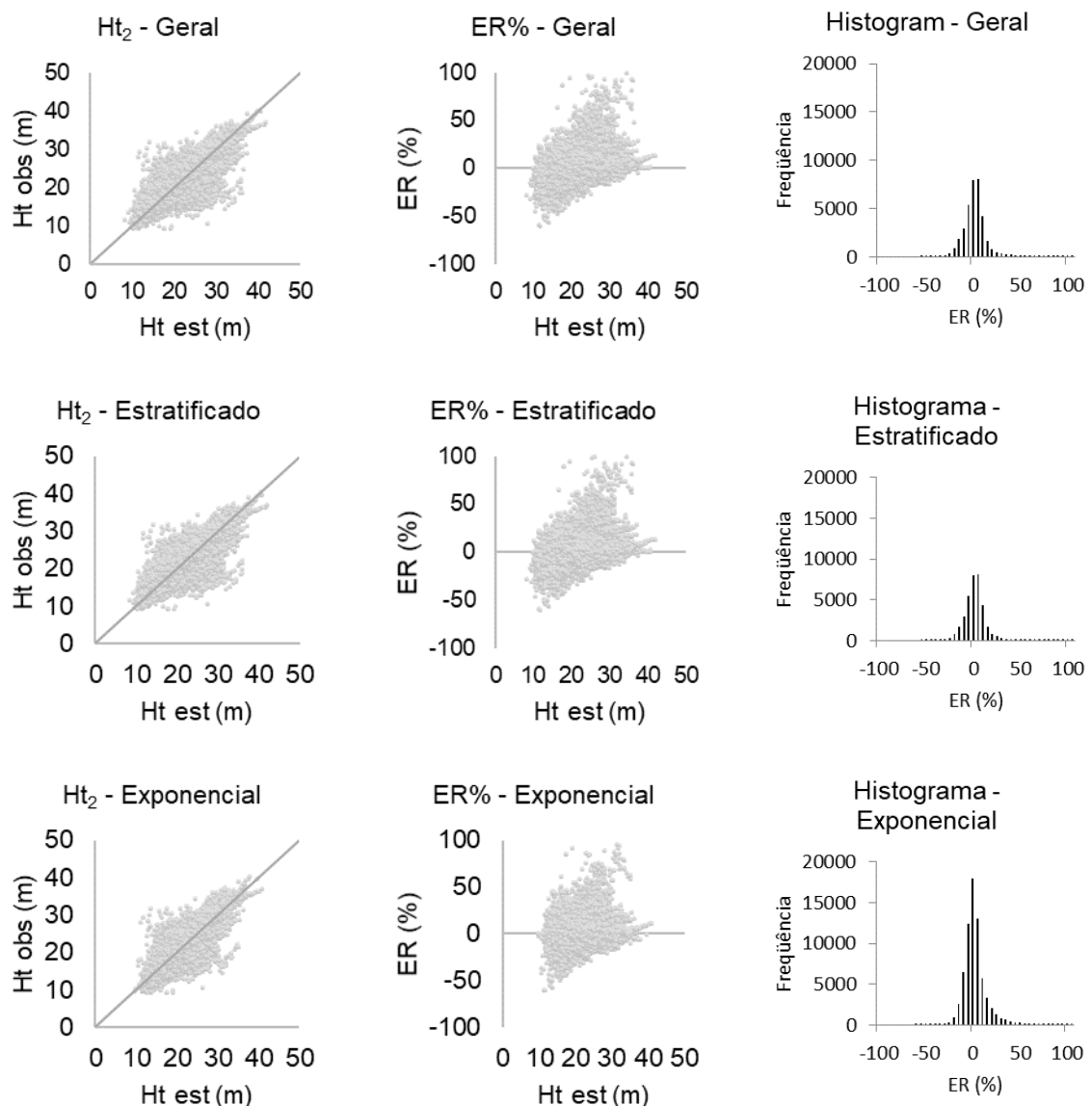
Tabela 5 - Estimativas dos parâmetros dos modelos de regressão ajustados para estimar a altura total futura (H_{t2}) de árvores de eucalipto na área de estudo e correspondentes estatísticas de precisão e de exatidão

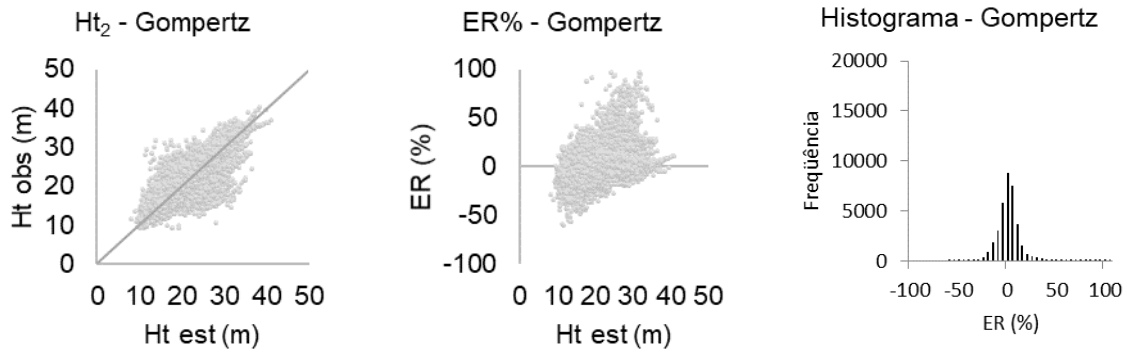
Modelo (H_{t2})	Estratificação	β_0	β_1	β_2	Bias	RQEM	r_{yy}
Pienaar e Shiver (1981)		5,2086	-0,3241		0,2056	11,4434	0,8593
Pienaar e Shiver (1981)	Classe I	7,1669	-0,0905		0,1519	11,3586	0,8587
	Classe II	5,7985	-0,3949				
	Classe III	5,2188	-0,4156				
Exponencial Modificado		3,4895	-545,8175	-0,0570	0,2329	10,5613	0,8610
Gompertz Modificado		1,0038	0,4601	0,0371	0,2900	11,3651	0,8578

Analisando os gráficos observa-se que a distribuição das estimativas e dos valores observados de H_{t2} tendem a se concentrar em torno da linha de 45º no gráfico de dispersão, com propensão a subestimar e superestimar a variável de interesse. Os gráficos

de erro relativo demonstram uma capacidade de erro superestimado da variável nos valores de Ht maiores que 20 metros, essa tendencia se demonstra atenuada no modelo Exponencial modificado. Os histogramas de frequência dos resíduos no geral tendem a normalidade, com pequenas concentrações de -50 a -30 e 50 a 100, a maior concentração das frequências de resíduos entre -30 e 30 é observada no modelo Exponencial modificado.

Figura 6 - Análise gráfica dos resíduos dos modelos de regressão para variável Ht₂





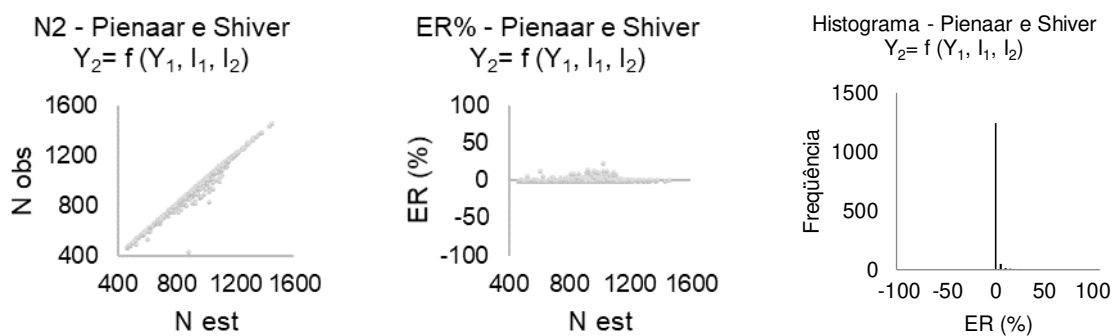
Para estimativa da sobrevivência futura (N_2) o modelo de Pienaar e Shiver (1981) apresentou bom desempenho estatísticos (Tabela 6) com coeficiente de correção próximo de 1 e valor de RQEM menor que 10%. Resultados semelhantes foram observados por Penido *et al.*, (2019), Salles (2016) e Nogueira *et al.*, (2005) ao estimar sobrevivência utilizando o modelo de Pienaar e Shiver (1981).

Tabela 6 - Parâmetros do modelo de regressão para variável N_2

Modelo (Ht ₂)	β_0	β_1	Bias	RQEM	r_{yy}
Pienaar e Shiver (1981)	0,00004	1,56309	3,7517	2,0895	0,9911

Nas análises gráficas (Figura 7) é possível observar alta correlação entre a variável observada e estimada, com a grande maioria dos pontos concentrados em torno do eixo de 45°. O erro relativo também apresentou essa tendência com baixos valores de erro se concentrando em torno do eixo "x". O histograma de erros relativos evidenciou o ocorrido resultando em grande acúmulo do erro em torno de zero.

Figura 7 - Análise gráfica dos resíduos do modelo de regressão Pienaar e Shiver (1981) para variável N_2



Pode-se observar que a mortalidade natural das árvores observadas na base de dados foi pequena, diminuindo ligeiramente o número de árvores ao longo dos

anos. Salienta-se que a estimativa da quantidade de árvores mortas é de difícil previsibilidade, aleatório e com causas não complementarmente identificadas em uma unidade de manejo ou povoamento (CASTRO *et al.*, 2013).

De acordo com Campos e Leite (2017) para projeção em povoamentos de eucalipto, se a mortalidade for insignificante entre cerca de três e sete anos. Há a alternativa de estimar a taxa de sobrevivência por classe de *dap*. Outra sugestão seria admitir uma taxa fixa de mortalidade ou sobrevivência para o povoamento em questão, tendo em vista, que se trata de um povoamento de clone de híbridos de *Eucalyptus grandis* X *Eucalyptus urophylla*.

5.3. Redes neurais artificiais e Redes neurais profundas

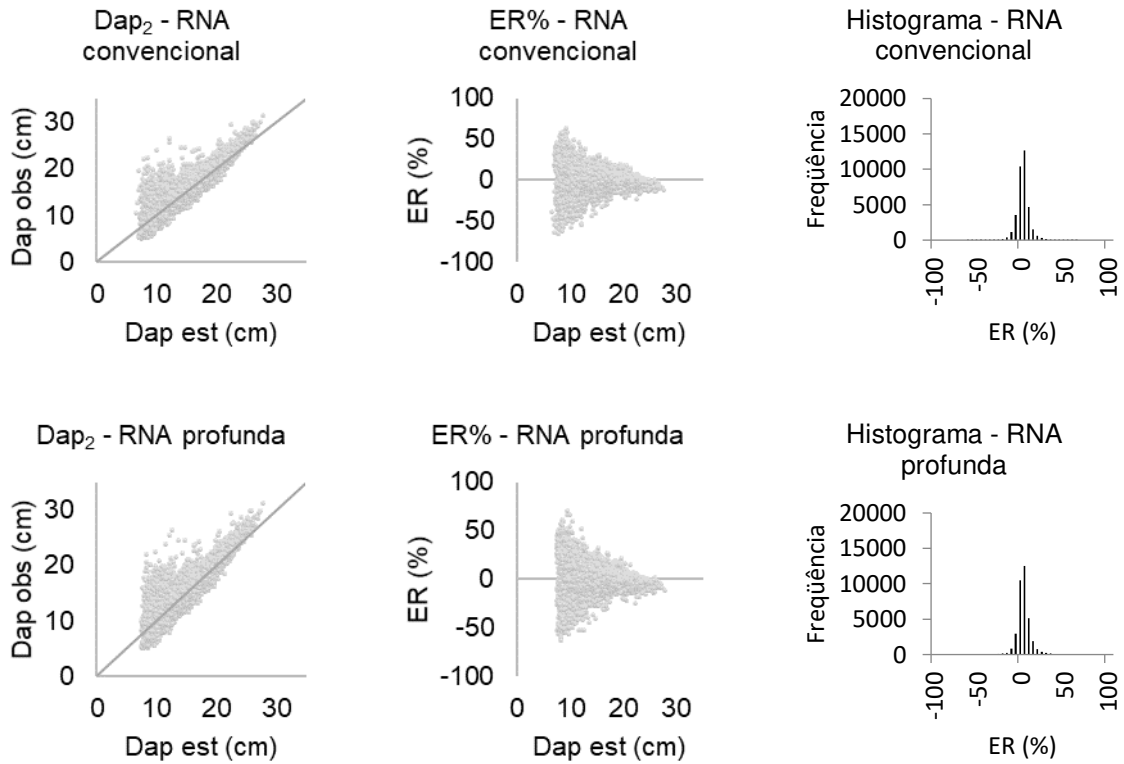
Os resultados estatísticos para avaliação das RNA convencional e profunda podem ser analisados na Tabela 7. Para a estimativa das variáveis dap_2 e Ht_2 observa-se que ocorreu melhor desempenho da RNA profunda com maior valor de r_{yy} e redução no RQEM e bias comparada a RNA convencional.

Tabela 7 - Parâmetros das RNA convencional e profunda para variáveis dap_2 e Ht_2 .

RNA	Saída	Bias	RQEM	r_{yy}
convencional	Dap_2	0,0152	7,4247	0,9424
	Ht_2	0,0776	9,6324	0,8837
profunda	Dap_2	-0,0723	7,2788	0,9468
	Ht_2	-0,0785	8,5847	0,9089

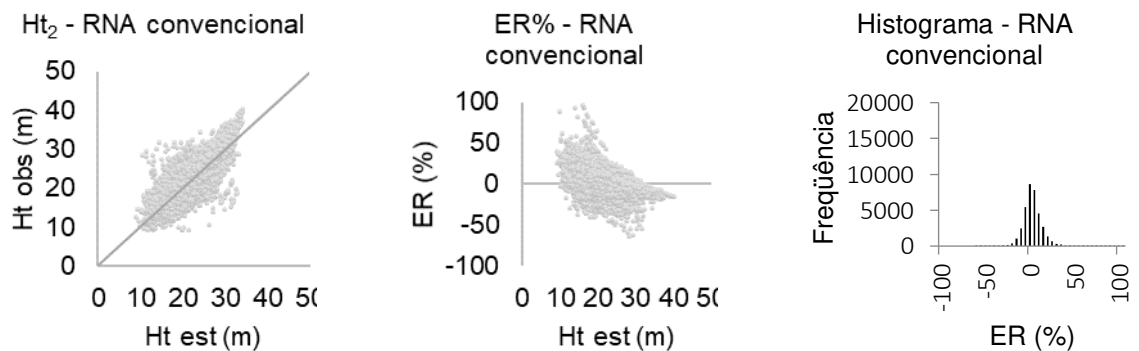
Analisando os gráficos (Figura 8) observa-se que a distribuição das estimativas e dos valores observados de dap_2 tendem a se concentrar em torno da linha de 45º no gráfico de dispersão, com propensão a superestimar a variável de interesse para RNA convencional e também RNA profunda. O gráfico de erro relativo, assim como nas regressões (Figura 7), novamente aponta maior dispersão dos erros nos dap de 10 a 20 cm, entretanto nas RNA é possível observar que além da subestimação há superestimação da variável estimada (dap_2). Os histogramas de frequência dos resíduos para estimativa do dap_2 apresentam pequenos acúmulos de frequência de -50 a 50, com leve tendência a normalidade.

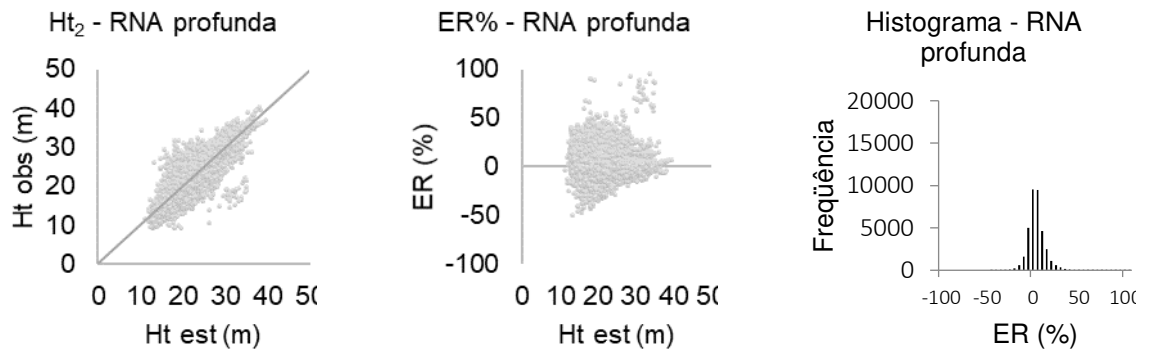
Figura 8 - Análise gráfica das RNA convencional e profunda para variável dap₂



Já para estimativa da Ht_2 é possível observar que a RNA profunda apresentou melhor desempenho, demonstrando que a distribuição das estimativas e dos valores observados tendem a se concentrar em torno da linha de 45º no gráfico de dispersão, que o gráfico de erro relativo apresenta os pontos aparentemente mais próximos do eixo “x” com alguns problemas de superestimação da variável, o histograma de frequência do erro tende a normalidade com uma maior concentração em torno de zero.

Figura 9 - Análise gráfica das RNA convencional e profunda para variável Ht_2





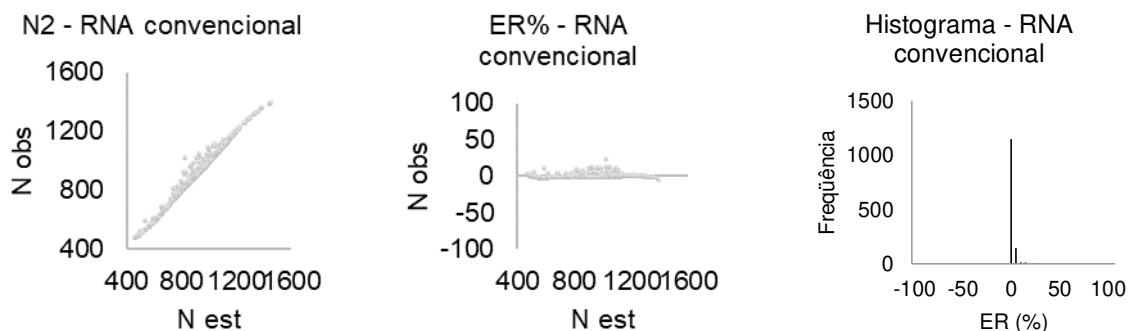
Para estimativa da sobrevivência futura (N_2), observa-se que as RNA convencional e profunda apresentaram bom desempenho estatísticos (Tabela 8) com coeficiente de correção próximo de 1 e valor de RQEM menor que 10%. Pode se observar melhor desempenho da RNA profunda comparada a RNA convencional, com menor valor de RQEM e aumento sutil no $r_{\hat{y}y}$.

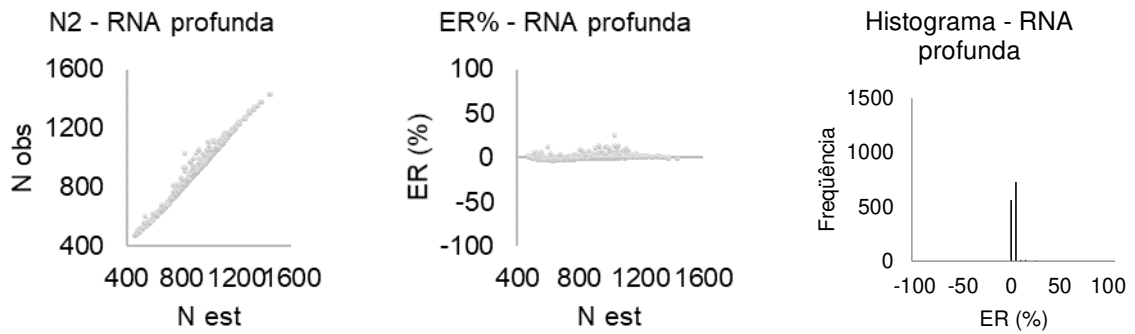
Tabela 8 - Parâmetros das RNA convencional e profunda para variáveis N_2 .

RNA	Saída	Bias	RQEM	$r_{\hat{y}y}$
RNA convencional	N_2	-3,9096	1,7778	0,9944
RNA Profunda	N_2	1,6093	1,5810	0,9947

Para as análises gráficas (Figura 10) é possível observar alta correlação entre a variável observada e estimada, com os pontos concentrados em torno do eixo de 45°. O erro relativo apresentou baixos valores de erro se concentrando em torno do eixo "x" e histograma de erros relativos apresentou grande acúmulo do erro em torno de zero.

Figura 10 - Análise gráfica das RNA convencional e profunda para variável N_2

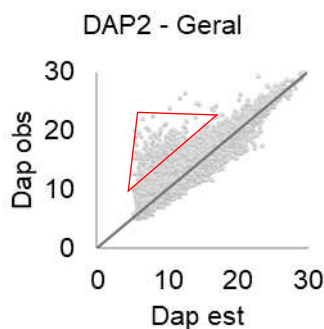




5.4. Análise dos resíduos

Quando se trata de árvores individuais, quase sempre irão ocorrer alguns casos discrepantes. Na Figura (11) é possível observar que a quantidade de pontos em destaque é de pequena proporção, as árvores que compõe esse agrupamento representam apenas 0,35% da totalidade dos dados. Em estudo de modelagem de árvore individual Abreu (2019) descreve sobre a influência dos dados discrepantes em aumentar a inclinação da curva ajustada e avaliou metodologias para reduzir esta influência na modelagem.

Figura 11 -Faixa de dados discrepantes na modelagem de árvore individual

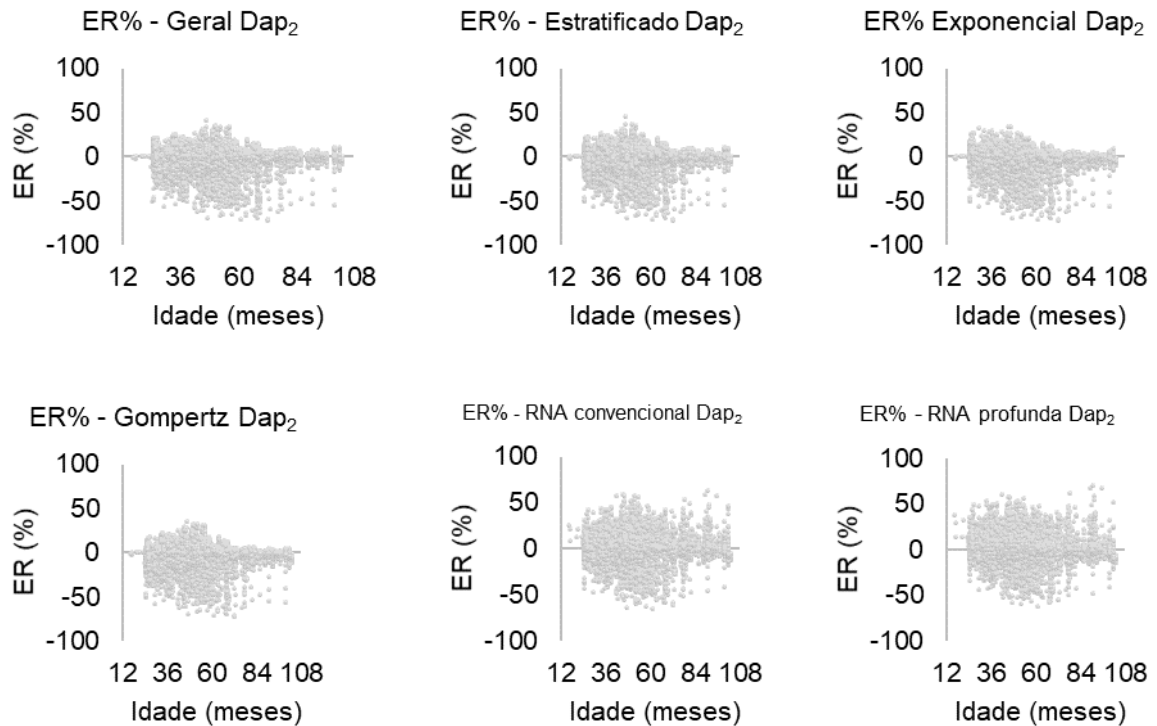


Os histogramas de resíduos das regressões e RNA convencional e profunda, apresentaram tendência a normalidade, demonstrando que apesar da presença dos dados discrepantes não ocorreu interferências significativas para projetar dap_2 , Ht_2 e N_2 .

Para buscar entender se o erro ocorre em função da idade, foram gerados os gráficos da Figura 12, 13 e 14, com erro relativo (ER%) em função da idade em meses. Para projeção do dap_2 verificou-se que para as regressões e RNA o erro ocorreu independente da idade e que nas maiores idades, de 84 a 108 meses, esse erro

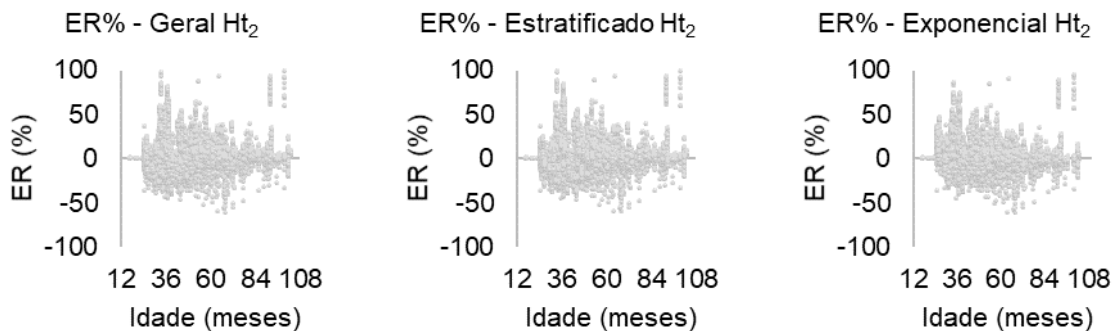
tendeu a ser menor em todos os métodos. Para as RNA convencional e profunda observa-se que de modo geral os erros ficaram mais dispersos.

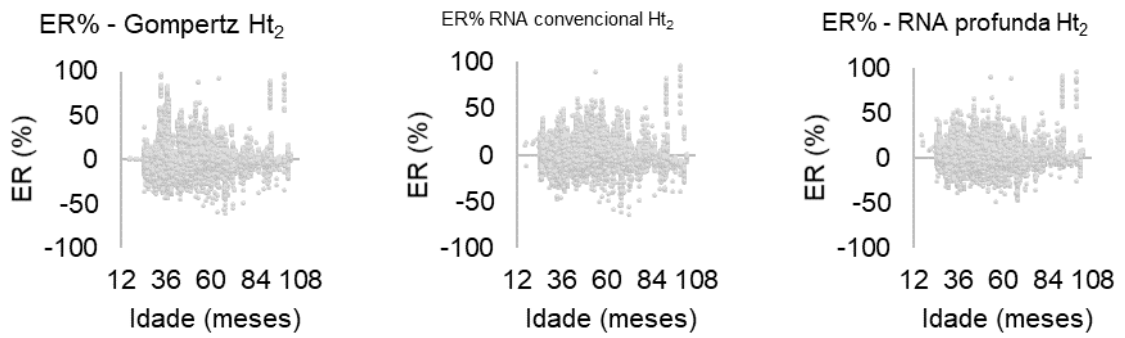
Figura 12 - Erro relativo em função das idades (meses) para projeção de Dap_2 das árvores individuais



Na Figura 13 verifica-se novamente que o erro ocorreu independente da idade com ocorrência maior de superestimação do erro em projetar a variável Ht_2 . Em comparação as regressões, as RNA convencional e profunda demonstraram os erros mais atenuados até a idade de 84 meses.

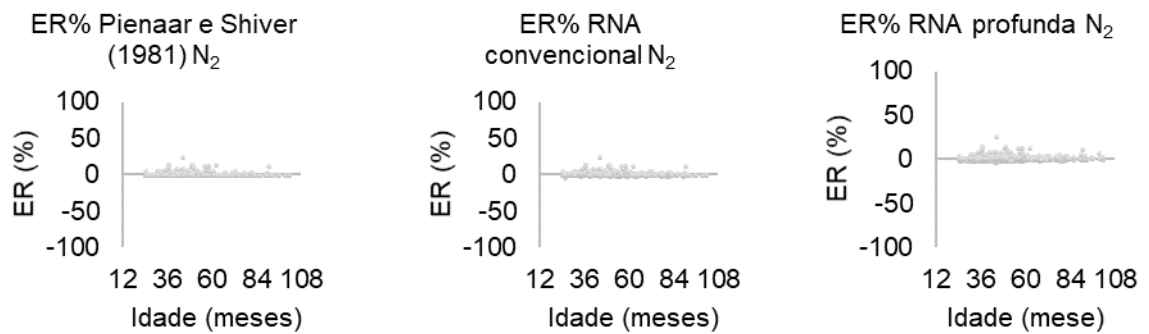
Figura 13 - Erro relativo em função das idades (meses) para projeção de Ht_2 das árvores individuais





Para projeção de sobrevivência (N_2) observou se na Figura 14 que o erro novamente ocorreu independente da idade, para todos os métodos o erro foi baixo com aproximação dos pontos em torno do eixo “x”.

Figura 14 - Erro relativo em função das idades (meses) para projeção de N_2



6. CONCLUSÕES

Com este estudo pode-se concluir que:

As RNA convencional e profunda foram mais eficientes para projetar dap_2 , Ht_2 e N_2 , apresentando melhor desempenho nos parâmetros estatísticos avaliados e nas análises gráficas. Portanto, ao construir modelos de árvores individuais para povoamentos de eucalipto é importante verificar o possível ganho de exatidão nas estimativas de diâmetro e altura com o uso de RNA, clássica e profunda, em relação ao emprego de regressão, ainda que utilizando as mesmas variáveis de entrada do método paramétrico.

REFERÊNCIAS

- ABREU, J. C. **Alternativas para estimar o volume de árvores individuais em formações florestais no estado de Minas Gerais**. 2019. 82 f. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa. 2019.
- ASSMANN, E., **The principles of forest yield study**. New York: Pergamon Press, 1970 506p.
- AVERY, T.E.; BURKHART, H. E. **Forest Measurements**. 4. ed., New York: McGraw-Hill Book, 1994, 408 p.
- BELCHER, D. W.; HODOWAY, M. R.; BRAND, G. J. **A Description of STEMS: The stand and tree evaluation and modeling system**. USDA: Forest Service, 1982. 17 p.
- BINOTI, D. H. B. **Computer systems applied to forest management**. 2012. 122 f. Tese (Doutorado em Manejo Florestal; Meio Ambiente e Conservação da Natureza; Silvicultura; Tecnologia e Utilização de) - Universidade Federal de Viçosa, Viçosa, 2012.
- BINOTI, M. L. M. D. S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**. V.37, n.4, p.639-645, 2013.
- BINOTI, M. L. M. S. **Emprego de redes neurais artificiais em mensuração e manejo florestal**. 2012. 130f. Tese (Doutorado) - Universidade Federal de Viçosa, Viçosa, 2012.
- BINOTI, M. L. M. S. **Rede neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. 2010. 54 f. Dissertação (Mestrado) – Universidade Federal de Viçosa, Viçosa, 2010.
- BRAGA, A.P.; CARVALHO, A.P.L.F.; LUDEMIR, T.B. **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro, RJ. Editora LTC. 262 f. 2000.
- CAMPOS, J. C.; LEITE, H. G. **Mensuração Florestal: Perguntas e respostas**. Viçosa: UFV, 5. Ed, 2017, 636 p.
- CARRANZA, F. **Recomendación de Música Basada en Contenido con Deep Learning**. Universidad Nacional de Córdoba, Argentina, 2018.
- CASTRO, R. **Modelagem de crescimento em nível de árvores individuais utilizando redes neurais e autômatos celulares**. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2011.
- Ciência e Ambiente**, Santa Maria, v. 20, n. 1, p. 53-69, 2000.
- Ciência e Ambiente**, Santa Maria, v. 20, n. 1, p. 53-69, 2000.
- CLUTTER, J. L.; FORTSON, J. C.; PIENAAR, L. V.; BRISTER, G. H.; BAILEY, R. L. **Timber management: A quantitative approach**. New York: John Wiley & Sons, 1983.

333 p.

COTRIM, A. N.; PEDRINI, H. **Identificação de Espécies de Árvores por Imagens de Tronco Utilizando Aprendizado de Máquina Profundo**. 2019. Instituto de Computação, Universidade Estadual de Campinas, Campinas-SP, 2019.

CRUZ, J. P.; LEITE, H. G.; SOARES, C. P. B.; CAMPOS, J. C. C.; SMIT, L.; NOGUEIRA, G. S.; OLIVEIRA, M. L. R. Modelos de crescimento e produção para plantios comerciais jovens de *Tectona grandis* em Tangará da Serra, Mato Grosso. **Revista Árvore**, Viçosa, v.32, n.5.

DANIELS, R. F.; BURKHART, H. E. **Simulation of individual tree growth and stand development in managed loblolly pine plantations**. Blacksburg: Virginia Polytechnic Institute and State University, 1975. (Publ. FWS – 5 – 75).

DATA CAMP. **Differences Between Machine Learning & Deep Learning**. Disponível em: <<https://www.datacamp.com/community/tutorials/machine-deep-learning>>. Acesso em: jun de 2020.

DATA SCIENCE ACADEMY. **Deeplearning Book, 2019**. Disponível em: <<http://www.deeplearningbook.com.br/>>. Acesso em: jun de 2020.

DAVIS, L. S.; JONHSON, K.N. **Forest Management**. 3. ed., New York: McGraw-Hill Book, 1987. 790 p.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA – EMBRAPA. **Transferência de Tecnologia Florestal: Eucalipto**. Disponível em: <<https://www.embrapa.br/florestas/transferencia-de-tecnologia/eucalipto/perguntas-e-respostas>>. Acesso em: jul de 2020.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS – FAO. **Global Forest Resources Assessment 2020 key findings**. Disponível em: <<http://www.fao.org/3/CA8753EN/CA8753EN.pdf>>. Acesso em: jun de 2020.

GLOVER, G. R.; HOOL, J. N. A basal area ratio predictor of loblolly pine plantation mortality. **Forest Science**, Washington, v. 25, n. 2, p. 275-282, 1979.

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

HASENAUER, H. Princípios para a modelagem de ecossistemas florestais.

HASENAUER, H.; MONSERUD, R.A. A crow model for Austrian species. **Forest Ecology and Management**, Amsterdam, v. 84, n. 1-3, p. 49-60, 1997.

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

IBGE. Instituto Brasileiro de Geografia e Estatística. **Produção da extração vegetal e da silvicultura 2019**. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/74/pevs_2019_v34_informativo.pdf>. Acesso em: jun. 2021.

INDÚSTRIA BRASILEIRA DE ÁRVORES. **Relatório 2020**: Indústria Brasileira de árvores. Disponível em: <<https://iba.org/datafiles/publicacoes/relatorios/iba-relatorioanual2020.pdf>>. Acesso em: mar de 2020.

INNOARCHTECH INSTITUTE. Deep Learning. Disponível em: <<https://www.innoarchitechinstitute.com/>>. Acesso em: jun de 2020.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**. 2015. (521): 436-444.

LEDELL, E. et al. **h2o: R Interface for the 'H2O' Scalable Machine Learning Platform**, R package version 3.28.0.4, R Foundation for Statistical Computing: Vienna, Austria, 2020.

LEDUC, D. J.; MATNEY, T. G.; BELLI, K. L.; BALDWIN, V. C. JR. Predicting diameter distribution of longleaf pine plantations: A comparison between artificial neural networks and other accepted methodologies. **Forest Service, United States Department of Agriculture**. Research paper. SRS-25, 2001.

LEE, Y. J. Predicting mortality for even-aged stands of lodgepole pine. **The Forestry Chronicle**, Mattawa, v. 47, n. 1, p. 29-32, 1971.

LIPPNANN, R. An introduction to computing with neural nets. **ASSP Magazine**, v.4, n.2, p.4-22, 1987.

MAGALHÃES, J.G. de S. **Prognose e regulação de florestas de eucalipto utilizando modelo de crescimento e inteligência computacional**. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Minas Gerais. 2014.

MARTINS, F.B.; SOARES, C.P.B.; LEITE, H.G.; SOUZA, A.L.; CASTRO, R.V.O. Índices de competição em árvores individuais de eucalipto. **Pesquisa Agropecuária Brasileira**, v.46, p.1089-1098, 2011.

MIRANDA, W. R. Netflix: Big Data e os algoritmos de recomendação. Intercom – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação. **XXII Congresso de Ciências da Comunicação na Região Sudeste** – Volta Redonda - RJ. jun 2017.

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

OLIVEIRA, Petronio Diego Silva de. **Uso de aprendizagem de máquina e redes neurais convolucionais profundas para a classificação de áreas queimadas em imagens de alta resolução espacial**. 2019. viii, 23 f., il. Dissertação (Mestrado em Geografia) — Universidade de Brasília, Brasília, 2019.

OLIVEIRA, Wellington. **Software para reconhecimento de espécies florestais a partir de imagens digitais de madeiras utilizando deep learning**. 2018. 103 f. Dissertação (Mestrado em Tecnologias Computacionais para o Agronegócio) – Universidade Tecnológica Federal do Paraná, Medianeira, 2018.

PEET, R. K.; CHRISTENSEN, N. L. Competition and tree death. **BioScience**, Washington, v. 37, n. 8, p. 586-595, 1987.

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). v 48. n 128.

PIENNAR, L. V.; SHIVER, B. D. Survival function for site prepared slash pine plantations in flat Woods of Georgia Northern Florida. **Southern Journal Forestry**, Bethesda, v. 5, n. 2, p. 59-62, 1981.

ROSSI, L. M. B.; KOEHLER, H. S.; SANQUETTA, C. S.; ARCE, J. E. Modelagem de mortalidade em florestas naturais. **FLORESTA**, Curitiba, PR, v. 37, n. 2, mai./ago. 2007.

SALLES, T. T. **Simulação de florestas equiâneas**. 2016. 138f. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa. 2016.

SALLES, T.T., LEITE, H.G., NETO, S.N. DE O., SOARES, C.P.B., PAIVA, H.N., SANTOS, F.L. Modelo de Clutter na modelagem de crescimento e produção de eucalipto em sistemas de integração lavoura-pecuária-floresta. **Pesquisa Agropecuaria Brasileira**, v.47, n.2, p.253-260, 2012.

SILVA, G. C. C.; CALEGARIO, N.; DILVA, A. A. L.; CRUZ, J. P.; LEITE, H. G. Site index curves in thinned and non-thinned eucaliptos stands. **Forest Ecology and Management**. v.408, p.36-44, 2018.

SILVA, I N da; *et al.* **Redes neurais artificiais**: para engenharia e ciências aplicadas: curso prático. São Paulo: Artliber, 2010. 399 p.

SILVA, M.L.M. da; 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.

SILVA, R. L. F. **Reconhecimento de ambientes em fotos do Instagram com deep learning**. Graduação em ciência da computação – UFP. Recife – PE. 2016.

SOBRINHO, M. V. O.; MENEZES, M. C.; NASCIMENTO, C. G. M.; BRAZ JUNIOR, G. **Reconhecimento de Sinais de Transito Utilizando Deep ^ Learning**.JIM 2016 - VI Jornada de Informática do Maranhão.

SOUZA, E. P.; GOMES, C. M.; BARROSO, D. H.; MIRANDA, V. L.; GURGEL-GONÇALVES, R. Aplicações do Deep Learning para diagnóstico de doenças e identificação de insetos vetores. **Saúde Debate**. Rio de janeiro, v. 43, N. Especial 2, p. 147-154, nov 2019.

STAGE, A. R. Prognosis model for stand development. **USDA For. Serv. Res. Pap. INT-137**. Washington, DC, 1973. 32 p.

STATSOFT, INC. **Statistica (data analysis software system)**, version 12.0. 2013.

TAQUARY, E. C. **Deep learning para identificação precisa de desmatamentos através do uso de imagens satelitárias de alta resolução.** 2019. Dissertação (Mestrado em Ciência da Computação). Universidade Federal de Goiás, Goiás. 2019.

TAQUETTI M. F. **Método da predição de parâmetros em modelos de crescimento em função do índice de local.** (Monografia) - Universidade Federal de Viçosa. Viçosa Minas Gerais. 2014.

VANCLAY, J. K. Growth models for tropical forests: a syntesis of models and methods. **Forest Science**, Bethesda, v. 41, n. 1, p. 7-42, 1995.

VEDRUSCOLO, D. G. S.; *et al.* Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. **Revista Brasileira de Biometria**, São Paulo, v.33, n.4, p.556-569, 2015.

VEGA, A. E. D. **Modelagem de povoamentos de eucalipto no nordeste da Bahia utilizando regressão e redes neurais artificiais.** 2018. Dissertação (Mestrado) – Universidade Federal de Viçosa, Viçosa, 2018.

WYKOFF, W. R. A basal area increment model for individual conifers in the northern rocky montains. **Forest Science**, Bethesda, v. 36, n. 4, p. 1077-1104, 1990.

WYKOFF, W. R.; CROOKSTON, N. L.; STAGE, A. R. **User's guide to the stand prognosis model.** [S.1.]: USDA Forest Service, 1982. 112 p. (General Technical Report., INT – 133).

XING, W.; DU, D. Dropout prediction in MOOCs: Using deep learning for personalized intervention. **Journal of Educational Computing Research**, v. 57, n. 3, p. 547-570, 2019.