

GIANMARCO GOYCOCHEA CASAS

**MODELAGEM DO CRESCIMENTO E DA PRODUÇÃO DE
POVOAMENTOS DE EUCALIPTO UTILIZANDO REGRESSÃO, REDES
NEURAS ARTIFICIAIS E REDES PROFUNDAS**

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

G724m
2021 Goycochea Casas, Gianmarco, 1991-
Modelagem do crescimento e da produção de povoamentos
de eucalipto utilizando regressão, redes neurais artificiais e redes
profundas / Gianmarco Goycochea Casas. – Viçosa, MG, 2021.
100 f. : il. (algumas color.) ; 29 cm.

Orientador: Helio Garcia Leite.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Florestas - Manejo. 2. Estruturas de dados (Computação).
3. Aprendizado do computador. 4. Inteligência artificial.

I. Universidade Federal de Viçosa. Departamento de Engenharia
Florestal. Programa de Pós-Graduação em Ciência Florestal.

II. Título.

CDO adapt. CDD 634.92285

GIANMARCO GOYCOCHEA CASAS

**MODELAGEM DO CRESCIMENTO E DA PRODUÇÃO DE
POVOAMENTOS DE EUCALIPTO UTILIZANDO REGRESSÃO, REDES
NEURAIS ARTIFICIAIS E REDES PROFUNDAS**

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: 15 de fevereiro de 2021.

Assentimento:



Gianmarco Goycochea Casas
Autor



Helio Garcia Leite
Orientador

DEDICO

A Deus, todo poderoso, pela vida.

Minha mãe, Mirtha Casas, e pai, Victor Goycochea (*in memoriam*)

Meu irmão, Paul Goycochea.

Meus avos Justo Casas (*in memoriam*) e Irma Perez.

AGRADECIMENTOS

À Universidade Federal de Viçosa (UFV) e ao Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG, pela oportunidade e pelo apoio financeiro para a realização deste trabalho, em colaboração da Fundação Arthur Bernardes – FUNARBE.

Ao professor Helio Garcia Leite, pela orientação, paciência e dedicação continua, além do altíssimo profissionalismo que demonstrou durante o período deste e outros trabalhos.

Aos funcionários do Departamento de engenharia florestal, os secretários da pós-graduação Josué Dilson e Alexandre, ao porteiro Chiquinho, aos trabalhadores de limpeza da área do departamento.

Aos funcionários da Diretoria de Relações Internacionais (DRI) da UFV, pela atenção e tramitação no processo da documentação internacional.

Ao Departamento de Letras, pelo curso de português para estrangeiros, à professora Idalena, pela contribuição e grande trabalho que faz com os estrangeiros.

Aos colegas do curso de pós-graduação na área de manejo florestal, Fardin, Lucas, Gabi, Simone, Carlos, Ricardo, Kemely, Debora, muito obrigado pelo apoio e incentivo.

A Merly Rocio, amiga e colega de trabalho, pela convivência e profissionalismo.

Ao colega Rodrigo Baselly, uns dos poucos pesquisadores na área de manejo florestal em Perú.

Ao professor Alfredo Huertas, seu filho Alfredinho com esposa Ana Paula e seu filho Lucio, aos colegas Alex Vergara e Arlan Toro, obrigado pela companhia.

Ao grupo de dança UFV, Contato e Improvisação no programa de Vivencias Corporais (2019), professor Josemateus, os colegas, Cristofer, Raíssa, Estevão, Ana, Gabriel, Lucas, Cafu, Arthur e Nayara, pela amizade e momentos da vida.

Ao grupo Taichi Chuan UFV (2019), Mestre Roberto, pela contribuição desta técnica.

Aos funcionários de cuidado das áreas da UFV, em especial às áreas verdes e estrutura civil, à natureza em geral, céu, lua, sol, aves, nuvens, tornando a universidade a mais linda.

A minha família, pais, irmão, tios, Silva Casas, Ramon Burga, Edgar Casas, Maritza, Walter casas, Luz (*in memoriam*), aos meus primos Luis e Ramon Burga, Eric e Evelyn Casas, Walter e Toño Casas.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Às demais pessoas que contribuíram neste trabalho.

"O importante é não parar de questionar. A curiosidade tem a sua própria razão para existir".

Albert Einstein

BIOGRAFIA

GIANMARCO GOYCOCHEA CASAS, filho de Victor Segundo Manuel Goycochea Novak e de Mirtha Irma Casas de Goycochea, nasceu em Lima, Perú, em 20 de junho de 1991.

Conclui o curso primário e secundário na escola 7037 Ariosto Matellini Espinoza, Chorillos, Perú.

Em 2011, ingressou-se no curso de engenharia florestal da Universidad Nacional de Cajamarca, Jaén, Perú, graduando-se em 2015.

Em julho de 2014, fez estágio na Universidad Nacional de Loja, Equador e em agosto do mesmo ano, fez estágio na Universidade Federal de Viçosa, Brasil, com ênfases em anatomia da madeira.

Em 2016, de março até julho, continuo de forma voluntaria com programas de estágio no sector de manejo florestal na Universidade Federal de Viçosa, Brasil.

Em 2017, trabalhou em pesquisas na área de Recursos Florestais e Engenharia Florestal.

Em agosto de 2018, retorna a Viçosa para ser parte do Programa de Pós-graduação em Ciência Florestal, na Universidade Federal de Viçosa, submetendo-se a defesa da dissertação em 15 de fevereiro de 2021.

RESUMO

GOYCOCHEA CASAS, Gianmarco, M.Sc., Universidade Federal de Viçosa, fevereiro de 2021. **Modelagem do crescimento e da produção de povoamentos de eucalipto utilizando regressão, redes neurais artificiais e redes profundas.** Orientador: Helio Garcia Leite.

Modelos de densidade variável sempre foram ajustados com dados de parcelas permanentes organizados em um banco de dados pareados em idades sequenciais i e $i+1$, sendo $i = 1, 2, \dots, N$ medições das parcelas. Para povoamentos equiâneos, de eucalipto, pinus e teca, sempre houve predomínio de intervalos de cerca de um ano de idade entre as medições das parcelas. Quando as redes neurais artificiais (RNA) começaram a ser empregadas como alternativa para modelagem de crescimento e produção surgiu a necessidade de organizar bancos de dados com intervalos irregulares, por exemplo, $i, i+1; i, i+2; \dots; i, N$. Isto foi necessário para que as redes pudessem prever a produção futura com diferentes idades iniciais e intervalos de predição, necessário para uso dos modelos em planejamento hierárquico. Como consequência surgiu a necessidade de verificar a melhor maneira de comparar estimativas geradas com modelos usuais de regressão, com predições feitas por RNA. Além de detalhes relativos à separação de dados para validação e treino e métodos estatísticos a serem empregados, surgiu a necessidade de compatibilizar as validações em relação à forma de organização dos bancos de dados. O trabalho foi desenvolvido em dois capítulos com o objetivo geral de comparar os métodos de regressão, redes neurais artificiais clássicas e redes profundas, na modelagem de crescimento e produção em povoamentos de eucalipto e verificar a exatidão de prognose a partir de idades precoces de aproximadamente 2 anos. No capítulo I, se analisaram três alternativas de organização dos dados, para ajuste dos modelos de Clutter e de Buckman. A primeira com dados emparelhados em idades de medição sequenciais i e $i+1$, sendo $i = 1, 2, \dots, N$ medições das parcelas. Na segunda foram considerados todos os intervalos de medição possíveis para cada parcela: $i, i+1; i, i+2; \dots; i, N$;

$i+1, i+2; \dots, N-1, N$. A terceira com os dados emparelhados mês (j) a mês, sempre com intervalo de um mês, ou seja, $j, j+1; j+1, j+2; j+M-1, M$, sendo M a idade da última medição, em meses. Foram avaliadas as projeções volumétricas e as idades técnicas de corte. Para o modelo de Clutter os resultados mostraram que a exatidão e consistência das projeções depende da forma de organização dos dados e para o ajuste do modelo de Buckman os resultados indicaram a necessidade de uso da alternativa 3, com observações mês a mês. No capítulo II, se analisaram as projeções volumétricas nas idades de projeção de 6 e 7 anos ao longo das idades precoces. Foram avaliadas e comparadas as técnicas: método de mínimos quadrados (modelo de densidade variável), redes neurais artificiais clássicas e redes neurais artificiais profundas. Se utilizou uma extensa e abrangente base de dados de parcelas permanentes instaladas e medidas em povoamentos de eucalipto, organizados pela alternativa 2. Se testou algumas configurações de RNA e se ajustou o modelo de Clutter por mínimos quadrados em dois estágios. Os resultados mostraram que as três técnicas geram estimativas consistentes, com um pequeno ganho de exatidão ao empregar redes profundas. Contudo, se pode concluir que para o emprego das técnicas citadas a forma de organizar os dados em intervalos diferentes de medição influencia na prognose e permanece-se ainda a dificuldade de aumentar a exatidão de projeções feitas a partir de idade muito precoce, mesmo com o emprego de redes profundas.

Palavras-chave: Manejo florestal. Estrutura de dados. Aprendizado de máquina. Inteligência artificial.

ABSTRACT

GOYCOCHEA CASAS, Gianmarco, M.Sc., Universidade Federal de Viçosa, February, 2021. **Growth and yield modelling of eucalypt stands using regression, neural network and deep learning.** Adviser: Helio Garcia Leite.

Variable density models were always fitted with data from permanent plots, organized in a paired database at sequential ages i and $i+I$, where $i = 1, 2, \dots, N$ plot measurements. Equianx, eucalyptus, pine, and teak stands, predominantly showed an interval of about one year of age between plot measurements. When artificial neural networks (ANN) began to be employed to model growth and production, the need arose to organize databases with irregular intervals, for example, $i, i+1; i, i+2; \dots; i, N$. This enabled networks to predict future production with different initial ages and prediction intervals, as required by models for hierarchical planning. Thus, it became necessary to identify the best way to compare estimates generated using regression models with predictions made by ANNs. Statistical methods were employed to separate data, for validation and training. The validations had to be made compatible with the data organization forms. The work was developed in two chapters, with the objective of comparing results using regression methods, and classical and deep artificial neural networks, to model growth and production in eucalyptus stands and to verify the accuracy of early age prognosis of approximately two years. In Chapter I, we analyzed three data organization alternatives for fitting Clutter and Buckman models. In the first, we paired data at sequential measurement ages i and $i+I$, where $i = 1, 2, \dots, N$ plot measurements. In the second, we considered all possible interval measures for each plot: $i, i+1; i, i+2; \dots; i, N; i+I, i+2; \dots, N-1, N$. In the third, we paired data from month (j) to month, with a constant interval of one month, i.e., $j, j+1; j+1, j+2; j+M-1, M$, where M is the age of the last measurement, in months. We evaluated the volumetric projections and technical cutting ages. For the Clutter model, the results show that the accuracy and consistency of the projections depend on the data organization method and for the Buckman model, the results indicate the need to use the

third alternative, with month-to-month observations. In Chapter II, we analyzed volumetric projections at six and seven years of age along with the early ages. We evaluated and compared techniques using the least squares method (variable density model), classical and deep artificial neural networks. We used an extensive and comprehensive database of permanent plots, installed and measured in eucalyptus stands, and organized using the second alternative. We tested ANN configurations and adjusted Clutter's two-stage least squares model. The results were consistent using the three techniques, with a small increase in accuracy when employing deep networks. We conclude that while using the techniques mentioned above, the method used to organize the data in different interval measures influences the predictions. Even with the use of deep networks, it is difficult to increase the accuracy of projections made from a very early age.

Keywords: Forest management. Data structure. Machine learning. Artificial intelligence.

SUMÁRIO

	Página
INTRODUÇÃO GERAL.....	14
REFERÊNCIAS BIBLIOGRÁFICAS	17
CAPÍTULO I	20
AVALIAÇÃO DE DOIS MODELOS DE DENSIDADE VARIÁVEL AJUSTADOS COM DIFERENTES ESTRUTURAS DE BANCO DE DADOS	20
1. INTRODUÇÃO.....	20
2. REVISÃO BIBLIOGRÁFICA.....	24
2.1. Crescimento e produção.....	24
2.2. Funções de crescimento na Engenharia Florestal	25
2.3. Modelos de crescimento e produção	26
3. MATERIAL E MÉTODOS.....	30
3.1. Origem e organização dos dados.....	30
3.2. Índices de sítio e ajuste dos modelos de crescimento e produção.....	32
4. RESULTADOS	36
5. DISCUSSÃO	46
6. CONCLUSÕES	49
7. REFERÊNCIAS BIBLIOGRÁFICAS	49
CAPITULO II.....	59

MODELAGEM DO CRESCIMENTO E DA PRODUÇÃO DE POVOAMENTOS DE EUCALIPTO UTILIZANDO REDES NEURAS ARTIFICIAIS CLÁSSICAS E REDES PROFUNDAS.....	59
1. INTRODUÇÃO.....	59
2. REVISÃO BIBLIOGRÁFICA.....	61
2.1. Conceitos básicos sobre redes neurais artificiais	61
2.2. Principais funções de ativação	63
2.3. Arquitetura de redes neurais artificiais	64
2.4. Algoritmos em redes neurais artificiais.....	68
2.5. Redes neurais artificiais profundas	69
3. MATERIAL E MÉTODOS.....	72
3.1. Origem e organização dos dados.....	72
3.2. Índices de sítio e ajuste dos modelos de crescimento e produção.....	73
3.3. Configuração e treinamento das redes neurais artificiais.....	74
4. RESULTADOS	80
4.1. Avaliação dos modelos	81
4.2. Avaliação das projeções volumétricas a partir de idade precoce	83
5. DISCUSSÃO.....	87
6. CONCLUSÕES	91
7. REFERENCIAS BIBLIOGRÁFICAS	91
CONCLUSÕES GERAIS.....	100

INTRODUÇÃO GERAL

Os modelos em nível de povoamento (MNP) predominam no Brasil, em função do uso da madeira, prevalecendo a produção de polpa de celulose ou de carvão vegetal (CAMPOS e LEITE, 2017). Na década de 1980, Trévizol Júnior (1985) ajustou o modelo de Clutter para a espécie de *Eucalyptus grandis*, sendo esse um dos primeiros estudos com o esse modelo no Brasil. A partir daquele momento, os modelos de densidade variável foram utilizados com grande importância para o sector de manejo florestal brasileiro. Estes modelos são ajustados com uma estrutura no banco de dados proveniente do inventário florestal contínuo, no qual os intervalos são aproximadamente fixos entre as medições sucessivas, ou seja, em um banco de dados pareados em idades sequenciais i e $i+1$, sendo $i = 1, 2, \dots, N$ medições das parcelas.

Por enquanto, até o termino deste estudo, não foram encontrados trabalhos relacionados aos modelos de crescimento e produção de Clutter (1963) e de Buckman (1962) com uma estruturação de dados diferente ao usual. Por outro lado, os autores Goelz e Burk (1992) e Huang (2002) mencionaram que a organização dos dados com todos os intervalos possíveis apresentava resultados mais consistentes para modelagem da altura dominante. Além, dessa organização, existem outras formas de organização dos dados, por exemplo: I) o maior intervalo ascendente, II) o maior intervalo, III) todos os intervalos ascendentes e sem sobreposição, IV) intervalos sem sobreposição e V) todos os intervalos possíveis ascendentes. Assim, muitos autores ajustaram modelos para a classificação da capacidade produtiva em que a variável dependente era a altura dominante com diferentes estruturas de dados, demonstrando sua influência nas avaliações estatísticas (HUANG, 1997; BAILEY e CIESZEWSKI, 2000; HIRIGOYEN et al., 2018).

As redes neurais artificiais (RNA) têm sido cada vez mais utilizadas para diversas finalidades na ciência florestal, mas quando começaram a ser empregadas como alternativa para modelagem de crescimento e produção surgiu a necessidade de organizar bancos de

dados com intervalos irregulares, por exemplo, $i, i+1; i, i+2; \dots; i, N$ medições das parcelas. Um exemplo de uso desse tipo de organização dos dados é o trabalho conduzido por Binoti et al., (2015). Isto foi necessário para que as redes pudessem prever a produção futura com diferentes idades iniciais e intervalos de predição, necessário para uso dos modelos em planejamento hierárquico.

Apesar do ganho de exatidão com o emprego de RNA ainda persiste o problema da ineficiência da projeção a partir de cerca de dois anos de idade (no caso de povoamentos de eucalipto). Considerando a amplitude relativamente pequena das variáveis altura dominante (Hd), área basal (B) e volume (V) em idades iniciais (2 a 3 anos), uma possibilidade seria usar estratificação máxima e ajustar de modelos mais simples (CAMPOS e LEITE, 2017). Tais modelos quando aplicados para cada estrato específico resultaria em diferenciação nas projeções. Entretanto esse maior erro nas projeções a partir de idades muito precoces ainda ocorre e isto pode ser visto em diversos estudos publicados recentemente, por exemplo, em Dávila Vega (2018).

O aprendizado profundo (*deep learning*) faz parte de um campo mais amplo de aprendizado de máquina e inteligência artificial que usa redes neurais artificiais (BHARATENDRA, 2019). A diferença em relação a uma rede convencional é que as redes profundas permitem que modelos computacionais compostos por várias camadas de processamento aprendam representações de dados com vários níveis de abstração (LECUN et al., 2015; NIELSEN, 2015; GOODFELLOW et al., 2016; AGGARWAL, 2018).

O aprendizado profundo tem sido cada vez mais utilizado em pesquisas envolvendo estruturas complexas, em diferentes áreas ambientais e naturais (SUN et al., 2017; HAMDI et al., 2019; Xi et al., 2020; KENTSCH et al., 2020; FERREIRA et al., 2020). Desse modo, é importante verificar a possibilidade de aumentar a exatidão de prognose feita a partir de idades iniciais precoces, em povoamento de eucalipto, utilizando esta técnica. Para o emprego desta técnica ou de uma rede neural artificial clássica é também importante investigar se a forma de organizar os dados tem efeito sobre a referida prognose.

No presente estudo foram avaliadas três alternativas de organização dos dados de parcelas permanentes de inventário de eucalipto para fins de modelagem do crescimento e da produção. Na primeira os dados foram emparelhados em idades de medição sequenciais i e

$i+1$, sendo $i = 1, 2, \dots, N$ medições das parcelas. Na segunda foram considerados todos os intervalos de medição possíveis para cada parcela: $i, i+1; i, i+2; \dots; i, N; i+1, i+2; \dots, N-1, N$. E na terceira alternativa os dados foram emparelhados mês (j) a mês, sempre com intervalo de um mês, ou seja, $j, j+1; j+1, j+2; j+M-1, M$, sendo M a idade da última medição, em meses. Esses três bancos de dados foram utilizados para ajuste dos modelos de Clutter (1963) e de Buckman (1962), sendo este último na versão modificada por Silva (2006) e, também, para o treinamento de redes neurais artificiais clássicas e de redes profundas.

O objetivo geral do trabalho foi de comparar os métodos de regressão, redes neurais artificiais clássicas e redes profundas, na modelagem de crescimento e produção em povoamentos de eucalipto e verificar a exatidão de prognose a partir de idades precoces utilizando esses métodos e três alternativas de organização dos dados. Os objetivos específicos foram:

- i) Ajustar os modelos Clutter e Buckman utilizando três alternativas de organização dos dados e avaliar a exatidão das estimativas de produção geradas ao aplicar as equações.
- ii) Avaliar a eficiência de redes neurais profundas na projeção da produção de povoamentos de eucalipto, em relação ao emprego de redes neurais clássicas;
- iii) Avaliar a possibilidade de ganho de exatidão em projeções de produção de eucalipto feitas a partir de aproximadamente 2 anos com emprego dos modelos, redes e abordagens para organização dos dados.

Neste estudo, foram formuladas as seguintes hipóteses:

H_{o1} : a exatidão das estimativas do modelo de Clutter independe da forma de organização dos dados das parcelas permanentes *versus* H_{a1} : não H_{o1} ;

H_{o2} : a exatidão das estimativas do modelo de Buckman independe da forma de organização dos dados das parcelas permanentes *versus* H_{a2} : não H_{o2} ;

H_{o3} : a inexatidão de produções obtidas com projeções feitas para idades de 6 e 7 anos em povoamentos de eucalipto, a partir de idade precoce (cerca de 2 anos), independe do método utilizado: regressão, redes neurais artificiais ou redes profundas *versus* H_{a3} : não H_{o3} .

REFERÊNCIAS BIBLIOGRÁFICAS

AGGARWAL, C. C. **Neural networks and deep learning**. Springer, 2018.

BAILEY, R. L.; CIESZEWSKI, C. J. Development of a well-behaved site-index equation: jack pine in north-central Ontario: comment. **Canadian Journal of Forest Research**, v. 30, n. 10, p. 1667, 2000.

BHARATENDRA, R. **Advanced deep learning with R become an expert at designing, building, and improving advanced neural network models using R**. B3 2PB, UK. Birmingham, 2019.

BINOTI, M. L. M. D. S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, n. 1, p. 97-105, 2015.

BUCKMAN, R. E. **Growth and yield of red pine in Minnesota**. US Dept. of Agriculture, 1962.

CAMPOS, J.C.C.; LEITE, H.G. **Mensuração florestal: perguntas e respostas**. 5.ed. atual. e ampl. Viçosa: Ed. da UFV, 2017. 636p.

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

DÁVILA VEGA, A. E. **Modelagem de povoamentos de eucalipto no nordeste da Bahia utilizando regressão e redes neurais artificiais**. 2018. 52f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa. 2018.

FERREIRA, M. P. et al. Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. **Forest Ecology and Management**, v. 475, p. 118397, 2020.

GOELZ, J. C. G.; BURK, T. E. Development of a well-behaved site index equation: jack pine in north central Ontario. **Canadian Journal of Forest Research**, v. 22, n. 6, p. 776-784, 1992.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A.; BENGIO, Y. **Deep learning**. Cambridge: MIT press, 2016.

HAMDI, Z. M; BRANDMEIER, M.; STRAUB, C. Forest damage assessment using deep learning on high resolution remote sensing data. **Remote Sensing**, v. 11, n. 17, p. 1976, 2019.

HIRIGOYEN, A.; FRANCO, J.; DIÉGUEZ, U. Modelo dinámico de rodal para Eucalyptus globulus (L.) en Uruguay. **Agrociencia Uruguay**, v. 22, n. 1, p. 63-80, 2018.

HUANG, S. Validating and localizing growth and yield models: procedures, problems and prospects. In: **Proceedings of IUFRO Workshop on Reality, models and parameter estimation: the forestry scenario**. Portugal: Sesimbra, 2002.

HUANG, S. **Development of a subregion-based compatible height-site index-age model for black spruce in Alberta**. Alberta Environmental Protection, Land and Forest Service, Forest Management Division, 1997.

KENTSCH, S.; LOPEZ CACERES, M. L.; SERRANO, D.; ROURE, F.; DIEZ, Y. Computer vision and deep learning techniques for the analysis of drone-acquired forest images, a transfer learning study. **Remote Sensing**, v. 12, n. 8, p. 1287, 2020.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, v. 521, n. 7553, p. 436-444, 2015.

NIELSEN, M. A. **Neural networks and deep learning**. San Francisco, CA: Determination press, 2015.

SILVA, A. L.; CAMPOS, J. C. C.; LEITE, H. G.; SOUZA, A. L.; LOPES, P. F. Growth and yield prediction using the modified Buckman model. **Revista Árvore**, vol. 30, n. 5, p. 787-793, 2006.

SUN, Y.; LIU, Y.; WANG, G.; ZHANG, H. Deep learning for plant identification in natural environment. **Computational intelligence and neuroscience**, v. 2017, 2017.

TREVIZOL JÚNIOR, T. L. **Análise de um modelo compatível de crescimento e produção em plantações de Eucalyptus grandis (W.Hill ex-Maiden)**. 1985. 74 f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, 1985.

XI, Z.; HOPKINSON, C.; ROOD, S. B.; PEDDLE, D. R. See the forest and the trees: Effective machine and deep learning algorithms for wood filtering and tree species classification from terrestrial laser scanning. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 168, p. 1-16, 2020.

CAPÍTULO I

AVALIAÇÃO DE DOIS MODELOS DE DENSIDADE VARIÁVEL AJUSTADOS COM DIFERENTES ESTRUTURAS DE BANCO DE DADOS

1. INTRODUÇÃO

A modelagem do crescimento e da produção de povoamentos equiâneos é parte essencial do manejo florestal (DAVIS e JOHNSON, 1987; CAMPOS e LEITE, 2017). Os modelos ajustados são utilizados para estimar estoques de colheita para diferentes prescrições alternativas e assim subsidiar os planos de manejo, especialmente o planejamento hierárquico. Essa modelagem é importante também para determinar a produtividade média em diferentes sítios, sendo esta informação importante em estudos relacionados com a sustentabilidade da produção.

Os modelos de crescimento e produção podem ser explicitamente de predição ou de projeção. No primeiro caso o crescimento e a produção futura independem da produção atual; no segundo, a produção atual (em área basal e, ou, volume) é utilizada como uma variável preditora (variável independente) nas relações funcionais que compõem o modelo (CAMPOS e LEITE, 2017). Conforme esses autores, um modelo de predição pode ser transformado e aplicado de modo que a produção futura seja diferenciada em função de produções atuais, porém, com uma mesma tendência de crescimento.

O processo de estimar estoques de madeira para idades definidas em um modelo de planejamento estratégico é conhecido como *prognose* (CAMPOS e LEITE, 2017). Esse termo é usado como referência a qualquer procedimento que gere como resultado uma expectativa de produção futura. Conforme Burkhart e Tomé (2012) e Castro et al. (2013), a

predição descreve a mudança de tamanho do indivíduo no povoamento com o tempo. Entretanto, uma maneira mais objetiva e consistente de diferenciar predição e projeção é que no segundo caso a produção atual é uma variável independente e no primeiro ela é desconsiderada.

Assumindo uma relação funcional $v=f(t)$ e a transformação $v_2 = v_1 f(t_2)/f(t_1)$, em que v é a produção, t é a idade de medição e 1 e 2 indicam idades atual e futura, então: $v_2 = v_1 f(t_2)/f(t_1) + \varepsilon$ é um modelo estatístico de projeção e $v_2 = v_1 f(t_2)/f(t_1)$, obtido após o ajuste de $v=f(t)$, por transformação matemática simples, é um modelo de predição transformado para permitir projetar a produção assumindo mesma tendência de crescimento em todos os talhões (CAMPOS e LEITE, 2017), ou seja, curvas de produção anamórficas.

Conforme Campos e Leite (2017), dentre outros, a modelagem do crescimento e da produção pode ser feita utilizando parcelas ou árvores individuais como unidade de modelagem. No primeiro caso tem-se os modelos em nível de povoamento (MNP) e os de distribuição de diâmetros (MDD) e no segundo os modelos de árvores individuais (MAI). Esta classificação aparece nos livros clássicos de mensuração florestal, desde os mais antigos até os que foram publicados atualmente, em diversos países.

Da mesma forma a discussão e inferências sobre a escolha de uma categoria de modelo está sempre presente nos livros-texto, de hoje e de muitos anos atrás. Campos e Leite (2017), por exemplo, reforçam o que já foi mencionado por muitos autores ao longo das décadas: a escolha de um tipo de modelo depende dos objetivos do manejo e dos dados disponíveis. MAI e MDD são modelos úteis quando o manejo visa produzir multiprodutos da madeira incluindo referência ao tamanho (comprimento e diâmetros) das toras. Nesse caso geralmente o manejo requer aplicação de desbastes. Por outro lado, os MNP são indicados quando o objetivo é estimar produção por unidade de área, para um ou mais usos da madeira (MENDONÇA et al, 2014; AZEVEDO et al., 2016; SALEKIN et al., 2020; BLANCO et al., 2020).

Os modelos em nível de povoamento (MNP) predominam no Brasil, em função do uso da madeira, na maioria das vezes para produção de polpa de celulose ou de carvão vegetal. Os MNP foram classificados por Campos e Leite (2017) em modelos $Y=f(I)$, $Y=f(I,S)$ e $Y=f(I,S,B, N$ ou $q)$, em que Y é produção (volume ou massa por hectare), S é o índice de

sítio, B é área basal, N é a frequência de árvores por hectare e q é o diâmetro quadrático médio. Em todos os tipos de modelo o ajuste das relações funcionais pode considerar o não a produção atual.

Considerando Y_a e Y_f as produções nas idades atual e futura, I_a e I_f , e X_i outras variáveis independentes, como o índice de sítio, então, conforme mencionado anteriormente, as duas opções mencionadas podem ser assim resumidas: $Y_f = f(I_f X_i)$ e $Y_f = f(Y_a, I_a, I_f, X_i)$. Estas são opções de forma de ajuste uma vez que a aplicação, conforme os autores citados, deve ser feita sempre por $Y_f = Y_a * f(I_f X_i) / f(I_a X_i)$. A segunda opção requer pelo menos duas medições de cada parcela permanente, o que pode resultar em perda de dados no processo de modelagem.

Esses aspectos são discutidos em Campos e Leite (2017) com referência a modelos exponenciais e sigmóides, além dos modelos de Clutter (CLUTTER, 1963) e de Buckman (BUCKMAN, 1962; SILVA et al., 2006). Em suas discussões eles reportam que o uso de relações funcionais mais simples requer uma estratificação mais intensa dos dados para a modelagem. E ainda, que o emprego dos modelos de Clutter e de Buckman requerem a organização do banco de dados com o pareamento das informações considerando as idades de medição sequenciais.

O modelo de Buckman (1962) não havia sido utilizado em escala comercial no Brasil até 2017 (CAMPOS e LEITE, 2017) e provavelmente ainda não foi utilizado até o momento da conclusão do presente estudo. Isto, apesar de diversos estudos mostrarem sua flexibilidade e consistência como modelo de densidade variável (BECK e DELLA-BIANCA, 1972; BURKHART e SPRINZ, 1984; TREVIZOL JÚNIOR, 1985; RESENDE, 1991; SILVA et al., 2006 e GUERA et al., 2019). Por outro lado, o modelo de Clutter sempre foi utilizado em diferentes regiões do Brasil desde a década de 80 (TREVIZOL JÚNIOR, 1985; SOARES et al., 2004; SALLES et al., 2012; MIGUEL et al., 2016; PENIDO et al., 2020).

Na literatura, diferentes autores demonstraram que a avaliação das estimativas e dos parâmetros estatísticos depende tanto do modelo quanto da estrutura de dados utilizada (CAO, 1993; HUANG, 1997; BAILEY e CIESZEWSKI, 2000; CIESZEWSKI, 2002; STRUB e CIESZEWSKI, 2006; HIRIGOYEN et al., 2018). Essas demonstrações foram

feitas para regressões ajustadas para classificação da capacidade produtiva em que a variável dependente era a altura dominante.

Por outro lado, não foram encontrados trabalhos relacionados aos modelos de CLUTTER (1963) e de BUCKMAN (1962). Exemplos de estruturas de dados são: I) o maior intervalo ascendente, II) o maior intervalo, III) todos os intervalos ascendentes e sem sobreposição, IV) intervalos sem sobreposição, V) todos os intervalos possíveis ascendentes e VI) todos os intervalos possíveis, sendo este último apresenta resultados mais consistentes para modelagem da altura dominante (GOELZ e BURK, 1992; HUANG, 2002).

É muito provável que o modelo de Clutter sempre foi ajustado utilizando uma estrutura de dados na qual as informações foram organizadas pareando as informações sempre nas idades i e $i+1$. Por exemplo, se as medições foram feitas aos 2, 3, 4 e 5 anos então, para ajuste do modelo de Clutter os dados foram pareados em 2-3, 3-4 e 4-5, ou seja, quatro registros ou linhas do banco de dados, referentes a uma parcela permanente, são transformados em três registros.

Além de técnicas de regressão, outras técnicas como metaheurística apresentam características próprias na organização do banco de dados: Quando uma metaheurística, como redes neurais artificiais-RNA, é utilizada para modelagem de crescimento e produção, o banco de dados deve ter uma estrutura com todos os intervalos de tempo (medições) variáveis. É necessário utilizar diferentes intervalos para organização dos dados, ou seja, para cada parcela, ao invés de parear as informações em i e $i+1$, $i+1$ e $i+2$, ..., $in-1$ e in , é necessário organizar os dados assim: i e $i+1$, i e $i+2$, ..., i e in , $i+1$ e $i+2$, $i+1$ e $i+3$, ..., $in-1$ e in . Assim, considerando as idades de 2, 3, 4 e 5, resultaria em 2-3, 2-4, 2-5, 3-4, 3-5, 4-5 (6 registros). Isto é necessário porque na aplicação seja possível prever a produção para diferentes idades a partir de uma produção observada numa idade atual. Nos últimos anos cada vez mais tem sido utilizado metaheurísticas e técnicas de aprendizado de máquinas para modelagem do crescimento e produção de povoamentos florestais (ARAÚJO JÚNIOR et al., 2019; SOUZA et al., 2019; LEITE et al., 2020; GAVILÁN-ACUÑA et al., 2021). Uma dessas aplicações é a modelagem do crescimento e da produção (ALCÂNTRA et al., 2018; MEDEIROS et al., 2019; FREITAS et al., 2020).

Como na maioria dos estudos envolvendo o uso de aprendizado de máquinas para prognose são feitas comparações com a modelagem por regressão, surge a seguinte pergunta: existe algum ganho ou perda de eficiência na modelagem com o modelo de Clutter ao empregar as duas abordagens de organização dos dados das parcelas permanentes? No presente estudo foram avaliadas três alternativas de organização da base de dados de parcelas permanentes mensuradas em povoamentos de eucalipto, para ajuste dos modelos de Clutter (1963) e de Buckman (1962), sendo este último na versão modificada por Silva (2006).

2. REVISÃO BIBLIOGRÁFICA

2.1. Crescimento e produção

Os termos crescimento e produção florestal definem diferentes aspectos da dinâmica de desenvolvimento dos povoamentos florestais. Estudos de crescimento e produção são importantes para propósito de manejo florestal (BURKHART, 1979; CAMPOS, 1980; CRUZ et al., 2008). Conforme Davis e Johnson (1987) e Campos e Leite (2017), o estudo do crescimento e da produção está inserido em um dos elementos essenciais do manejo florestal, que é a predição ou projeção de estoques de colheita ao longo de horizontes de planejamento estabelecidos para fins de regulação da produção.

O crescimento longitudinal de todas as partes da árvore é do meristema primário e o crescimento em diâmetro é originário do meristema secundário ou cambio (FINGER, 1992; SCOLFORO, 1998; DRESCHER, 2004). Assim, o *crescimento* é o aumento de dimensões de um ou mais indivíduos em um determinado período de tempo, sendo dimensões de variáveis como o diâmetro, altura, volume, massa, área basal, entre outros (VANCLAY, 1994). A *produção* é o crescimento acumulado até uma certa idade (HUSCH et al., 2002; WEISKITTEL et al., 2011).

Outros termos usuais em estudos de crescimento são, conforme Campos e Leite (2017): *incremento corrente* (aumento da produção no período de um ano ou um mês), *incremento médio* (taxa média de aumento da produção até uma certa idade), *incremento*

periódico (diferença entre produções observadas em duas idades quaisquer) e *incremento periódico anual*, que é o incremento periódico dividido pelo intervalo de tempo entre as duas produções mencionadas.

Os tipos de crescimento podem ser diferenciados em função da aplicação de desbastes e da ocorrência de mortalidade. Conforme Davis e Johnson (1987) e exemplos apresentados em Campos e Leite (2017), se não há mortalidade (V_M) então a produção líquida (Y_l) é igual a produção bruta (Y_b) e o crescimento líquido (C_l) é igual ao crescimento bruto (C_b). Ainda, sendo V_1 e V_2 as produções observadas nas ocasiões atual I_1 e futura I_2 , então, $Y_b = V_2 + V_d + V_M$, $Y_l = Y_b - V_M$, $C_b = V_2 - V_1 + V_d + V_M$ e $C_l = C_b - V_M$. Conhecer esses componentes e tipos de crescimento é importante para o estabelecimento de planos de manejo e análise de problemas de decisão sobre aplicação de práticas silviculturais como o desbaste.

2.2. Funções de crescimento na Engenharia Florestal

Existem várias aplicações de funções de crescimento na Engenharia Florestal (KIVISTE e CONZÁLEZ, 2002): estimar crescimento de árvores individuais (diâmetro, altura, área seccional, volume, densidade e massa); estimar o crescimento do povoamento (altura dominante, diâmetro quadrático médio, área basal, volume e massa). Ainda segundo os autores, as estimativas de crescimento e produção são utilizadas, dentre outros usos, para estimar a produtividade média, idade técnica de corte e taxas de crescimento.

Funções de crescimento são utilizadas também para estudar relações entre variáveis das árvores e dos povoamentos; por exemplo, para modelagem o afilamento do fuste (KOZAK et al., 1969; DEMAERSCHALK, 1972; GARAY, 1979; PIRES e CALEGÁRIO, 2007), relações hipsométricas (BATISTA et al., 2001; MORAES NETO et al., 2010; LEITE et al., 2011; ANDRADE et al., 2018) e razões entre produções para diferentes usos da madeira (CAMPOS e LEITE, 2017).

A evolução no tamanho (diâmetro, altura e volume) de uma árvore ao longo de sua vida segue uma tendência sigmoideal, descrita por um modelo sigmoide. Esse modelo tem duas características fundamentais: possui um ponto de inflexão e uma assíntota horizontal (KIVISTE e CONZÁLEZ, 2002; CAMPOS e LEITE, 2017).

Ao empregar modelos sigmóides, além das características gerais de existência de ponto de inflexão e de um valor assintótico da variável dependente, expressa por um dos parâmetros da função, podem ocorrer algumas restrições de acordo com cada tipo de análise, por exemplo (PIENNAAR e TURNBULL, 1973, GOELZ e BURK, 1992; HOSOKAWA et al., 1998): altura (Y) igual a 1,3 m quando o dap (X) for igual zero e diâmetro comercial (Y) igual a zero quando $X = 1$, sendo X a razão entre altura comercial (h) e a altura total (Ht) da árvore.

Quando as funções de crescimento são empregadas para variáveis do povoamento, em certos casos, a tendência deixará de ser sigmoide após certo intervalo de tempo, devido a mortalidade e, ou práticas silviculturais ou prescrições de manejo aplicadas. Por exemplo, assumindo que a altura dominante (Hd) de povoamentos de eucalipto não é influenciada pelo desbaste por baixo (MEDEIROS, 2016), então a tendência entre Hd e idade é sigmoide. Porém, se no desbaste forem eliminadas árvores dominantes, então poderá haver mudança na tendência de crescimento. Sob o ponto de vista biológico é esperada a ocorrência de um valor assintótico para Hd .

Da mesma forma, ao relacionar a produção (m^3ha^{-1}) com a idade, em um povoamento manejado sem aplicação de desbastes, poderá resultar em tendências sigmoide ou exponencial entre Y (produção) e X (idade). Se o corte raso for efetuado antes da ocorrência de “altas” taxas de mortalidade e a idade inicial não for muito tardia, a tendência deverá ser sigmoide; caso contrário a tendência poderá ser exponencial. Isto tem implicações diretas na escolha da função de crescimento ou modelo a ser utilizado para cada base de dados. Conforme Campos e Leite (2017), para selecionar uma função ou modelo é importante considerar os aspectos biológicos, os dados disponíveis, os indicadores estatísticos e o uso pretendido para o modelo.

2.3. Modelos de crescimento e produção

Um modelo de crescimento e produção (MCP) pode ser construído por uma ou mais equações, tabelas e gráficos, utilizadas para estimar crescimento e produção de povoamentos

florestais (CLUTTER et al., 1983; NASCIMENTO, 2016; CAMPOS e LEITE, 2017). Eles podem ser utilizados para simular a dinâmica natural de um povoamento e prever a sua produção ao longo do tempo, em diferentes possibilidades de exploração (VANCLAY, 1994).

Davis e Johnson (1987), classificou os MCP em três categorias: modelos de distribuição de diâmetros (MDD), que permitem estimar frequência e produção por classe de diâmetro, modelos em nível de árvores individuais (MAI), que estimam frequência e produção por árvore individual e modelos em nível de povoamento (MNP), que estimam a produção em nível de povoamento.

Conforme Campo e Leite (2017), as produções obtidas por classe de *dap* (MDD) e por árvore individual (MAI) podem ser somadas para obter a produção em nível de povoamento. Ainda conforme esses autores, independentemente da categoria existem duas abordagens para a modelagem do crescimento e da produção: estratificação mais intensa e emprego de funções de crescimento com poucas variáveis independentes (somente a idade ou a idade e o índice de sitio) e estratificação menos intensa seguida do ajuste de modelos contendo mais variáveis independentes (por exemplo, idade, índice de sítio e área basal).

A escolha de um MCP depende dos objetivos do manejo e dos dados disponíveis (CAMPOS e LEITE, 2017). No Brasil há um predomínio de MNP uma vez que na maioria das vezes o interesse do usuário é estimar apenas um ou dois tipos de sortimentos, em nível de povoamento. Se o objetivo é estimar a produção para diferentes usos da madeira, porém, sem informar sobre dimensões das toras ou sobre multiprodutos, então um MNP é suficiente, desde que sejam incluídas variáveis dummy na modelagem, ou equações de razão entre produções para os diferentes usos (CAMPOS e LEITE, 2017). Isto significa que MAI e MDD serão necessários apenas se houver necessidade de diferenciar a produção por classe de tamanho das toras ou de estimar a produção de multiprodutos da madeira.

Os modelos em nível de povoamento foram aplicados a florestas tropicais há muitas décadas (PORTÉ e BARTELINK, 2002). No Brasil os primeiros registros de uso de MNP são possivelmente da década de 1980 (TREVISOL JÚNIOR, 1985; SOUZA, 1988; LEITE, 1990, SOARES, 1999). Conforme Campos e Leite (2017), no Brasil, em povoamentos de eucalipto, a atualização das estimativas dos parâmetros dos MNP é feita pelo menos a cada dois anos e a principal MNP é o de Clutter (1963).

Os modelos de Clutter e de Buckman foram desenvolvidos na mesma década. Ambos são de densidade variável e foram idealizados com a mesma premissa: a taxa de crescimento em volume depende da taxa de crescimento em área basal, além da idade e do índice de sítio. A diferença entre esses modelos é que enquanto no de Clutter a produção futura em área basal é obtida por projeção direta, no de Buckman a área basal futura é obtida pela soma de incrementos em área basal.

Conforme Campos e Leite (2017), o modelo de Clutter é: **tipo povoamento total** (a variável estimada é o volume por unidade de área, independentemente da classe de tamanho da árvore), **de densidade variável** (permite estimar a produção para diferentes níveis de área basal inicial), **compatível** (a integração analítica do crescimento em área basal resulta na produção e a derivada da produção resulta no crescimento), **explícito** (a produção em volume é calculada diretamente) e **Consistente** (a área basal pode ser projetada ano a ano ou para intervalos diferentes de 1 ano, resultando nas mesmas estimativas de área basal futura). Excluindo as interações entre as três variáveis independentes do modelo de LnV_2 , as relações funcionais que compõem o modelo de Clutter são:

$$LnB_2 = LnB_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) + \alpha_1 \left(1 - \frac{I_1}{I_2} \right) S_1 + \varepsilon$$

$$LnV_2 = \beta_0 + \beta_1 \left(\frac{1}{I_2} \right) + \beta_2 S_1 + \beta_3 LnB_2 + \varepsilon$$

em que:

V_2 = produção futura ($m^3 ha^{-1}$);

I_1 e I_2 = idade atual e futura (meses);

S_1 = índice de sítio na idade atual (m);

B_2 e B_1 = área basal atual e área basal futura ($m^2 ha^{-1}$);

Ln = logaritmo neperiano;

α e β = parâmetros;

ε = erro aleatório

Buckman (1962) pode ter sido o primeiro autor a explicitar uma relação matemática biológica entre crescimento e produção (CAMPOS e LEITE, 2017). Na proposta de Buckman (1962) o modelo de densidade variável era composto por uma equação de incremento em área basal, uma equação para projetar a altura dominante (Hd) e uma equação

de volume de povoamento, $V = f(BHd)$. Silva et al. 2006, modificaram estas relações substituindo a relação funcional $V = f(BHd)$ por aquela definida por Clutter (1963) para LnV_2 , resultando em:

$$Ln dB_2 = \beta_0 + \beta_1 I_2^{-1} + \beta_2 S_1 + \beta_4 Ln B_{1_1} + \varepsilon$$

$$Ln V_2 = \beta_0 + \beta_1 I_2^{-1} + \beta_2 S_1 + \beta_3 Ln B_2 + \varepsilon, B_2 = B_1 + \sum_{j=I_1}^{j=I_2} dB_j$$

Outra modificação importante feita por Silva et al. (2006) foi a substituição da variável B_i por B_1 no modelo de incremento em área basal proposto por Buckman (1962). Esta modificação foi justificada pelo fato da tendência de crescimento em área basal em povoamentos equiâneos ser dependente do peso e idade do desbaste. No modelo de 1962 B_i é a área basal inicial única e na proposta de Silva et al. (2006) o incremento desde uma idade atual I_1 até uma idade futura I_2 (dB_2) é sempre em função da área basal remanescente da idade atual I_1 .

Apesar das melhorias propostas por Silva et al. (2006), há ainda uma certa dificuldade em relação ao intervalo de idades para definição da variável dB . No caso de povoamentos equiâneos nem todas as parcelas permanentes são mensuradas com mesmo intervalo de tempo e isto, dependendo da forma de organização do banco de dados, pode gerar inconsistência quando da aplicação do modelo ajustado. Esses detalhes foram apresentados por Campos e Leite (2017) e uma maneira de evitar a inconsistência é considerar dB como incremento sempre mensal. Outra alternativa é ajustar os intervalos de medição para um ano por meio de interpolações.

Os dados para ajuste dos modelos de Clutter e de Buckman são geralmente obtidos de parcelas permanentes, na maioria das vezes de inventários florestais contínuos. Esses dados são organizados de modo a contemplar as variáveis presentes nas relações funcionais. Portanto, por serem modelos explicitamente de projeção, os dados sequenciais são pareados de alguma forma.

Algumas possíveis estruturas de banco de dados (GOELZ e BURK, 1992; HUANG, 1997, 2002) contemplam, em relação às idades de medição de cada parcela: maior intervalo

ascendente, maior intervalo, todos os intervalos ascendentes e sem sobreposição, intervalos sem sobreposição, todos os intervalos possíveis ascendente, ou todos os intervalos possíveis.

Alguns autores destacam a estrutura de dados com todos os possíveis intervalos de crescimento como a mais adequada (GOELZ e BURK, 1992, 1996; PARRESOL, 1998; HUANG, 1999; ARANDA, 2004; DORADO, 2004). Entretanto, no caso dos modelos de Clutter e de Buckman sempre foi usual contemplar todos os intervalos ascendentes sem sobreposição, ou seja, parear os dados considerando de I_1 para I_2 , de I_2 para I_3 , ..., de I_{n-1} para I_n , sendo n a idade da última medição da parcela (TREVISOL JÚNIOR, 1985; SOUZA, 1988; SOARES et al., 1998, DIAS et al., 2005; SILVA et al., 2006; MIRANDA et al., 2016; CAMPOS e LEITE, 2017).

Bancos de dados organizados com todos os intervalos possíveis ascendente tem sido utilizado para treinamento de redes neurais artificiais (BINOTI et al., 2014; BINOTI et al., 2015; ALCÂNTRA et al., 2018; DÁVILA VEGA, 2018; FREITAS et al., 2020). A validação do modelo de Clutter e da rede neural artificial em tais estudos foi feita com sub-amostras de dados separados para validação, na maioria das vezes constituída de cerca de 30% da totalidade de parcelas disponíveis. Fica, portanto, um questionamento sobre a eficiência das comparações entre esses dois enfoques de modelagem do crescimento e da produção: regressão e RNA (ou outras técnicas de IA).

3. MATERIAL E MÉTODOS

3.1. Origem e organização dos dados

Os dados utilizados neste estudo foram obtidos de 878 parcelas permanentes, com área em torno de 400 m², mensuradas em povoamentos de eucalipto localizados nas regiões Cocais (46 projetos), Sabinópolis (32 projetos) e Virginópolis (31 projetos), em Minas Gerais, Brasil (Figura 1). Em cada parcela e idade de medição foram determinadas as variáveis: altura dominante (m), área basal (m²ha⁻¹) e volume (m³ha⁻¹).

As medições foram feitas em idades que variaram de 19 a 96 meses e os genótipos cultivados foram clones híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis*. Considerando I_j uma idade de medição da parcela, $j = 1, 2, \dots, n$ meses, foram definidas três alternativas de organização do banco de dados, conforme Tabela 1, em que:

- I. Dados pareados considerando os intervalos ascendentes sem sobreposição ($I_1-I_2, I_2-I_3, \dots, I_n - I_{n+1}$). Esta estrutura é historicamente utilizada para os ajustes dos modelos de crescimento e produção de densidade variável e outros modelos de projeção;
- II. Dados pareados considerando todos os intervalos de idades ascendentes possíveis ($I_1-I_2, I_1-I_3, \dots, I_2-I_3, I_2-I_4, \dots, I_n - I_{n+1}, I_n - I_{n+2}$);
- III. Dados pareados considerando os intervalos ascendentes sem sobreposição, porém, mês a mês. Para estabelecer esta estrutura feitas interpolações lineares e utilizados apenas parcelas com 5 ou mais medições. A decisão de usar parcelas com 5 ou mais medições foi para se ter maior consistência nas extrapolações.

Figura 1 - Localização das regiões de área de estudo no estado de Minas Gerais, Brasil

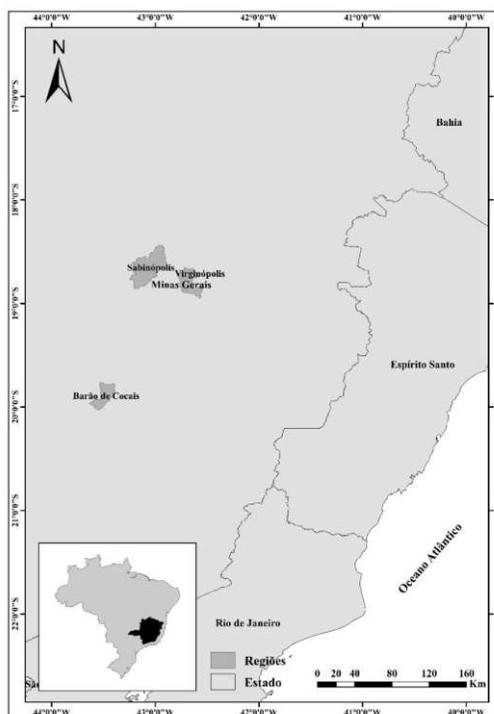


Tabela 1 - Demonstração das estruturas dos bancos de dados, utilizadas para o ajustamento dos modelos de Clutter e de Buckman

Estrutura	Plot	Medição ₁	Medição ₂	Idade ₁ (mês)	Idade ₂ (mês)
I	1	1	2	33	43
	1	2	3	43	55
	1	3	4	55	68
	1	4	5	68	81
	1	5	6	81	91
II	1	1	2	33	43
	1	1	3	33	55
	1	1	4	33	68

	1	4	5	68	81
	1	4	6	68	91
III	1	5	6	81	91
	1	1	2	33	34
	1	2	3	34	35
	1	3	4	35	36

	1	55	56	87	88
	1	56	57	88	89
1	57	58	89	90	

3.2. Índices de sítio e ajuste dos modelos de crescimento e produção

Os índices de sítio de cada parcela em cada medição foram obtidos pelo método da curva-guia (CLUTTER et al., 1983), utilizando o modelo Gompertz e uma idade-índice de 72 meses. Esse modelo foi ajustado para toda a área, sendo gerado um feixe de curvas anamórficas de índices de sítio. A relação funcional ajustada foi $HD = \beta_0 e^{-e^{\beta_1 - \beta_2 I}} + \varepsilon$ e os índices de sítio foram obtidos pela transformação $Hd_2 = Hd_1 f(I_2) / f(I_1)$, em que I_1 é a idade de medição e I_2 a idade-índice, ou seja,

$$S = Hd e^{-e^{\beta_1 - 72\beta_2}} e^{e^{\beta_1 - \beta_2 I}}, \quad (1)$$

em que Hd é a altura dominante (m), S é o índice de sitio (m) e β_i são as estimativas dos parâmetros do modelo Gompertz.

Os modelos de crescimento e produção definidos para este estudo foram o de Clutter,

$$\text{Ln}B_2 = \text{Ln}B_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) + \alpha_1 \left(1 - \frac{I_1}{I_2} \right) S_1 + \varepsilon_1 \quad (2)$$

$$\text{Ln}V_2 = \beta_0 + \beta_1 \left(\frac{1}{I_2} \right) + \beta_2 S_1 + \beta_3 \text{Ln}B_2 + \varepsilon_2 \quad (3)$$

e o de Buckman, modificado por Silva et al. (2006).

$$\text{Ln}dB_2 = \beta_0 + \beta_1 I_2 + \beta_2 S_1 + \beta_3 I_2^{-1} + \beta_4 \text{Ln}B_1 + \varepsilon_1 \quad (4)$$

$$\text{Ln}V_2 = \beta_0 + \beta_1 I_2^{-1} + \beta_2 S_1 + \beta_3 \text{Ln}B_2 + \varepsilon_2, \quad B_2 = B_1 + \sum_{j=I_1}^{j=I_2} dB_j \quad (5)$$

em que:

V_2 = produção futura (m^3ha^{-1});

I_1 e I_2 = idades atual e futura (meses);

S_1 = índice de sitio na idade atual (m);

B_1 e B_2 = áreas basais atual e futura (m^2ha^{-1});

dB_2 = incremento em área basal entre I_1 e I_2 ;

Ln = logaritmo neperiano;

α e β = parâmetros;

ε_1 e ε_2 ; $\varepsilon_k \sim NID(0; \sigma^2)$.

Os dados foram separados em duas sub-amostras de 50% das parcelas. A separação foi feita utilizando o software *Select* 2.0 (BINOTI, 2012), tendo como meta a minimização da raiz quadrada do erro quadrático médio ao ajustar o modelo de Clutter. Foi utilizado o algoritmo de busca aleatória disponível no software, com 100 simulações. Desse modo, 50% das parcelas foram utilizadas para ajuste e 50% para validação. Os ajustes foram feitos para cada uma das três estruturas de dados indicadas na Tabela 1.

O modelo de Clutter foi ajustado pelo método de mínimos quadrados em dois estágios e o de Buckman por mínimos quadrados ordinários. O software EViews 6 (Quantitative Micro Software, 2007) foi utilizado para os ajustamentos e as análises gráficas foram geradas com o software R (R CORE TEAM, 2020) com o pacote ggplot2 (WICKHAM, 2016). Para a obtenção da produção em área basal pelo modelo e Buckman foi utilizado o método de repetidas soluções da função de crescimento (SILVA et al., 2006; CAMPOS e LEITE, 2017).

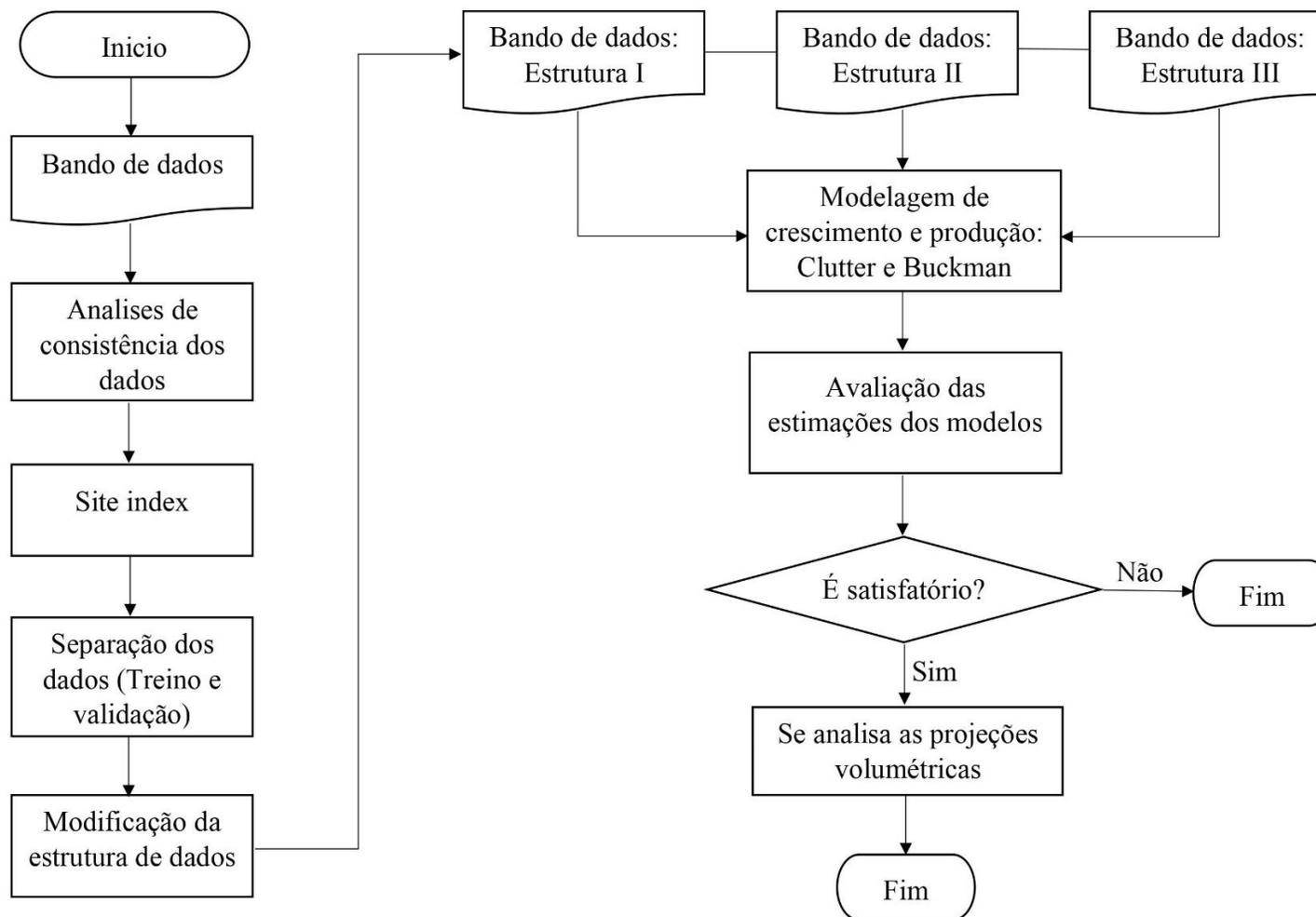
A validação dos modelos ajustados com as diferentes estruturas de dados foi feita pela análise comparativa das estatísticas: coeficiente de determinação (R^2), a média das diferenças absolutas (MDA), raiz quadrada do erro quadrático médio (RQEM%), *Bias* e erro relativo percentual (ER%). Além disso, foram elaborados gráficos de correlação entre produções observadas (Y_i) e estimadas (\hat{Y}_i), ER% em função da produção observada e percentagem de casos por classe de ER%. Para a estrutura de dados III as produções interpoladas foram excluídas da validação. Um fluxograma com a sequência das análises está na Figura 2 e os estimadores utilizados foram:

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (6); \quad MDA = \left(n^{-1} \sum_{i=1}^n |\hat{Y}_i - Y_i| \right) \quad (7); \quad RQEM\% = 100 \bar{Y}_i^{-1} \sqrt{n^{-1} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (8);$$

$$Bias = \sum_{i=1}^n \frac{(\hat{Y}_i - Y_i)}{n} \quad (9); \quad Bias\% = 100 \bar{Y}_i^{-1} \sum_{i=1}^n \frac{(\hat{Y}_i - Y_i)}{n} \quad (10); \quad ER\% = 100 \left(\frac{\hat{Y}_i - Y_i}{Y_i} \right) \quad (11), \text{ em que}$$

n é o número de observações, \hat{Y}_i é a produção estimada, Y_i é a produção observada e \bar{Y} é a média das produções observadas.

Figura 2 - Fluxograma dos procedimentos e avaliações realizados neste estudo



4. RESULTADOS

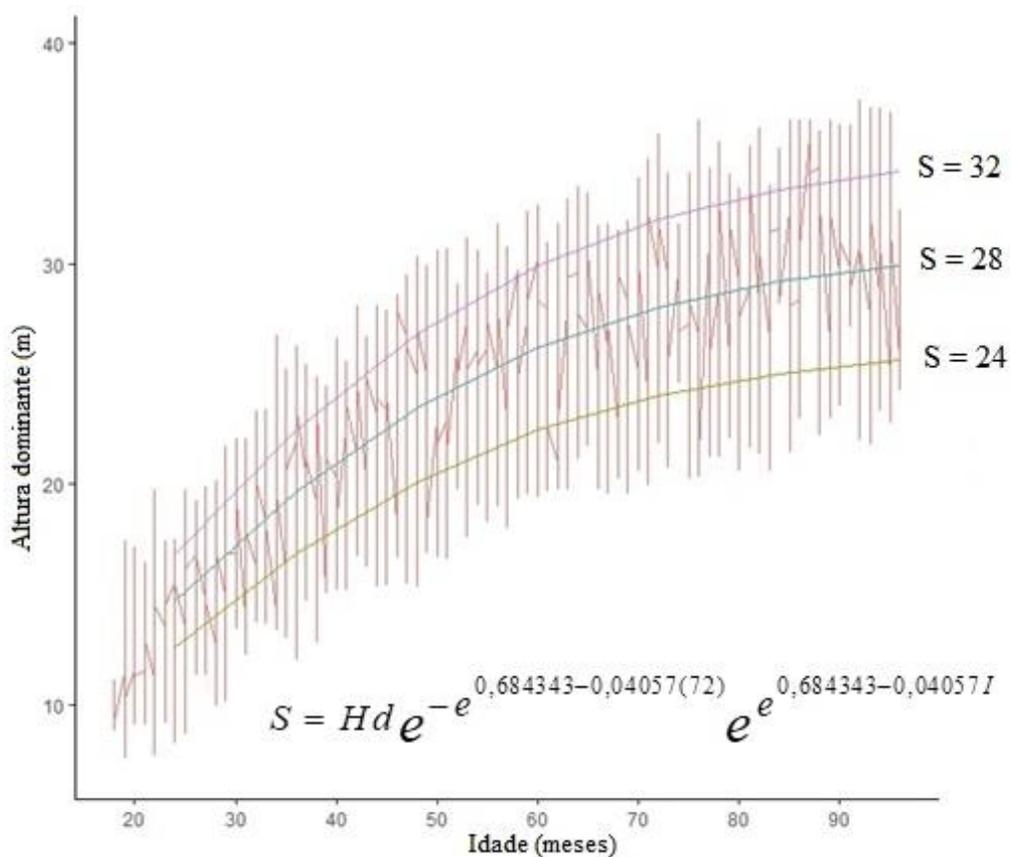
A equação de altura dominante ajustada a toda a base de dados foi:

$$Hd = 30,78586e^{-0,684343-0,04057I}, \text{ com } r_{\hat{y}y} = 0,8813.$$

Os índices de sítio (S) de cada parcela, em cada medição, foram estimados por:

$$S = Hd e^{-0,684343-0,04057(72)} e^{0,684343-0,04057I}, \text{ em que } Hd \text{ é a altura dominante observada na idade } I \text{ (meses). As curvas de índices de sítio podem ser observadas na Figura 3.}$$

Figura 3 - Curvas de índice de sítio para clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG, para idade-índice de 72 meses



As estimativas dos parâmetros das equações dos modelos de Clutter e de Buckman, ajustados para cada estrutura de banco de dados, são apresentadas na Tabela 2. Na mesma tabela constam as estatísticas de precisão e de exatidão obtidas em cada ajustamento. Todos os coeficientes foram significativos ($p < 0,01$). Em todos os casos, o ajuste com o modelo de Clutter apresentou um melhor desempenho. Nas estruturas I e II, observou-se que o modelo de Buckman modificado apresentou pior desempenho ($R^2 = 0,7664$ - RQEM% = 20,1905 - Bias% = -5,7514; $R^2 = 0,6308$ - RQEM% = 22,2517 - Bias% = -4,4525). Para esse modelo a estrutura III foi melhor ($R^2 = 0,9345$ - RQEM% = 15,8859 - Bias% = -3,5490), mas inferior ao desempenho do modelo de Clutter para esta mesma estrutura III ($R^2 = 0,9456$ - RQEM% = 13,7365 - Bias% = 0,1664).

Ao utilizar as estruturas I e II foram observadas maiores superestimação e subestimação, com o modelo de Buckman. Isto ocorreu em toda amplitude de idades sendo evidenciado a necessidade da mudança na estrutura de dados usual. As estimativas obtidas com o modelo de Buckman ajustado com a estrutura de dados III foi mais consistente, com erros relativos percentuais e resíduos melhor distribuídos.

A exatidão foi maior ao empregar o modelo de Clutter, em relação ao modelo de Buckman, com as três estruturas de dados, sempre com maior amplitude de erros nas idades iniciais (Tabela 2 e Figuras 4 a 7). Os resultados deste estudo mostram que a exatidão e consistência das projeções dependem da forma de organização do banco de dados, independentemente do modelo de densidade variável utilizado (Clutter ou Buckman).

Tabela 2 - Estimativas dos parâmetros dos modelos de Clutter e de Buckman ajustados com os dados organizados conforme Tabela 1 e respectivas estatísticas de precisão e de exatidão.

Organização de dados		I		II		III	
Modelo		Clutter	Buckman	Clutter	Buckman	Clutter	Buckman
Parâmetros estimados	α_0	4,0722	5,8046	3,8250	5,9999	4,8095	-2,7324
	α_1	-0,0138	-0,0297	-0,0059	-0,0121	-0,0398	-0,0147
	α_2	-	0,0387	-	0,0594	-	-0,0061
	α_3	-	-77,6705	-	-86,3499	-	38,0746
	α_4	-	-0,8806	-	-1,4494	-	0,5344
	β_0	1,5925	1,5265	1,7197	1,6039	2,6132	2,1149
	β_1	-	0,0232	-	0,0186	-	0,0069
	β_2	18,2778	-17,5504	17,7620	-16,4358	32,0684	-20,1348
	β_3	0,0238	-17,5504	0,0194	-16,4358	0,0318	-20,1348
	β_3	1,1391	1,1622	1,1380	1,1767	0,8072	1,1340
Critérios de desempenho	R^2	0,9247	0,7664	0,8669	0,6308	0,9456	0,9345
	MDA (m^3ha^{-1})	16,9298	30,7131	22,0843	34,7635	14,243	17,218
	RQEM (%)	11,0035	20,1905	13,1523	22,2517	13,7365	15,8859
	Bias (m^3ha^{-1})	-0,7999	-11,7048	-2,6863	-10,1982	0,2535	-5,4071
	Bias (%)	-0,3930	-5,7514	-1,1728	-4,4525	0,1664	-3,5490

Figura 4 - Correlação para as três estruturas de dados com modelos Clutter e Buckman modificado em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

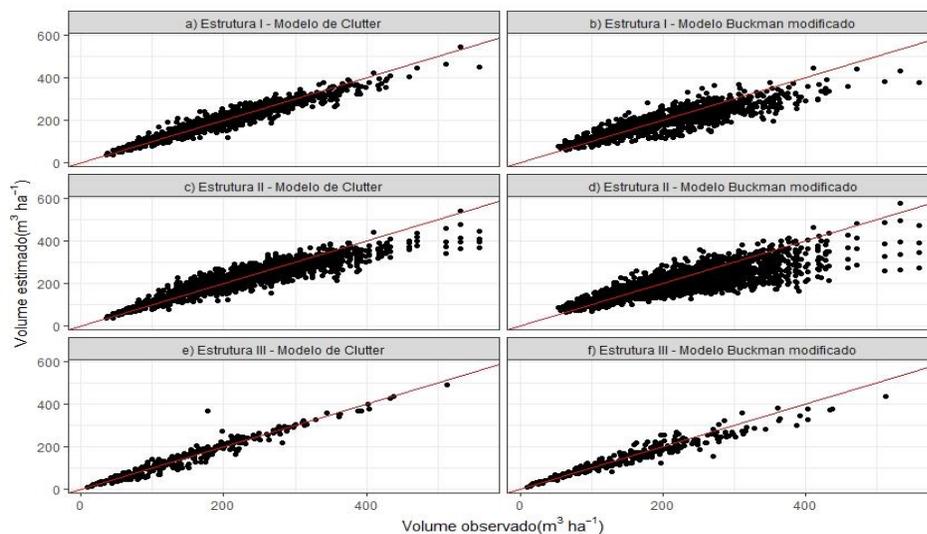


Figura 5 - Dispersão dos resíduos em relação ao volume observado as três estruturas de dados com modelos Clutter e Buckman modificado em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

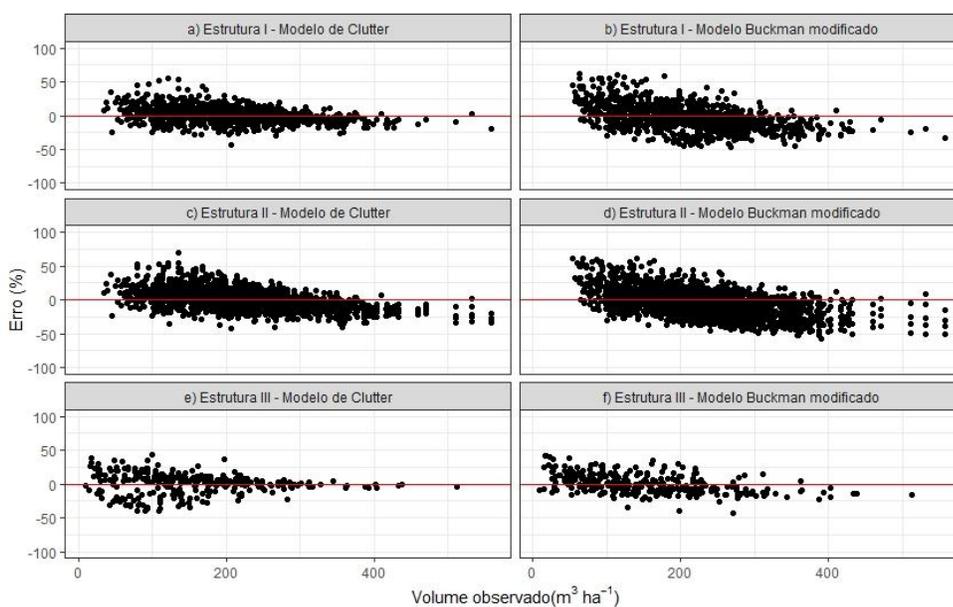


Figura 6 - Dispersão dos resíduos em relação à idade para as três estruturas de dados com modelos Clutter e Buckman modificado em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

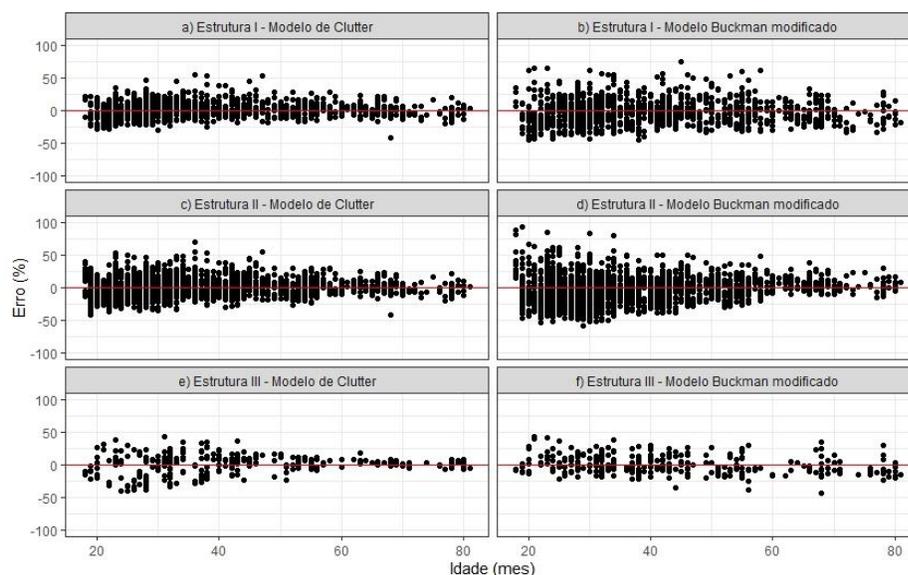
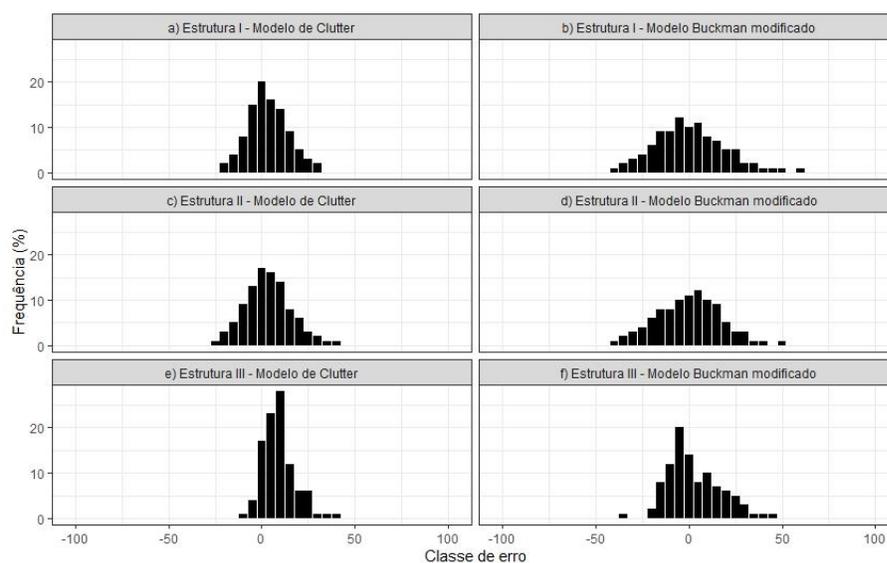


Figura 7 - Histograma da frequência em relação a classe de erro para as três estruturas de dados com modelos Clutter e Buckman modificado em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG



Na Tabela 3 são apresentadas as idades técnicas de corte (rotação) e as correspondentes produções, por classe de sítio, obtidas ao aplicar os modelos de Clutter, ajustado com os dados organizados pelas estruturas I, II e II e ao aplicar o modelo de Buckman com os dados organizados pela estrutura III. As estruturas I e II com modelo de Buckman foram desconsideradas nestas determinações devido às tendenciosidades e baixa exatidão observadas. Foram utilizadas áreas basais iniciais de $9,36 \text{ m}^2\text{ha}^{-1}$ (Classe I; $S = 32$), $6,94 \text{ m}^2\text{ha}^{-1}$ (Classe II; $S = 28$) e $3,6 \text{ m}^2\text{ha}^{-1}$ (Classe III; $S = 24$), com idade inicial de 19 meses. Estimativas de incremento e produção para diferentes idades são apresentadas nas Tabelas 4, 5 e 6 e Figura 8.

Tabela 3 - Idade técnica de colheita, produtividades médias (IMA) correspondentes e produtividade média aos 7 anos de idade (IMA_7), por classe de sítio, estrutura de organização dos dados e modelo (Clutter e Buckman), para povoamentos de clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* localizado nas regiões de Cocais, Sabinópolis e Virginópolis, MG

Organização de dados	Modelo	Índice de sítio (m)	Idade técnica de corte (mês)	IMA na idade técnica de corte ($\text{m}^3\text{ha}^{-1}\text{ano}^{-1}$)	IMA_7 ($\text{m}^3\text{ha}^{-1}\text{ano}^{-1}$)
I	Clutter	S = 32	48	60,22	52,92
		S = 28	56	50,10	46,65
		S = 24	71	38,06	37,60
II	Clutter	S = 32	48	60,20	52,71
		S = 28	55	49,72	45,98
		S = 24	70	37,28	36,68
III	Clutter	S = 32	52	55,85	50,54
		S = 28	59	49,10	46,45
		S = 24	72	40,56	40,07
	Buckman	S = 32	55	62,64	58,54
		S = 28	60	52,03	49,64
		S = 24	68	36,34	35,66

Tabela 4 - Tabela de produção para três índice sitio com a estrutura de dados I utilizando o modelo de Clutter em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virgíópolis, MG

CLASSE I; S = 32 - Área basal inicial de 9,46 m ² ha ⁻¹ *						CLASSE II; S = 28 - Área basal inicial de 6,94 m ² ha ⁻¹ *					CLASSE III; S = 24 - Área basal inicial de 3,6 m ² ha ⁻¹ *				
Modelo de Clutter															
Idade (mês)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)
20	0,68	10,14	58,99	-	2,95	0,63	7,57	38,49	-	1,92	0,47	4,07	17,26	-	0,86
21	0,66	10,79	66,17	7,18	3,15	0,62	8,20	43,99	5,50	2,10	0,48	4,55	20,46	3,20	0,97
..
47	0,26	21,58	235,75	5,20	5,02	0,30	19,67	192,96	4,95	4,11	0,33	15,59	134,64	4,38	2,87
48	0,25	21,83	240,84	5,09	5,02	0,29	19,96	197,82	4,86	4,12	0,33	15,92	138,98	4,34	2,90
49	0,25	22,07	245,83	4,99	5,02	0,28	20,25	202,60	4,78	4,14	0,32	16,24	143,27	4,29	2,92
..
55	0,21	23,40	273,70	4,41	4,98	0,24	21,80	229,57	4,30	4,17	0,28	18,02	168,01	4,00	3,06
56	0,20	23,61	278,02	4,32	4,97	0,24	22,04	233,79	4,22	4,18	0,28	18,29	171,95	3,95	3,07
57	0,20	23,80	282,25	4,23	4,95	0,23	22,27	237,94	4,15	4,17	0,27	18,56	175,85	3,90	3,09
..
70	0,14	25,93	330,28	3,28	4,72	0,17	24,82	285,68	3,30	4,08	0,21	21,62	222,00	3,26	3,17
71	0,14	26,07	333,50	3,22	4,70	0,17	24,98	288,93	3,25	4,07	0,20	21,82	225,22	3,22	3,17
72	0,13	26,20	336,66	3,16	4,68	0,16	25,15	292,12	3,19	4,06	0,20	22,02	228,39	3,17	3,17
..
83	0,11	27,50	367,91	2,60	4,43	0,13	26,73	323,91	2,66	3,90	0,16	24,00	260,53	2,72	3,14
84	0,10	27,60	370,47	2,55	4,41	0,13	26,86	326,53	2,62	3,89	0,16	24,16	263,22	2,69	3,13

*área basal inicial na idade de 19 meses

Tabela 5 - Tabela de produção para três índice sitio com a estrutura de dados II utilizando o modelo de Clutter em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

CLASSE I; S = 32 - Área basal inicial de 9,46 m ² ha ⁻¹ *						CLASSE II; S = 28 - Área basal inicial de 6,94 m ² ha ⁻¹ *					CLASSE III; S = 24 - Área basal inicial de 3,6 m ² ha ⁻¹ *				
Modelo de Clutter															
Idade (mês)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)
20	0,68	10,14	59,74	-	2,99	0,62	7,56	39,60	-	1,98	0,46	4,06	18,04	-	0,90
21	0,66	10,80	66,94	7,20	3,19	0,61	8,18	45,14	5,54	2,15	0,47	4,53	21,30	3,26	1,01
..
47	0,26	21,64	235,77	5,16	5,02	0,29	19,36	192,21	4,83	4,09	0,32	15,06	133,61	4,24	2,84
48	0,26	21,90	240,82	5,05	5,02	0,28	19,65	196,96	4,74	4,10	0,31	15,37	137,80	4,19	2,87
49	0,25	22,14	245,76	4,94	5,02	0,28	19,92	201,62	4,66	4,12	0,30	15,67	141,94	4,15	2,90
..
54	0,21	23,27	268,99	4,46	4,98	0,24	21,19	223,70	4,26	4,14	0,27	17,08	161,93	3,90	3,00
55	0,21	23,48	273,36	4,36	4,97	0,24	21,43	227,88	4,18	4,14	0,27	17,35	165,77	3,85	3,01
56	0,20	23,68	277,63	4,28	4,96	0,23	21,66	231,99	4,11	4,14	0,26	17,61	169,57	3,80	3,03
..
69	0,15	25,88	326,07	3,31	4,73	0,17	24,18	279,15	3,25	4,05	0,20	20,53	214,39	3,16	3,11
70	0,14	26,02	329,31	3,24	4,70	0,16	24,34	282,35	3,20	4,03	0,20	20,72	217,50	3,12	3,11
71	0,14	26,16	332,49	3,18	4,68	0,16	24,50	285,49	3,14	4,02	0,19	20,91	220,57	3,07	3,11
..
83	0,11	27,61	366,48	2,57	4,42	0,13	26,19	319,32	2,57	3,85	0,15	22,95	254,21	2,59	3,06
84	0,11	27,71	369,00	2,52	4,39	0,12	26,31	321,85	2,53	3,83	0,15	23,10	256,77	2,55	3,06

*área basal inicial na idade de 19 meses

Tabela 6 - Tabela de produção para três índice local com a estrutura de dados III utilizando o modelo de Clutter e Buckman em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

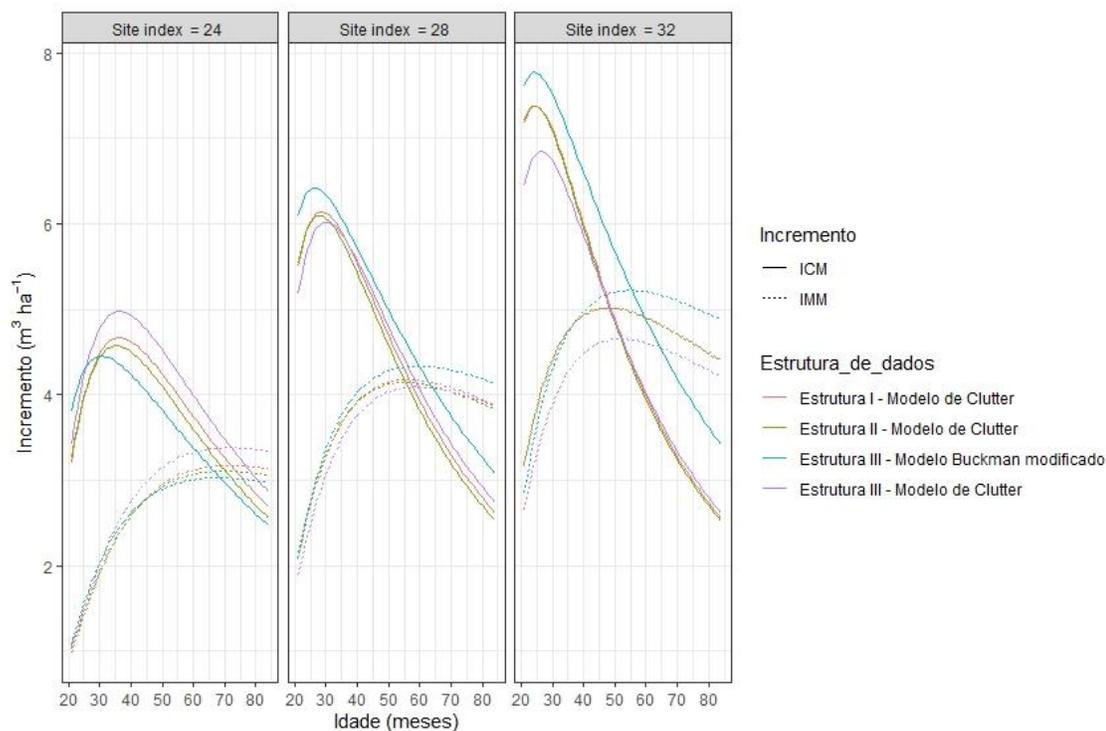
CLASSE I; S = 32 - Área basal inicial de 9,46 m ² ha ⁻¹ *						CLASSE II; S = 28 - Área basal inicial de 6,94 m ² ha ⁻¹ *					CLASSE III; S = 24 - Área basal inicial de 3,6 m ² ha ⁻¹ *				
Modelo de Clutter															
Idade (mês)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)	dB (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	ICM (m ³ ha ⁻¹)	IMM (m ³ ha ⁻¹)
20	0,63	10,09	49,11	-	2,45	0,63	7,57	34,31	-	1,71	0,49	4,09	18,38	-	0,91
21	0,60	10,69	55,56	6,45	2,64	0,62	8,20	39,49	5,17	1,88	0,50	4,60	21,79	3,41	1,03
..
51	0,20	21,23	237,32	4,77	4,65	0,27	20,90	206,35	4,72	4,04	0,34	18,09	161,66	4,47	3,17
52	0,19	21,43	242,01	4,68	4,65	0,26	21,17	211,00	4,64	4,05	0,33	18,42	166,08	4,42	3,19
53	0,19	21,62	246,60	4,59	4,65	0,25	21,43	215,56	4,56	4,06	0,33	18,75	170,45	4,36	3,21
..
58	0,16	22,50	268,29	4,17	4,62	0,22	22,62	237,29	4,19	4,09	0,29	20,31	191,48	4,09	3,30
59	0,16	22,66	272,39	4,09	4,61	0,22	22,84	241,41	4,12	4,09	0,29	20,60	195,53	4,04	3,31
60	0,15	22,82	276,40	4,01	4,60	0,21	23,06	245,47	4,05	4,09	0,28	20,89	199,52	3,99	3,32
..
71	0,11	24,31	315,99	3,27	4,45	0,16	25,14	285,89	3,37	4,02	0,23	23,70	240,00	3,43	3,38
72	0,11	24,42	319,21	3,22	4,43	0,16	25,30	289,21	3,32	4,01	0,22	23,93	243,38	3,38	3,38
73	0,11	24,54	322,37	3,16	4,41	0,16	25,46	292,47	3,26	4,00	0,22	24,15	246,71	3,33	3,38
..
83	0,09	25,55	351,16	2,66	4,23	0,13	26,91	322,40	2,78	3,88	0,18	26,18	277,63	2,90	3,34
84	0,09	25,64	353,78	2,62	4,21	0,12	27,04	325,14	2,74	3,87	0,18	26,36	280,49	2,86	3,33

*área basal inicial na idade de 19 meses

CLASSE I; S = 32 - Área basal inicial de 9,46 m ² ha ⁻¹ *						CLASSE II; S = 28 - Área basal inicial de 6,94 m ² ha ⁻¹ *					CLASSE III; S = 24 - Área basal inicial de 3,6 m ² ha ⁻¹ *				
Modelo de Bukman															
Idade (mês)	<i>dB</i> (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	<i>ICM</i> (m ³ ha ⁻¹)	<i>IMM</i> (m ³ ha ⁻¹)	<i>dB</i> (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	<i>ICM</i> (m ³ ha ⁻¹)	<i>IMM</i> (m ³ ha ⁻¹)	<i>dB</i> (m ² ha ⁻¹)	Área basal (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	<i>ICM</i> (m ³ ha ⁻¹)	<i>IMM</i> (m ³ ha ⁻¹)
20	0,88	10,34	52,05	-	2,60	0,77	7,71	37,29	-	1,86	0,55	4,15	18,50	-	0,92
21	0,84	11,18	59,66	7,60	2,84	0,73	8,44	43,37	6,08	2,06	0,54	4,69	22,30	3,80	1,06
..
54	0,27	26,24	281,84	5,34	5,21	0,26	22,20	233,20	4,71	4,31	0,22	15,88	159,46	3,63	2,95
55	0,27	26,51	287,10	5,26	5,22	0,25	22,46	237,86	4,65	4,32	0,21	16,10	163,06	3,59	2,96
56	0,26	26,78	292,28	5,18	5,21	0,25	22,71	242,44	4,58	4,32	0,21	16,31	166,61	3,55	2,97
..
59	0,25	27,55	307,36	4,95	5,21	0,23	23,43	255,81	4,39	4,33	0,20	16,93	177,00	3,42	3,00
60	0,24	27,79	312,24	4,87	5,20	0,23	23,66	260,14	4,33	4,33	0,19	17,13	180,38	3,37	3,00
61	0,24	28,03	317,04	4,80	5,19	0,22	23,88	264,41	4,26	4,33	0,19	17,32	183,71	3,33	3,01
..
67	0,21	29,38	344,38	4,38	5,14	0,20	25,15	288,78	3,91	4,31	0,17	18,42	202,85	3,08	3,02
68	0,20	29,59	348,70	4,32	5,12	0,19	25,35	292,64	3,86	4,30	0,17	18,59	205,90	3,04	3,02
69	0,20	29,79	352,96	4,25	5,11	0,19	25,54	296,45	3,80	4,29	0,16	18,76	208,91	3,00	3,02
70	0,20	29,99	357,15	4,19	5,10	0,19	25,73	300,20	3,75	4,28	0,16	18,92	211,88	2,96	3,02
..
83	0,15	32,30	406,39	3,46	4,89	0,15	27,92	344,43	3,12	4,15	0,13	20,83	247,12	2,50	2,97

*área basal inicial na idade de 19 meses

Figura 8 - Curvas de incremento para três índice sitio com as estruturas de dados I – II – III com o modelo de Clutter e estrutura de dados III com o modelo de Buckman modificado em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG



5. DISCUSSÃO

Neste estudo se analisou o efeito da forma de organização dos dados no resultado estatístico do ajuste dos modelos de Clutter e de Buckman. Esses modelos usualmente são ajustados com uma estrutura de banco de dados com intervalos de idade fixos entre duas medições consecutivas, que pode gerar perda de eficiência, conforme Cruz et al. (2008) e Salles et al. (2012).

A ideia de modificar a estrutura do banco de dados com os intervalos de crescimento para ajustar modelos de crescimento e produção surgiu do princípio da implementação de métodos não paramétricos (VIEIRA et al., 2018; MONGUS et al., 2018). Além disso, pelo fato da possibilidade das suposições estatísticas serem violadas devido às características

longitudinais dos dados. Entretanto, deve-se considerar que os intervalos de crescimento podem ser utilizados para obter os parâmetros (DORADO, 2004) e também podemos organizar os dados para as avaliações das projeções volumétricas.

A aplicação de um modelo de crescimento e produção indica um apoio à decisão no manejo florestal. No caso do modelo de Clutter, podem surgir dificuldades estatísticas devido às características dos dados disponíveis (CAMPOS e LEITE, 2017). Em nosso estudo, essa dificuldade foi apresentada no ajuste da área basal, obtendo na variável de índice local o parâmetro (αI) negativo (Tabela 2); no entanto, esse parâmetro tende a ser zero, o que indica não ser prejudicial no desempenho do ajuste do modelo, sendo todas as estimativas significativas em 1% ($p < 0,001$) e de bom desempenho estatístico para cada estrutura de dados estabelecida neste estudo.

O modelo de Clutter é amplamente utilizado no campo florestal, sendo grande o número de estudos realizados e preferido pelos pesquisadores, tornando-se referência fundamental no trabalho de medição florestal (SOARES et al., 1998; SOARES et al., 2004; DIAS et al., 2005; SANTANA et al., 2005; SOARES et al., 2005; GORGENS et al., 2007; DEMOLINARI et al., 2007; CRUZ et al., 2008; SALLES et al., 2012; CASTRO et al., 2013; MENDONÇA et al., 2014; MIRANDA et al., 2014; NASCIMENTO et al., 2015; CASTRO et al., 2015; ROVEDA et al., 2016; MANHÃES, 2019).

Nas estruturas I e II, o modelo de Buckman resultou em estimativas tendenciosas, havendo uma melhoria e aumento de exatidão ao empregar a estrutura de dados III. No caso do modelo de Clutter não houve tendenciosidade, mas a exatidão foi um pouco diferenciada em função da forma de organizar os dados.

Alguns autores destacaram o uso do modelo Buckman modificado para estimativas volumétricas (SILVA et al., 2006; GUERA et al., 2019), mas o modelo não é usado no Brasil. No entanto pode-se observar que o modelo de Buckman modificado apresentou um bom resultado com a estrutura de dados III, em comparação ao emprego do modelo de Clutter. Talvez o modelo não tenha sido utilizado no Brasil devido a inconsistências nas projeções, como consequência da forma de organizar os dados para os ajustamentos.

Para a obtenção da área basal se utilizou o método de repetidas soluções, que consiste em soluções repetidas com base na área basal, o que implica a solução da equação de crescimento para um determinado índice de sitio, idade e densidade do povoamento. O crescimento é então adicionado à densidade inicial da plantação, um mês é adicionado à idade e a equação é resolvida novamente (SILVA et al., 2006). É possível que para próximos estudos, essa estrutura de dados se baseia numa equação para obter dados mensalmente, o que poderia intervir em violações das suposições estatísticas.

A idade técnica de corte determinada com o modelo de Clutter ajustado com dados organizados pelas estruturas I e II foram muito próximas. Maiores diferenças foram encontradas ao utilizar a estrutura III. Nesse último caso (estrutura III), há maior proximidade nas estimativas de idades técnicas de corte estimadas com os modelos de Clutter e de Buckman. Isto evidencia a necessidade de se considerar a forma de organização dos dados para ajuste de modelos de projeção explícita, como de Clutter e de Buckman. Este último modelo por exemplo deve ser ajustado com os dados organizados pela estrutura III.

Os dados para ajuste do modelo de Buckman em princípio podem ser organizados em intervalos constantes de um ano para todas as parcelas, ou de um mês. Porém, a organização em intervalos de um ano limita a aplicação do modelo em planejamento estratégico. Isto porque nesse tipo de planejamento muitas vezes é necessário fazer projeções da produção de cada unidade de manejo para um único mês de referência, por exemplo, dezembro de cada ano). Nesse caso depreende-se a necessidade de poder projetar mensalmente e não anualmente, uma vez que as datas de medição inicial das parcelas não são as mesmas em uma grande base de dados.

Os resultados obtidos neste estudo, com o modelo de Clutter, são importantes em estudos comparativos entre esse modelo e metaheurísticas como redes neurais artificiais (RNA). Com base nos resultados do estudo, estimativas obtidas com RNA, onde os dados são organizados conforme estrutura II, podem ser comparadas com as correspondentes estimativas obtidas pelo modelo de Clutter independentemente da forma de estruturação dos dados utilizada para ajuste do mesmo.

6. CONCLUSÕES

A exatidão das estimativas de produção obtidas ao empregar o modelo de Clutter é influenciada pelas alternativas de organização dos dados.

O intervalo de idade para pareamento dos dados visando o ajuste do modelo de Clutter podem ser desiguais, porém, a organização com interpolação mensal não é indicada para esse modelo.

O ajuste do modelo de Buckman requer um mesmo intervalo entre as medições, para todas as parcelas e, para povoamentos de eucalipto, esse intervalo deve ser mensal, de modo a permitir o uso efetivo do modelo em aplicações no planejamento estratégico.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ALCÂNTRA, A. E. M.; DE ALBUQUERQUE SANTOS, A. C.; DA SILVA, M. L. M.; BINOTI, D. H. B.; SOARES, C. P. B.; GLERIANI, J. M.; LEITE, H. G. Use of artificial neural networks to assess yield projection and average production of eucalyptus stands. **African Journal of Agricultural Research**, v. 13, n. 42, p. 2285-2297, 2018.

ANDRADE, V. C. L.; SCHMITT, T. S.; BINOTI, D. H. B.; DE PÁDUA CHAVES, S.; LEITE, H. G. Modelos de altura-diâmetro para um plantio clonal de eucalipto no Sul do Tocantins. **Advances in Forestry Science**, v. 5, n. 4, p. 467-471, 2018.

ARANDA, U. D. **Modelo dinámico de crecimiento para masas de " Pinus sylvestris" L. procedentes de repoblación en Galicia**. Tesis Doctoral, Universidade de Santiago de Compostela, 2004.

ARAÚJO JÚNIOR, C. A.; SOUZA, P. D. D.; ASSIS, A. L. D.; CABACINHA, C. D.; LEITE, H. G.; SOARES, C. P. B.; SILVA, A, A, L; CASTRO, R. V. O. Artificial neural networks, quantile regression, and linear regression for site index prediction in the presence of outliers. **Pesquisa Agropecuária Brasileira**, v. 54, 2019.

AZEVEDO, G. B.; OLIVEIRA, E. K. B.; AZEVEDO, G. T. O. S.; BUCHMANN, H. M.; MIGUEL, E. P.; REZENDE, A. V. Modelagem da produção em nível de povoamento e por distribuição diamétrica em plantios de eucalipto. **Sci. For**, v. 44, n. 110, p. 383-392, 2016.

BAILEY, R. L.; CIESZEWSKI, C. J. Development of a well-behaved site-index equation: jack pine in north-central Ontario: comment. **Canadian Journal of Forest Research**, v. 30, n. 10, p. 1667, 2000.

BATISTA, J. L. F.; COUTO, H. T. Z.; MARQUESINI, M. Desempenho de modelos de relações hipsométricas: estudo em três tipos de floresta. **Scientia Forestalis**, v. 60, p. 149-163, 2001.

BECK, D. E.; DELLA-BIANCA, L. **Growth and yield of thinned yellow-poplar**. US Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, 1972.

BINOTI, D. H. B.; DA SILVA BINOTI, M. L. M.; LEITE, H. G. Configuração de redes neurais artificiais para estimação do volume de árvores. **Revista Ciência da Madeira (Brazilian Journal of Wood Science)**, v. 5, n. 1, p. 10-12953/2177-6830. v05n01a06, 2014.

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.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, n. 1, p. 97-105, 2015.

BLANCO, J. A.; AMEZTEGUI, A.; RODRÍGUEZ, F. Modelling Forest Ecosystems: A crossroad between scales, techniques, and applications. **Ecol. Model**, v. 425, p. 109030, 2020.

BUCKMAN, R. E. **Growth and yield of red pine in Minnesota**. US Dept. of Agriculture, 1962.

BURKHART, H. E. Growth and yield of southern pines—State of the art. In: **Proc. of the Southern forest economists workshop**. 1979.

BURKHART, H. E.; SPRINZ, P. T. Compatible cubic volume and basal area projection equations for thinned old-field loblolly pine plantations. **Forest Science**, v. 30, n. 1, p. 86-93, 1984.

BURKHART, H. E.; TOMÉ, M. **Modeling forest trees and stands**. Springer Science & Business Media, 2012.

CAMPOS, J. C. C. Análise do crescimento e produção em plantações desbastadas. **Revista Árvore, Viçosa**, v. 4, n. 2, p. 157-169, 1980.

CAMPOS, J.C.C.; LEITE, H.G. **Mensuração florestal: perguntas e respostas**. 5.ed. atual. e ampl. Viçosa: Ed. da UFV, 2017. 636p.

CAO, Q. V. Estimating coefficients of base-age-invariant site index equations. **Canadian Journal of Forest Research**, v. 23, n. 11, p. 2343-2347, 1993.

CASTRO, R. V. O.; CUNHA, A. B.; SILVA, L. V.; LEITE, H. G.; SILVA, A. A. L. Modelagem do crescimento e produção para um povoamento de Eucalyptus utilizando dois métodos para quantificação do índice de local. **Scientia Forestalis**, v. 43, n. 105, p. 83-90, 2015.

CASTRO, R. V. O.; SOARES C. P. B.; MARTINS, F. B.; LEITE, H. G.. Crescimento e produção de plantios comerciais de eucalipto estimados por duas categorias de modelos. **Pesquisa Agropecuária Brasileira**, v. 48, n. 3, p. 287-295, 2013.

CIESZEWSKI, C. J. Comparing fixed-and variable-base-age site equations having single versus multiple asymptotes. **Forest Science**, v. 48, n. 1, p. 7-23, 2002.

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

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. 333p.

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**, v. 32, n. 5, p. 821-828, 2008.

DÁVILA VEGA, A. E. **Modelagem de povoamentos de eucalipto no nordeste da Bahia utilizando regressão e redes neurais artificiais**. 2018. 64 f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2018.

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

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

DEMOLINARI, R. D. A.; SOARES, C. P. B.; LEITE, H. G.; SOUZA, A. L. Crescimento de plantios clonais de eucalipto não desbastados na região de Monte Dourado (PA). **Revista Árvore**, v. 31, n. 3, p. 503-512, 2007.

DIAS, A. N.; LEITE, H. G.; CAMPOS, J. C. C.; COUTO, L.; CARVALHO, A. F. Emprego de um modelo de crescimento e produção em povoamentos desbastados de eucalipto. **Revista Árvore**, v. 29, n. 5, p. 731-739, 2005.

DORADO, F. C. **Modelo dinámico de crecimiento para las masas de " Pinus radiata D. Don" en Galicia: simulación de alternativas selvícolas con inclusión del riesgo de incendio**. Tesis Doctoral. Universidade de Santiago de Compostela, 2004.

DRESCHER, R. **Crescimento e produção de *Tectona grandis* Linn F., em povoamentos jovens de duas regiões do estado de Mato Grosso-Brasil**. 2004. Tese de Doutorado. Universidade Federal de Santa Maria, 2004.

FINGER, C. A. G. **Fundamentos de biometria florestal**. Santa Maria: UFSM/CEPEF/FATEC, 1992. 269 p.

FREITAS, E. C. S.; DE PAIVA, H. N.; NEVES, J. C. L.; MARCATTI, G. E.; LEITE, H. G. Modeling of eucalyptus productivity with artificial neural networks. **Industrial Crops and Products**, v. 146, p. 112149, 2020.

GARAY, L. Tropical forest utilization system. VIII. A taper model for entire stem profile including buttressing. **Seatlhe. Coll. Forest. Resour., Inst. Forest Prod. Univ. Wash**, 1979.

GAVILÁN-ACUÑA, G.; OLMEDO, G. F.; MENA-QUIJADA, P.; GUEVARA, M.; BARRÍA-KNOPF, B.; WATT, M. S. Reducing the Uncertainty of Radiata Pine Site Index Maps Using an Spatial Ensemble of Machine Learning Models. **Forests**, v. 12, n. 1, p. 77, 2021.

GOELZ, J. C. G.; BURK, T. E. Development of a well-behaved site index equation: jack pine in north central Ontario. **Canadian Journal of Forest Research**, v. 22, n. 6, p. 776-784, 1992.

GOELZ, J. C. G.; BURK, T. E. Measurement error causes bias in site index equations. **Canadian Journal of Forest Research**, v. 26, n. 9, p. 1585-1593, 1996.

GORGENS, E. B.; LEITE, H. G.; NOGUEIRA, G. S.; DIAS, A. N. Tendência de crescimento de povoamento de eucalipto após aplicação de desbaste. **Revista Árvore**, Viçosa, v. 31, n. 5, p. 879-885, 2007.

GUERA, O. G. M.; SILVA, J. A. A.; FERREIRA, R. L. C.; LAZO, D. A. Á.; MEDEL, H. B. Alternatives to Growth and Yield Prognosis for *Pinus caribaea* var. *caribaea* Barrett & Golfari. **Floresta e Ambiente**, v. 26, n. 4, 2019.

HIRIGOYEN, A.; FRANCO, J.; DIÉGUEZ, U. Modelo dinámico de rodal para *Eucalyptus globulus* (L.) en Uruguay. **Agrociencia Uruguay**, v. 22, n. 1, p. 63-80, 2018.

HOSOKAWA, R. T.; MOURA, J. B.; CUNHA, U. S. **Introdução ao manejo e economia de florestas**. Curitiba: Editora UFPR, 1998, 162p.

HUANG, S. Development of compatible height and site index models for young and mature stands within an ecosystem-based management framework. **Empirical and process-based models for forest tree and stand growth simulation**, p. 61-98, 1999.

HUANG, S. Validating and localizing growth and yield models: procedures, problems and prospects. In: **Proceedings of IUFRO Workshop on Reality, models and parameter estimation: the forestry scenario**. Portugal: Sesimbra, 2002.

HUANG, S. **Development of a subregion-based compatible height-site index-age model for black spruce in Alberta**. Alberta Environmental Protection, Land and Forest Service, Forest Management Division, 1997.

HUSCH, B.; BEERS, T. W.; KERSHAW JR, J. A. **Forest mensuration**. John Wiley & Sons, 2002.

KIVISTE, A.; CONZÁLEZ, J. G. A. **Funciones de crecimiento de aplicación en el ámbito forestal**. INIA, 2002.

KOZAK, A.; MUNRO, D. D.; SMITH, J. H. G. Taper functions and their application in forest inventory. **The Forestry Chronicle**, v. 45, n. 4, p. 278-283, 1969.

LEITE, H.; CASTRO, R.; SILVA, A.; JÚNIOR, C.; BINOTI, D.; CASTRO, A. F.; BINOTI, M. Classificação da capacidade produtiva de povoamentos de eucalipto utilizando diâmetro dominante. **Silva Lusitana**, v. 19, n. 2, p. 181-195, 2011.

LEITE, H.G. **Ajuste de um modelo de estimação de frequência e produção por classe de diâmetro, para povoamentos de Eucalyptus saligna Smith**. 1990. 97f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 1990.

LEITE, R. V. et al. Estimating stem volume in eucalyptus plantations using airborne LiDAR: A comparison of area-and individual tree-based approaches. **Remote Sensing**, v. 12, n. 9, p. 1513, 2020.

MANHÃES, L. P.; SILVA, G. F.; BINOTI, D. H. B.; SILVA, W. B.; MENDONCA, A. R. **Uma abordagem bayesiana para monitoramento do crescimento e produção florestal de Eucalipto**. 2019. Dissertação (Mestrado em Ciência Florestal) Universidade Federal do Espírito Santo, 2019.

MEDEIROS, R. A. **Potencial produtivo, manejo e experimentação em povoamentos de Tectona grandis L.f. no Estado de Mato Grosso**. 2016. 182f. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa. 2016.

MEDEIROS, R. A.; DE PAIVA, H. N.; SOARES, Á. A. V.; MARCATTI, G. E.; TAKIZAWA, F. H.; DOMICIANO, C. A. R.; LEITE, H. G. Productive potential of *Tectona grandis* in Midwest Brazil. **Advances in Forestry Science**, v. 6, n. 4, p. 803-809, 2019.

MENDONÇA, A. R. D.; CALEGARIO, N.; SILVA, G. F. D.; SOUZA, A. L. D.; TRUGILHO, P. F.; CARVALHO, S. P. C.; POSSATO, E. L. Modelagem da produção de sortimentos em povoamentos de eucalipto. **Cerne**, v. 20, n. 4, p. 587-594, 2014.

MIGUEL, E. P.; DE OLIVEIRA, C. S.; MARTINS, T. O.; MATIAS, R. A. M.; REZENDE, A. V.; ANGELO, H.; MARTINS, I. S. Growth and yield models by total stand (MPT) in *Eucalyptus urophylla* (st Blake) plantations. **Aust J Basic Appl Sci**, v. 10, n. 13, p. 79-85, 2016.

MIRANDA, R. O. V.; DIAS, A. N.; FILHO, A. F.; SOARES, I. D.; CRUZ, J. P. Modelagem do crescimento e produção em classes de precipitação pluviométrica para *Eucalyptus* sp. **Floresta**, v. 45, n. 1, p. 117-128, 2014.

MIRANDA, R. O. V.; FIGUEIREDO, A. Ajuste do modelo de Clutter utilizando o programa Gretl. **Revista de Ciências Agrárias - Amazon Journal of Agricultural and Environmental Sciences**, v. 59, n. 4, p. 417–421, 2016.

MONGUS, D.; VILHAR U.; SKUDNIK, M.; ZALIK, B.; JESENKO, D. Predictive analytics of tree growth based on complex networks of tree competition. **Forest Ecology and Management**, v. 425, p. 164-176, 2018.

MORAES NETO, S. P.; PULROLNIK, K.; VILELA, L.; MUNHOZ, D. J. M.; GUMARÃES JÚNIOR, R.; MARCHÃO, R. L. **Modelos hipsométricos para *Eucalyptus cloeziana* e *Eucalyptus grandis* em Sistemas Agrossilvipastoril**. Planaltina, DF: Embrapa Cerrados, 2010.

NASCIMENTO, F. A. F. DO; DIAS, A. N.; FIGUEIREDO FILHO, A.; MIRANDA, G. D. M.; ARCE, J. E. Sistema de crescimento e produção para povoamentos de *Pinus taeda* na região norte de Santa Catarina. **Cerne**, v. 21, n. 2, p. 235–242, 2015.

NASCIMENTO, R. G. M. **Moelagem e simulação do crescimento e produção de floresta tropical manejada na Amazônia Oriental**. Thesis (Pós-Graduação em Engenharia Florestal). Curitiba: Universidade Federal do Paraná. 2016. p. 174.

PARRESOL, B. R. **White pine site index for the southern forest survey**. US Department of Agriculture, Forest Service, Southern Research Station, 1998.

PENIDO, T. M. A.; LAFETÁ, B. O.; NOGUEIRA, G. S.; ALVES, P. H.; GORGENS, E. B.; DE OLIVEIRA, M. L. R. Modelos de crescimento e produção para a estimativa volumétrica em povoamentos comerciais de eucalipto. **Scientia Forestalis**, v. 48, n. 128, 1 dez. 2020.

PIENAAR, L. V.; TURNBULL, K. J. The Chapman-Richards generalization of Von Bertalanffy's growth model for basal area growth and yield in even-aged stands. **Forest Science**, v. 19, n. 1, p. 2-22, 1973.

PIRES, L. M.; CALEGARIO, N. Ajuste de modelos estocásticos lineares e não-lineares para a descrição do perfil longitudinal de árvores. **Revista Árvore**, v. 31, n. 5, p. 845-852, 2007.

PORTÉ, A.; BARTELINK, H. H.. Modelling mixed forest growth: a review of models for forest management. **Ecological modelling**, v. 150, n. 1-2, p. 141-188, 2002.

QUANTITATIVE MICRO SOFTWARE. **EViews 6 Command Reference**. Irvine CA, USA, 2007.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria: R Foundation for Statistical Computing, 2020.

RESENDE, R.R. **Emprego de um modelo de crescimento e produção para determinação da rotação em povoamento de eucalipto**. 1991. 81 f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 1991.

ROVEDA, M.; DIAS, A. N.; FILHO, A. F.; SCAVINSK, V.; CASTRO R. R. Testes de identidade de modelos aplicados na construção de curvas de índices de local e na produção de povoamentos de *Eucalyptus grandis*. **Scientia Forestalis**, Piracicaba, v. 44, n. 109, p. 115-127, 2016.

SALEKIN, S.; MASON, E. G.; MORGENROTH, J.; MEASON, D. F. A preliminary growth and yield model for *Eucalyptus globoides* Blakely plantations in New Zealand. **New Zealand Journal of Forestry Science**, v. 50, 2020.

SALLES, T. T.; LEITE, H. G.; NETO, S. N. 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 Agropecuária Brasileira**, v. 47, n. 2, p. 253-260, 2012.

SANTANA, C.; MELLO, A. A.; EISFELD, R. L.; SANQUETTA, C. R. Sistema de equações para simulação do crescimento e da produção em povoamentos de *Eucalyptus grandis* Hill ex Maiden sem desbaste baseado no modelo de clutter. **Ambiência, Guarapuava**, v. 1, n. 2, p. 239-256, 2005.

SCOLFORO, J. R. S. **Biometria Florestal – Medição e Volumetria de árvores**. Lavras – MG: UFLA/FAEPE/DCF, 1998. 310p.

SILVA, A. L.; CAMPOS, J. C. C.; LEITE, H. G.; SOUZA, A. L.; LOPES, P. F. Growth and yield prediction using the modified Buckman model. **Revista Árvore**, vol. 30, n. 5, p. 787-793, 2006.

SOARES, C. P. B.; LEITE, H. G.; CAMPOS, J. C. C.; SEDIYAMA, G. C. A Inclusão da precipitação anual em um modelo de crescimento e produção. **Revista Árvore**, v. 22, n. 4, p. 475-482, 1998.

SOARES, C. P. B.; LEITE, H. G.; GORGENS, E. B. Equações para estimar o estoque de carbono no fuste de árvores individuais e em plantios comerciais de eucalipto. **Revista Árvore**, Viçosa, MG, v. 29, n. 5, 2005.

SOARES, C. P. B.; LEITE, H. G.; OLIVEIRA, M. L. R.; CARVALHO, A. Especificação de um modelo de crescimento e produção florestal. **Revista Árvore**, v. 28, n. 6, p. 831-837, 2004.

SOARES, C.P.B. **Um modelo para o gerenciamento da produção de madeira em plantios comerciais de eucalipto**. 1999. 71f. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, 1999.

SOUZA, G. S. A. et al. Multi-sensor prediction of *Eucalyptus* stand volume: A support vector approach. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 156, p. 135-146, 2019.

SOUZA, R. N. **Efeito de dois espaçamentos na produção em peso e volume de *Eucalyptus grandis* (W. Hill ex-Maiden)**. 1988. 74f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, 1988.

STRUB, M.; CIESZEWSKI, C. J. Base–age invariance properties of two techniques for estimating the parameters of site index models. **Forest Science**, vol. 52, n. 2, p. 182-186, 2006.

TREVIZOL JÚNIOR, T. L. **Análise de um modelo compatível de crescimento e produção em plantações de Eucalyptus grandis (W.Hill ex-Maiden)**. 1985. 74 f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, 1985.

VANCLAY, Jerome K. Sustainable timber harvesting: simulation studies in the tropical rainforests of north Queensland. **Forest ecology and management**, v. 69, n. 1-3, p. 299-320, 1994.

VIEIRA, G. C.; MENDONÇA, A. R.; SILVA, G. F.; ZANETTI, S. S.; SILVA, M. M.; Santos, A. R. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. **Science of the Total Environment**, v. 619, p. 1473-1481, 2018.

WEISKITTEL, A. R.; HANN, D. W.; KERSHAW JR, J. A.; VANCLAY, J. K. **Forest growth and yield modeling**. John Wiley & Sons, 2011.

WICKHAM, H. **ggplot2: Elegant Graphics for Data Analysis**. Springer-Verlag New York, 2016.

CAPITULO II

MODELAGEM DO CRESCIMENTO E DA PRODUÇÃO DE POVOAMENTOS DE EUCALIPTO UTILIZANDO REDES NEURAS ARTIFICIAIS CLÁSSICAS E REDES PROFUNDAS

1. INTRODUÇÃO

Um dos componentes ou elementos do manejo florestal é a prognose, feita usualmente com o emprego de modelos de crescimento e produção. Apesar do expressivo avanço nos estudos de modelagem do crescimento e da produção de povoamentos equiâneos, com emprego de regressão, ainda persistem alguns problemas. Um desses problemas é baixa exatidão das projeções feitas a partir de idades iniciais, por exemplo, a partir de dois a três anos de idade. Esta dificuldade está relacionada à natureza heterocedástica das relações entre altura dominante - Hd (e volume - V) com a idade (I), com menor variância em idades iniciais. De fato na maioria das vezes a dispersão de Hd e V em idade de 1,5 a 2,5 anos, em povoamentos de eucalipto, é relativamente muito pequena. Com o passar dos anos a variância de Hd e de V vai aumentando, caracterizando a heterocedasticidade que ocorre naturalmente.

A inclusão de mais variáveis nos modelos, como o índice de sítio e a área basal, pode resultar em ganho de exatidão nas projeções (CAMPOS e LEITE, 2017). Entretanto, se a amplitude de Hd aos 2 anos é pequena, então o índice de sítio (S) também apresentará pequeno desvio padrão nessa idade. Portanto, independentemente da projeção, de uma idade I_a para uma idade I_f , ser feita por $V_f = V_a f(I_f) / f(I_a)$ ou por $V_f = V_a f(I_f, S_a) / f(I_a, S_a)$, se I_a for uma idade muito inicial (1,5 a 3 anos), a inexatidão e insegurança na estimativa V_f é muito provável.

Metaheurísticas como redes neurais artificiais (RNA) tem sido cada vez mais utilizadas para diversas finalidades na ciência florestal (ARAÚJO JÚNIOR et al., 2019;

SOUZA et al., 2019; FREITAS et al., 2020; SILVA et al., 2020; LOPES et al., 2020; LEITE et al., 2020; GAVILÁN-ACUÑA et al., 2021). De modo geral a exatidão de estimativas de produção em nível de povoamento tem sido maior quando se emprega RNA, em comparação aos modelos de crescimento e produção (BINOTI et al., 2014a; CASAS et al., 2018; ALCÂNTRA et al., 2018; DÁVILA VEGA, 2018; TAVARES JÚNIOR et al., 2019; MEDEIROS et al., 2019).

Apesar do ganho de exatidão com o emprego de RNA, ainda persiste o problema mencionado anteriormente, da ineficiência da projeção a partir de cerca de 2 anos (no caso de povoamentos de eucalipto). Considerando a amplitude relativamente pequena das variáveis Hd , B e V nas idades iniciais (2 a 3 anos no caso de eucalipto), uma possibilidade seria usar estratificação máxima e ajuste de modelos mais simples (CAMPOS e LEITE, 2017). Tais modelos quando aplicados para cada estrato específico resultaria em diferenciação nas projeções. Entretanto dependendo do tamanho da área, a estratificação máxima mencionada pode ser inviável economicamente, devido ao grande esforço amostral. Além disso, na maioria das vezes o estrato irá contemplar mais de um talhão, e esses talhões agrupados terão pequena amplitude das variáveis Hd , B e V na idade inicial (cerca de dois anos), com possibilidade de haver diferenças significativas em V_f (por exemplo, ao final da rotação). Considerando que o emprego de RNA não requer grande frequência para cada combinação das variáveis de entrada, voltamos à possibilidade de uso da mesma, buscando identificar variáveis categóricas que possam contribuir para a referida exatidão ou tipos de RNA diferentes daquelas já usuais.

O aprendizado profundo (*deep learning*) faz parte de um campo mais amplo de aprendizado de máquina e inteligência artificial que usa redes neurais artificiais (BHARATENDRA, 2019). A diferença em relação a uma rede convencional é que as redes profundas permitem que modelos computacionais compostos por várias camadas de processamento aprendam representações de dados com vários níveis de abstração (GOODFELLOW et al., 2016; LECUN et al., 2015; NIELSEN, 2015; AGGARWAL, 2018).

É possível que o emprego de redes profundas possa contribuir para aumentar a exatidão das estimativas de produção em povoamentos de eucalipto, principalmente quando feitas a partir de idades precoces. O aprendizado profundo tem sido cada vez mais utilizado

em pesquisas envolvendo estruturas complexas, em diferentes áreas (DENG e YU, 2014; LECUN et al; 2015; SUN et al., 2017; HAMDI et al., 2019; Xi et al., 2020; KENTSCH et al., 2020; FERREIRA et al., 2020)

O objetivo do presente estudo foi de avaliar a eficiência de redes neurais artificiais profundas (*deep learning*) na projeção da produção de povoamentos de eucalipto, em relação ao emprego de redes neurais clássicas e avaliar a possibilidade de ganho de exatidão em projeções de produção de eucalipto feitas com aproximadamente de dois anos.

2. REVISÃO BIBLIOGRÁFICA

2.1. Conceitos básicos sobre redes neurais artificiais

O desenvolvimento científico relacionado às redes neurais artificiais (RNA) é motivado pela forma de como o cérebro humano processa informações, podendo organizar seus neurônios para determinado processamento, como reconhecimento de padrões, muito mais rápido e eficientemente do que um computador convencional (HAYKIN e ENGEL, 2001).

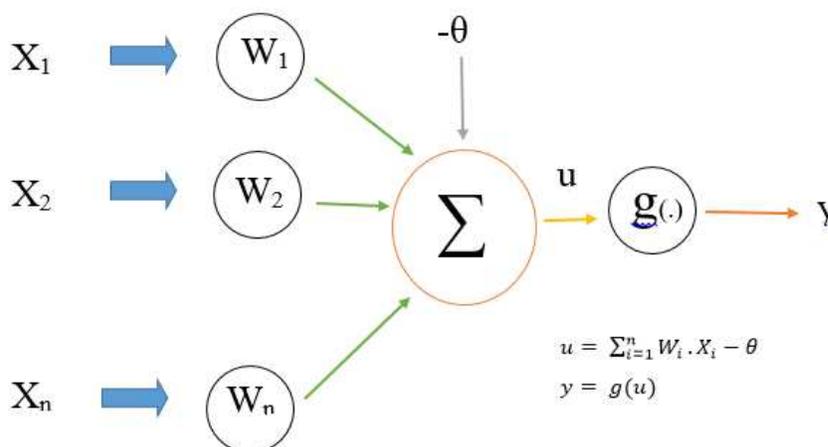
As RNA são redes de neurônios artificiais distribuídos em camadas e conectados de forma específica para executar uma determinada tarefa (BULLINARIA, 2014). Uma RNA consegue detalhar cada vez mais precisamente a produtividade média de grandes áreas de cultivo de eucalipto do que um modelo tradicional. Com elas é mais fácil incluir um grande número de variáveis, como características edáficas, climáticas e fisiográficas (ALCÂNTARA, 2018), histórico silvicultural dos talhões (FREITAS, 2020) e qualquer informação contida no cadastro florestal (ALCÂNTARA, 2018, FREITAS, 2020). Para estimação da produção corrente podem ser incluídas, por exemplo, métricas obtidas a partir de nuvens de pontos de LIDAR (LEITE et al., 2020).

Os elementos computacionais ou unidades processadoras, denominada unidades ou neurônios artificiais, são modelos simplificados dos neurônios biológicos que foram inspirados a partir de análises de geração e propagação de impulsos elétricos pela membrana

celular dos neurônios (HODGKIN e HUXLEY, 1952). Um modelo muito simplificado de um neurônio, proposto por McCulloch e Pitts (1943), incluía paralelismo e alta conectividade, com os 7 elementos básicos de uma rede neural artificial:

1. Sinais de entrada (X_1, X_2, \dots, X_n) = valores assumidos pelas variáveis de uma aplicação específica, geralmente são normalizados para incrementar a eficiência computacional dos algoritmos de aprendizagem;
2. Pesos sinápticos (W_1, W_2, \dots, W_n) = valores utilizados para ponderar cada uma das variáveis de entrada;
3. Combinador linear (Σ) cuja função é agregar todos os sinais de entrada a fim de produzir um valor de potencial de ativação;
4. Limiar de ativação (θ) = variável que especifica qual será o patamar apropriado para que o resultado produzido pelo combinador linear possa gerar um valor de disparo em direção à saída do neurônio;
5. Potencial de ativação (u) = resultado obtido pela diferença do valor produzido entre o combinador linear e o limiar de ativação;
6. Função de ativação (g) utilizada para limitar a saída do neurônio dentro de um intervalo de valores razoáveis; e
7. Sinal de saída (y) = valor final produzido pelo neurônio em relação a um determinado conjunto de sinais de entrada.

Figura 1 – Ilustração do neurônio artificial por proposto por McCulloch e Pitts (1943)



2.2. Principais funções de ativação

A escolha da função de ativação é uma parte crítica de uma rede neural, dependendo do que geralmente é previsto, da normalização dos dados e da variável de destino. Por exemplo, se a variável de destino prevista for real, a função de identidade deve ser usada. Se for desejável prever uma probabilidade de classe binária, faz sentido usar a função sigmoide e assim por diante (AGGARWAL, 2018)

Alguns exemplos de funções (AGGARWAL, 2018 e SOCHER e MUNDRA, 2016), ilustradas na Figura 2, são:

$$\Phi(v) = v \text{ (função identidade)}$$

$$\Phi(v) = \text{sign}(v) \text{ (função signo)}$$

$$\Phi(v) = \frac{1}{1+e^{-v}} \text{ (função logística)}$$

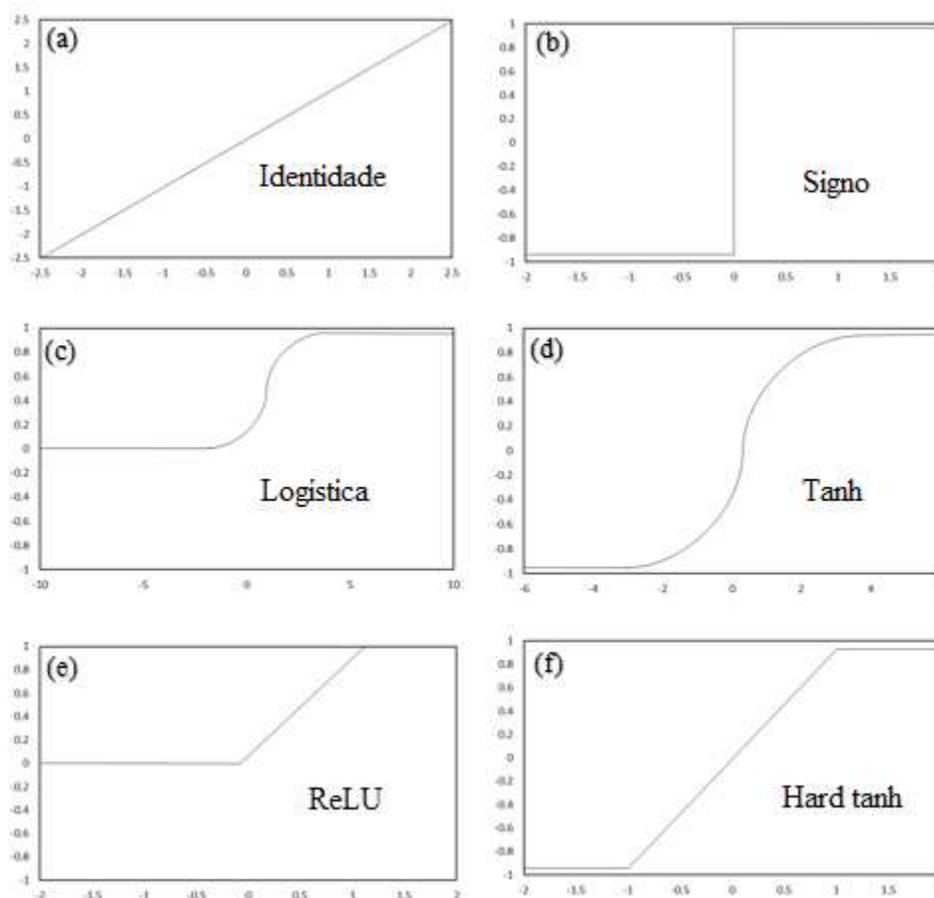
$$\Phi(v) = \frac{e^{2v}-1}{e^{2v}+1} \text{ (função tanh)}$$

$$\Phi(v) = \max\{v, 0\} \text{ (Rectified Linear Unit – ReLU)}$$

$$\Phi(v) = \max\{\min[v, 1], -1\} \text{ (hard tanh)}$$

A função sigmoide gera valores entre 0 e 1 enquanto a função tanh, que tem uma forma similar a função sigmoide, gera valores entre -1 e 1.

Figura 2 - Diferentes funções de ativação: (a) Identidade, (b) Signo, (c) Logística, (d) Tanh, (e) ReLU e (f) Hard tanh



2.3. Arquitetura de redes neurais artificiais

A RNA pode ser dividida por três partes ou camadas (SILVA et al., 2016), sendo: uma *camada de entrada*, que recebe as informações de dados, geralmente normalizados, uma *camada escondida* (intermediária, oculta ou invisível), cuja função é extrair as características associadas ao processo, onde é feito a maior parte do processamento interno da rede, e uma *camada de saída*, que contém o resultado final da rede, ou seja, a resposta da rede.

As principais arquiteturas das RNA aplicadas na Engenharia Florestal são *feedforward* de camadas múltiplas. Estas arquiteturas estão relacionadas a disposição de seus neurônios, formas de interligação entre eles e a constituição de suas camadas. Na maioria das

vezes são utilizadas redes neurais com uma arquitetura simples, com uma única camada oculta (FREITAS et al., 2020; LOPES et al., 2020).

Problemas mais complexos podem requerer redes com uma configuração de múltiplas camadas, transformando-se numa rede de aprendizagem profunda (*Deep Learning*) ou arquitetura *Deep Feed Forward*. Segundo Santana (2018), a diferença de uma rede convencional ou clássica, com o *Deep learning*, não é apenas em relação à quantidade de camada ocultas agregadas na rede (Figura 3), mas também à necessidade e impacto do volume dos dados, esforço computacional, quantidade de informação para a fase de treinamento da rede e aos algoritmos de otimização.

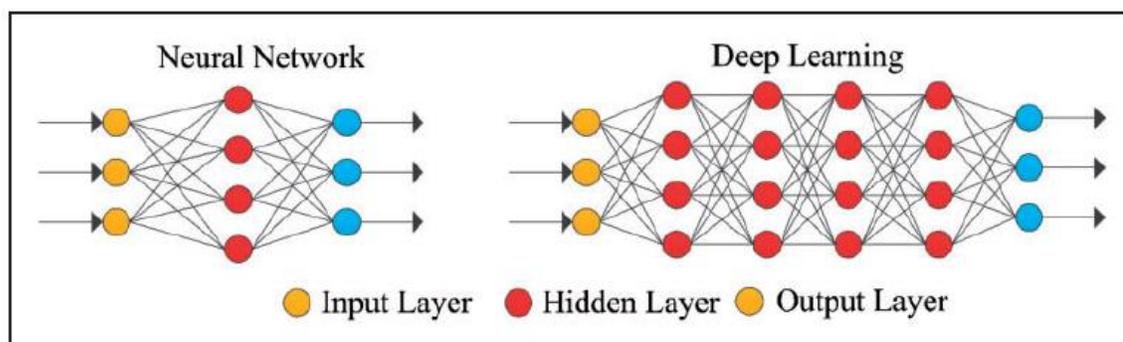


Figura 3 - Diferenças de rede neural convencional ou clássica e rede neural *Deep Learning* de acordo com as camadas ocultas. Fonte: XING e DU, 2019

Além destas duas, existem várias estruturas diferentes com uma infinidade de finalidades e seu funcionamento também depende da estrutura e todas são baseadas em redes neurais (CHAGAS, 2019). Dessas diferentes arquiteturas, Leijnen e Veen (2020), estabelecem uma visão geral das arquiteturas das redes neurais, algumas foram criados nos últimos anos e outra tem origem há muitas décadas (Figura 4):

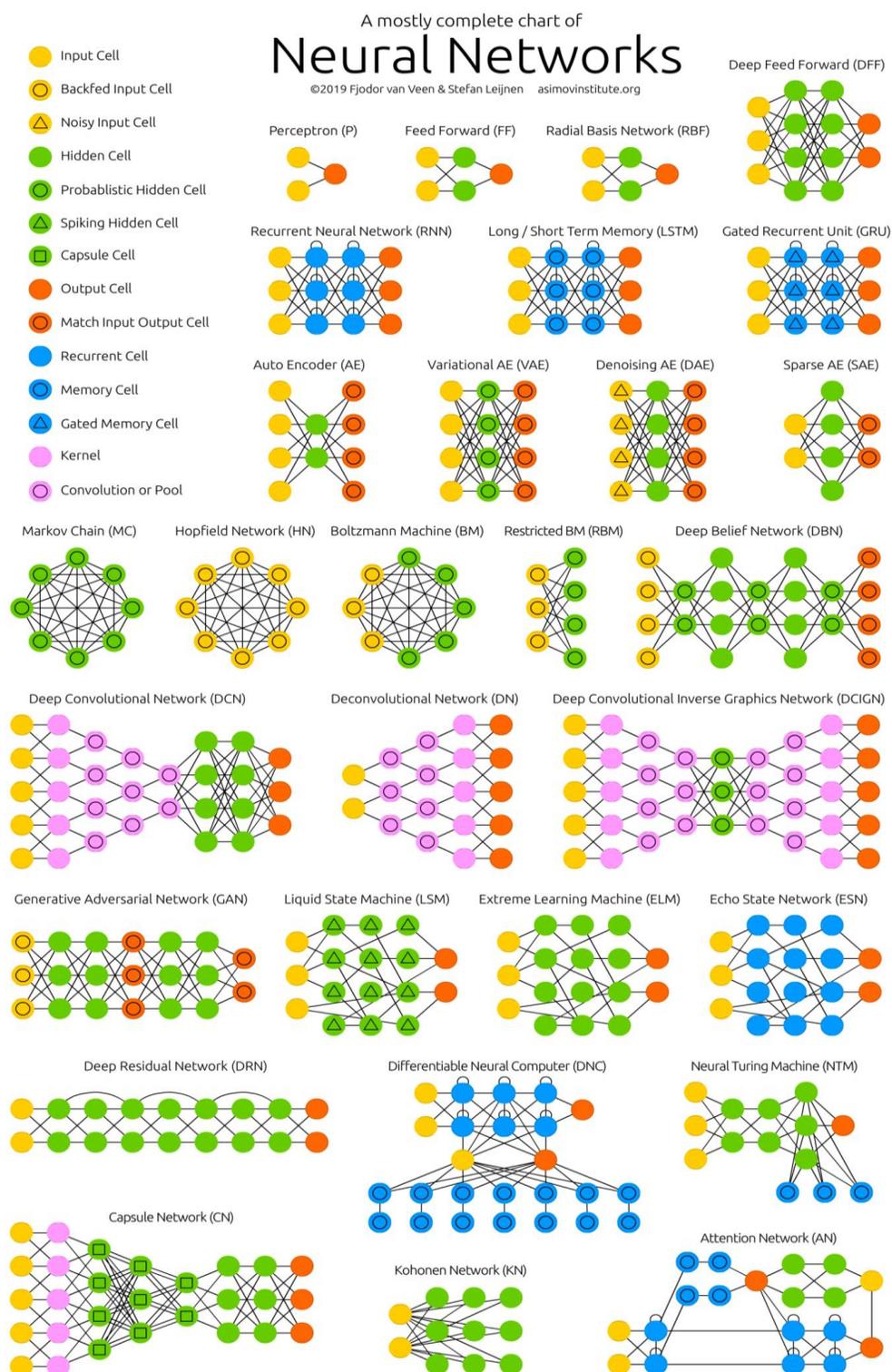
- Redes neurais *feed forward e perceptrons*
- *Recurrent Neural Networks*

- *Long Short-Term Memory*
- *Autoencoders*
- *Hopfield Networks and Boltzmann Machines*
- *Convolutional Networks*
- *Generative Adversarial Networks*
- *Liquid State Machines and Echo State Machines*
- *Deep Residual Networks*
- *Neural Turing Machines and Differentiable Neural Computers*
- *Attention Networks*
- *Kohonen Networks*
- *Capsule Networks*

A descrição e explicações sobre as arquiteturas da Figura 4, podem ser encontradas no artigo de Leijnen e Veen (2020), a quem agradecemos por terem autorizado a reprodução da ilustração no presente estudo.

Conforme mencionado na própria Figura elaborada por Leijnen e Veen, a ilustração mostra uma lista bastante completa de tipos de redes neurais artificiais.

Figura 4 - Taxonomia e arquiteturas de redes neurais artificiais. Esta Figura foi elaborada por Leijnen e Veen (2020)

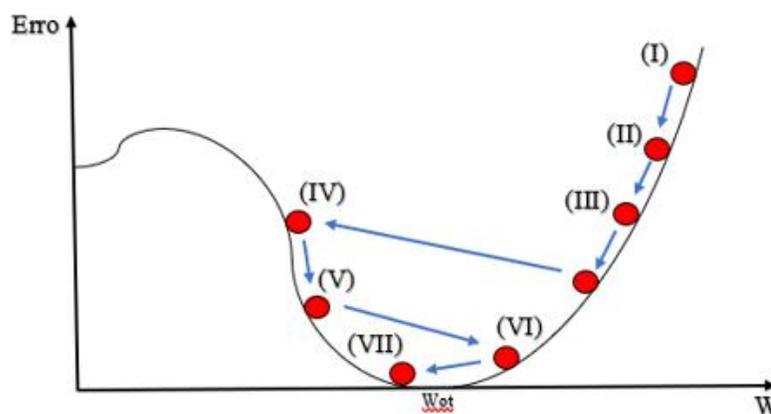


2.4. Algoritmos em redes neurais artificiais

Independentemente do tipo de pesquisa, existem muitas abordagens bem estabelecidas para o aprendizado de máquina. Cada uma delas usa uma estrutura algorítmica diferente para otimizar as previsões com base nos dados recebidos. O algoritmo mais utilizado é o *backpropagation*, mas com o tempo algumas variações deste método têm sido utilizadas, com o objetivo de tornar o processo de convergência mais eficiente. E tais aperfeiçoamentos tem-se o método de inserção do termo de *momentum*, o *Levenberg-Marquardt* e o *resiliente-propagation* (SILVA et al., 2016).

O método de *resiliente-propagation* foi desenvolvido por Martin Riedmiller e Heinrich Braun em 1992 e é muito usado para resolver problemas no manejo florestal (GÖRGENS et al., 2009; BINOTI et al., 2013; CASAS et al., 2018), assim o objetivo deste método, ao invés de considerar as variações das magnitudes do gradiente da função erro, levará somente em conta a variação de seu sinal (RIEDMILLER e BRAUN, 1993). Os processos de convergência do método *resiliente-propagation*, conforme Silva et al. (2016), são apresentados na Figura 5. Nesta Figura, (I), (II), (III) e (V) implicam em incrementos positivos (crescentes) na taxa de aprendizado e possuem variações de sinais de gradiente também positivos, enquanto (IV), (VI) e (VII) implicam em incrementos negativos (decrementos) e variações de sinais igualmente negativas.

Figura 5 - Ilustração do processo de convergência do método *resiliente-propagation*. Este gráfico foi elaborado por Silva et al., 2016.



O algoritmo de gradiente descendente é muito utilizado em redes neurais clássicas e profundas. Algumas estratégias desse método são (AGGARWAL, 2018):

- Diminuição da taxa de aprendizagem.
- Aprendizagem baseada em momentum (Nesterov Momentum).
- Taxa de aprendizagem de específica de parâmetro (AdaGrad, RMSProp, RMSProp com Nesterov Momentum, AdaDelta, Adam).
- Cliffs and Higher-Order Instability
- Gradient Clipping
- Derivadas de segunda ordem (Gradiente conjugado e otimização livre de hessian, método quase newton e BFGS, *saddle points*).
- Polyak Averaging
- Local and Spurious Minima

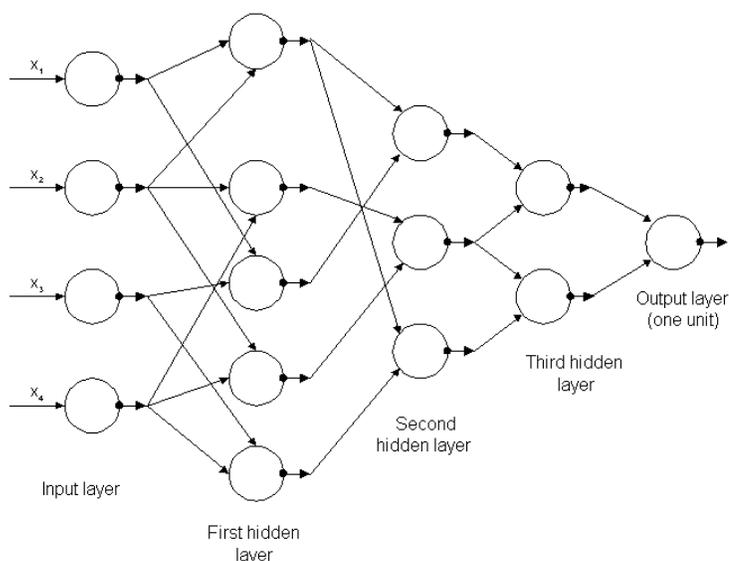
As descrições e explicações das estratégias do algoritmo gradiente descendente são encontrados no livro “*Neural networks and deep learning*” (AGGARWAL, 2018).

2.5. Redes neurais artificiais profundas

Os primeiros algoritmos de aprendizado profundo com múltiplas camadas de atribuições não lineares (Figura 6) têm suas origens com Alekseï Grigor’evich Ivakhnenko e Valentin Grigor’evich Lapa na década dos sessenta, autores do trabalho *Cybernetic*

predicting devices (IVAKHNENKO e LAPA, 1966) e *Cybernetics and Forecasting Techniques* (IVAKHNENKO e LAPA, 1967).

Figura 6 - Estrutura da primeira rede neural profunda treinada por Alekseï Grigor'evich Ivakhnenko. Fonte: (DETTMERS, 2015)



Deep Learning é inspirada em uma das técnicas mais disseminadas de Aprendizado de Máquina, as Redes Neurais Artificiais, e tem sido alvo de estudos em distintas áreas: reconhecimento de imagens (NAIR e HINTON, 2010), áudio (NOROUZI, 2009), caracteres (HINTON e SALAKHUTDINOV, 2006a) e reconhecimento facial (FAN et al., 2014). Não somente no meio acadêmico encontram-se aplicações fazendo uso de *Deep Learning*, como grandes corporações também têm dedicado esforços em projetos que a utilizam: Google (1), com carros que dispensam motoristas, óculos-computadores; Microsoft (2), com o Projeto Adam, ferramenta que promete eficiência e rapidez no reconhecimento visual; e Baidu (3) que criou um laboratório de *Deep Learning* para seus projetos, que também incluem motor de buscas e carro inteligente.

O aprendizado profundo permite que modelos computacionais compostos por várias camadas de processamento aprendam representações de dados com vários níveis de

abstração, que melhoraram drasticamente o estado da arte em reconhecimento de fala, reconhecimento visual de objetos, detecção de objetos e muitos outros domínios (HINTON et al., 2006b; RANZATO et al., 2007; BENGIO et al., 2007; GOODFELLOW et al., 2016; KRIZHEVSKY et al., 2012, 2017).

Métodos de aprendizado profundo são métodos de aprendizado de representação com vários níveis de representação, obtidos através da composição de módulos simples, que são compostos por transformações lineares e não lineares com o objetivo de produzir uma representação mais abstrata e útil (BENGIO et al., 2013).

O aprendizado profundo está fazendo grandes avanços na solução de problemas que resistiram às melhores tentativas da comunidade de inteligência artificial por muitos anos. Ele se mostrou muito bom em descobrir estruturas complexas em dados de alta dimensão e, portanto, é aplicável a muitos domínios da ciência, negócios e governo, inclusive a grandes companhias mundiais como *Google* e *Facebook* desenvolvem seu uso tecnológico com *deep learning* (TAIGMAN et al., 2014).

Se acredita que o aprendizado profundo terá muito mais sucessos no futuro próximo, porque requer muito pouca engenharia manualmente, para que possa tirar proveito fácil dos aumentos na quantidade de computação e dados disponíveis. Novos algoritmos e arquiteturas de aprendizado que estão sendo desenvolvidos para redes neurais profundas apenas acelerarão esse progresso, os professores e alunos fornece a definir tarefas e desafios que incentivem o desenvolvimento (HATTIE, 2002).

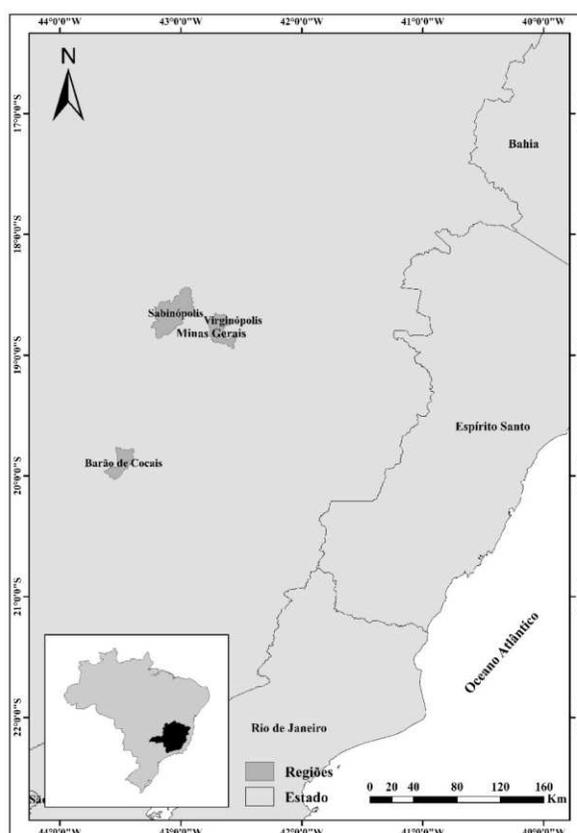
O conceito de *Deep Learning* é abrangente, podendo ser definido como uma sub-área de Aprendizado de Máquina, caracterizado pela utilização de várias camadas de informações não-lineares para extração de características (de forma supervisionada ou não-supervisionada), transformação e análise de padrões; uso de algoritmos de aprendizado de múltiplos níveis de representação para identificar relações entre os dados através de modelos estatísticos, normalmente fazendo uso de redes neurais (DENG e YU, 2014). *Deep Learning* teve o seu desenvolvimento favorecido pela busca por algoritmos de aprendizado não-supervisionados para aprender características de dados não-rotulados (HINTON, et al., 2006b; BENGIO, 2009; YANG, 2013).

3. MATERIAL E MÉTODOS

3.1. Origem e organização dos dados

Os dados utilizados neste estudo foram obtidos de 878 parcelas permanentes, com área em torno de 400 m², mensuradas em povoamentos de eucalipto localizados nas regiões Cocais (46 projetos), Sabinópolis (32 projetos) e Virginópolis (31 projetos), em Minas Gerais, Brasil (Figura 7). Em cada parcela e idade de medição foram determinadas as variáveis: altura dominante (m), área basal (m²ha⁻¹) e volume (m³ha⁻¹). As medições foram feitas em idades que variaram de 19 a 96 meses e os genótipos cultivados foram clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis*.

Figura 7 - Localização das regiões de área de estudo no estado de Minas Gerais, Brasil



O banco de dados para treinamento e validação foi organizado considerando todos os intervalos de idade possíveis para cada parcela, ou seja, dados pareados considerando todos os intervalos de idades ascendentes possíveis ($I_1-I_2, I_1-I_3, \dots, I_2-I_3, I_2-I_4, \dots, I_{n-1}-I_n$). Isto foi necessário para que as redes fossem treinadas visando generalização para diferentes idades iniciais e intervalos de projeção.

Os dados foram separados em duas sub-amostras de 50% das parcelas. A separação foi feita utilizando o software *Select* 2.0 (BINOTI, 2012a) tendo como meta a minimização da raiz quadrada do erro quadrático médio ao ajustar o modelo de Clutter. Foi utilizado o algoritmo de busca aleatória disponível no software, com 100 simulações. Desse modo, 50% das parcelas foram utilizadas para ajuste e 50% para validação.

3.2. Índices de sítio e ajuste dos modelos de crescimento e produção

Os índices de sítio de cada parcela e medição foram determinados pelo método da curva-guia (CLUTTER et al., 1983). O modelo utilizado foi o de Gompertz e a idade-índice foi de 72 meses. Assim, considerando a transformação clássica do método da curva-guia, $Hd_2 = Hd_1 f(I_2) / f(I_1)$, em que I_1 é a idade de medição e I_2 a idade-índice, e o modelo Gompertz,

$HD = \beta_0 e^{-e^{\beta_1 - \beta_2 I}} + \varepsilon$, determinamos os índices de sítio assim:

$$S = Hd e^{-e^{\beta_1 - 72\beta_2}} / e^{\beta_1 - \beta_2 I} \quad (1)$$

em que Hd é a altura dominante (m) observada na idade I , S é o índice de sitio (m) e β_i são as estimativas dos parâmetros do modelo Gompertz.

Foi ajustado o modelo de crescimento e produção de Clutter (1963) para servir de base comparativa, sendo:

$$\ln B_2 = \ln B_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) + \alpha_1 \left(1 - \frac{I_1}{I_2} \right) S_1 + \varepsilon_1 \quad (2)$$

$$\ln V_2 = \beta_0 + \beta_1 \left(\frac{1}{I_2} \right) + \beta_2 S_1 + \beta_3 \ln B_2 + \varepsilon_2 \quad (3)$$

em que

V_2 = produção futura (m^3ha^{-1});

I_1 e I_2 = idades atual e futura (meses);

S_1 = índice de sitio na idade atual (m);

B_1 e B_2 = áreas basais atual e futura (m^2ha^{-1});

Ln = logaritmo neperiano;

α e β = parâmetros;

ε_1 e ε_2 ; $\varepsilon_k \sim NID(0; \sigma^2)$.

O modelo logístico foi ajustado pelo algoritmo Livenberg-Marquardt e o modelo de Clutter foi ajustado pelo método de mínimos quadrados em dois estágios. O software *EViews* 6 (QUANTITATIVE MICRO SOFTWARE, 2007) foi utilizado para o ajustamento do modelo de Clutter e as análises gráficas de resíduos, para fins de verificação e validação, foram feitas com o software *R Studio* (R CORE TEAM, 2020) utilizando o pacote *ggplot2* (WICKHAM, 2016).

3.3. Configuração e treinamento das redes neurais artificiais

Rede neural clássica

Para o treino das redes neurais clássica, os dados foram normalizados assim:

$$x' = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})} \quad (4)$$

em que:

X' = valor normalizado,

X_{\min} = mínimo valor dos dados,

X_{\max} = máximo valor dos dados.

Foram treinadas redes neurais artificiais convencionais *Multilayer Perceptrons* (MLP) contendo três camadas: a camada de entrada, a camada oculta e a camada de saída. O

número de neurônios da camada oculta geralmente é definido em função do número de variáveis de entrada e níveis das variáveis categóricas, por:

$$\text{Número de neurônios na camada oculta} = \frac{n + \sum_{i=1}^I \sum_{j=1}^J X_{ij}}{2}, \quad (5)$$

em que n é o número de variáveis contínuas, I é o número de variáveis categóricas e J é o número de níveis da variável i , sendo $X_{ij} = 1$ para $i=1, 2, \dots, I$ e $j = 1, 2, \dots, J$.

Como neste estudo não foi incluída variáveis categóricas, o número de neurônios na camada oculta foi calculado em função ao número de variáveis contínuas, sendo assim:

$$\text{Número de neurônios na camada oculta} = \frac{n}{2} \quad (6)$$

As variáveis de entrada utilizadas foram: idade atual (I_1), idade futura (I_2), área basal atual (B_1), volume atual (V_1) e altura dominante (Hd). A variável de saída foi o volume futuro (V_2).

O algoritmo de treinamento utilizado foi o *Resilient Propagation* (RPROP⁺) e a função de ativação utilizada nas camadas oculta e de saída foi a logística. O critério de parada adotado foi de 0,0001 para o erro médio ou 3.000 para o número de ciclos, abrangendo todas as medições disponíveis. A RNA foi treinada e validada utilizando o software Neuroforest® (versão 4.0) (BINOTI, 2012b).

Rede neural profunda

Para criar e ajustar um modelo de rede neural profunda foi utilizada 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) e as análises gráficas de resíduos

foram geradas com o pacote *ggplot2* (WICKHAM, 2016). As variáveis de entrada foram as mesmas utilizadas no treinamento das redes clássicas e os dados foram padronizados para que tenham uma média zero e variância um, assim:

$$X_{estand} = \frac{x - média(X)}{desvio\ padrão(X)} \quad (7)$$

em que:

X_{estand} = valor normalizado,

X = Valor dos dados,

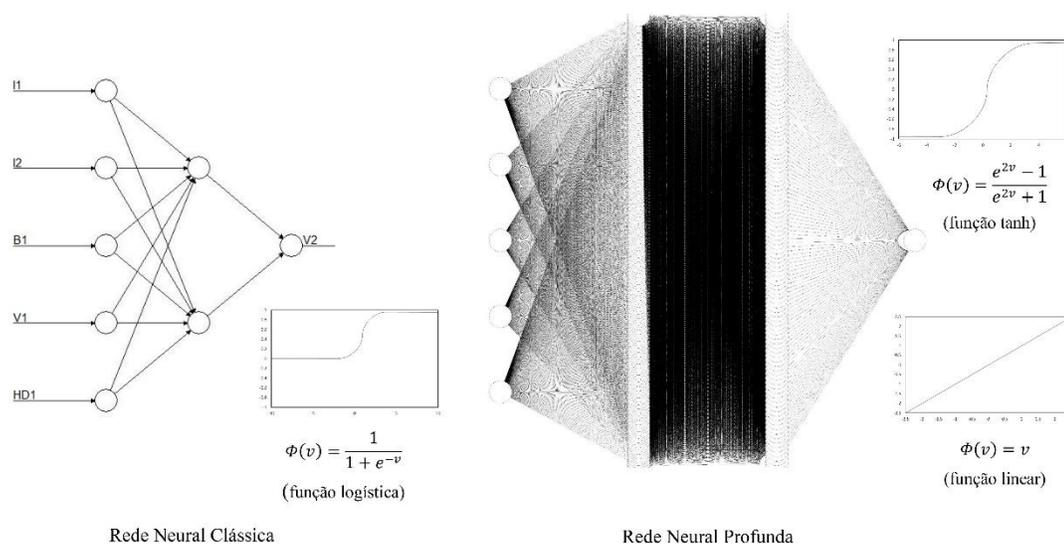
$Media(X)$ = média dos dados

$Desvio\ parado(X)$ = desvio padrão dos dados.

A arquitetura do modelo utilizado com H2O apresenta uma lista de mais de 50 parâmetros a serem configurados. Se criou uma camada entrada, duas camadas ocultas, com 200 e 100 neurônios ou unidades, e uma camada de saída. A função de ativação foi a *Tanh* na camada oculta, a *linear* na camada de saída, a função de erro que se intenta diminuir durante treino da rede foi *Quadratic* (erro quadrático) e a distribuição *gaussiana*. O número de interações de aprendizagem durante o treino foi 200.

O algoritmo de aprendizagem usado foi *aprendizagem adaptativa implementado (ADADELTA)* (ZEILER, 2012). O restante da configuração foi concordo ao padrão. Valores padronizados podem ser encontrados no pacote H2O R, utilizando o seguinte site oficial: <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/deep-learning.html> (última data de acesso: 01 de fevereiro de 2021) e, também, no repositório de consultas do site oficial: <https://github.com/h2oai>. Com esta configuração utilizada foi obtido um número total de 21.401 pesos. A arquitetura das redes clássica e profunda são ilustradas na Figura 8 e na Figura 9 é apresentado um fluxograma dos procedimentos seguidos desde a seleção dos dados até a validação.

Figura 8 - Arquitetura da rede neural clássica e profunda com sua respectiva função de ativação treinadas com dados de clones híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG



A validação foi feita com os dados separados pelo software Select conforme ITEM 3.1. As seguintes estatísticas foram utilizadas na validação: coeficiente de correlação linear entre produções observadas e correspondentes produções preditas (r), média das diferenças absolutas (MDA), raiz quadrada do erro quadrático médio (RQEM%), Bias e erro relativo percentual (ER%). Foram interpretados, também, gráficos de correlação entre produções observadas (Y_i) e estimadas (\hat{Y}_i), de ER% em função da produção observada e da percentagem de casos por classe de ER%. Os estimadores utilizados foram:

$$r_{y\hat{y}} = \frac{n^{-1} \sum_{i=1}^n (Y_{pi} - \hat{Y}_m)(Y_i - \bar{Y})}{\sqrt{n^{-1} \sum_{i=1}^n (Y_{pi} - \hat{Y}_m)^2 n^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad Y_m = n^{-1} \sum_{i=1}^n Y_{pi} \quad (8)$$

$$MDA = \left(n^{-1} \sum_{i=1}^n |Y_{pi} - Y_i| \right) \quad (9)$$

$$RQEM \% = 100 \bar{Y}_i^{-1} \sqrt{n^{-1} \sum_{i=1}^n (Y_i - Y_{pi})^2} \quad (10)$$

$$Bias = \sum_{i=1}^n \frac{(Y_{pi} - Y_i)}{n} \quad (11)$$

$$Bias\% = 100\bar{Y}_i^{-1} \sum_{i=1}^n \frac{(Y_{pi} - Y_i)}{n} \quad (12)$$

$$ER\% = 100 \left(\frac{Y_{pi} - Y_i}{Y_i} \right) \quad (13)$$

em que:

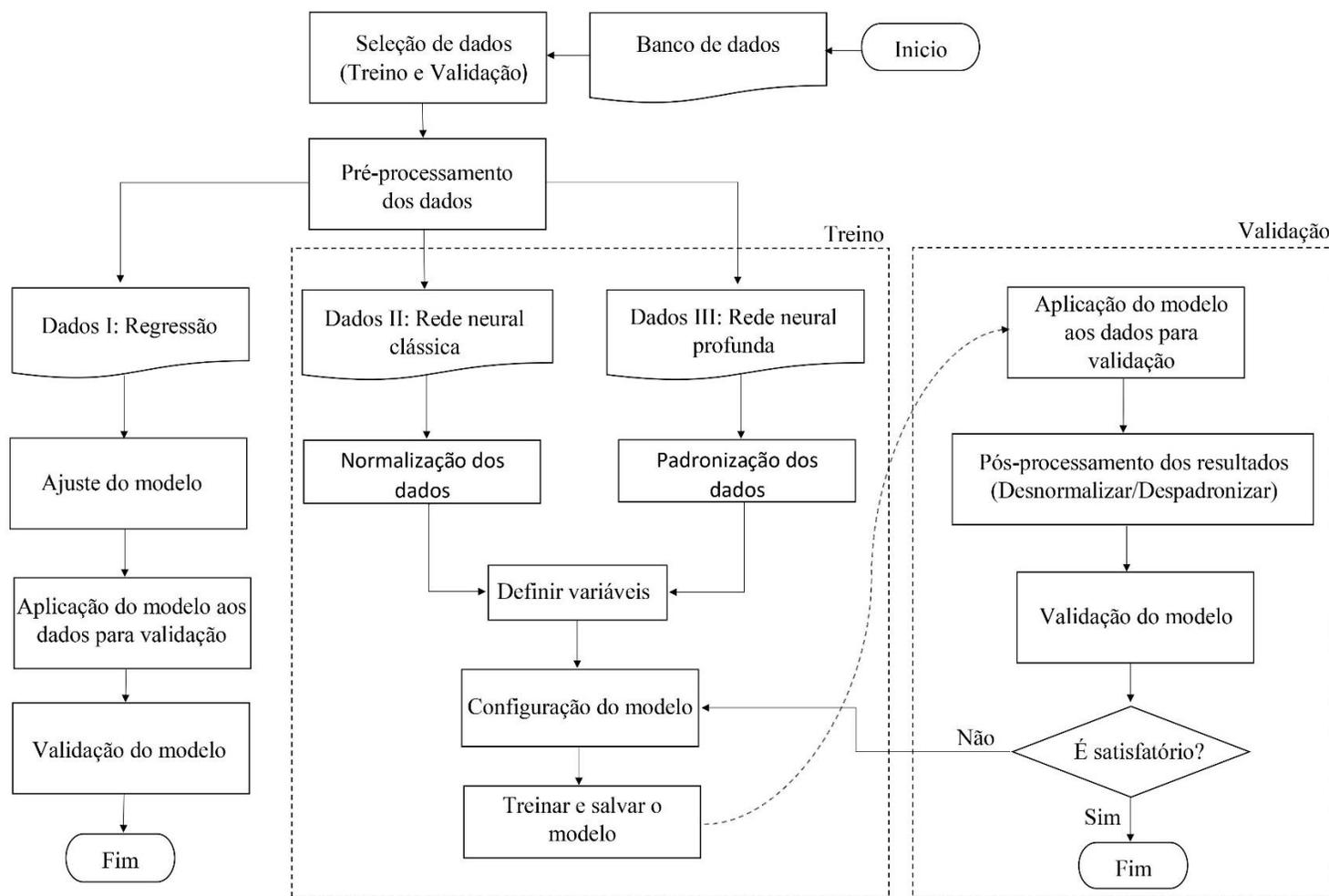
n : número de observações,

Y_{pi} : produção predita,

Y_i : produção observada,

\bar{Y} : média das produções observadas

Figura 9 - Fluxograma dos procedimentos e avaliações realizados neste estudo



4. RESULTADOS

A equação de altura dominante ajustada a toda a base de dados foi:

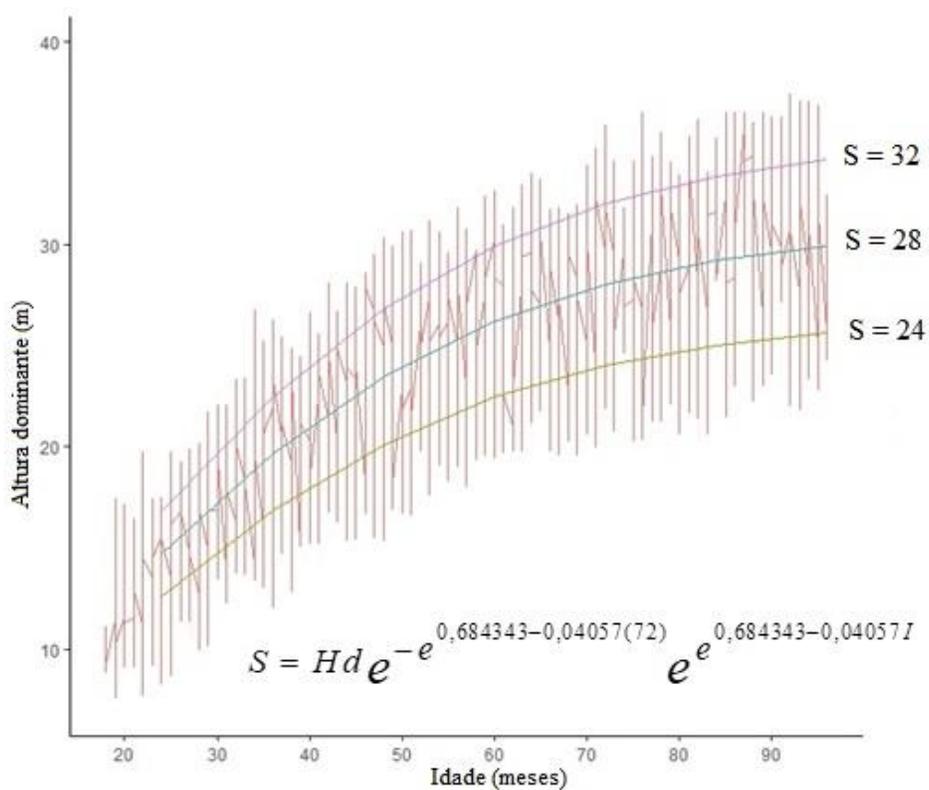
$$Hd = 30,78586e^{-0,684343-0,04057I}, \text{ com } r_{\hat{y}} = 0,8813.$$

Os índices de sítio (S) de cada parcela, em cada medição, foram estimados por:

$$S = Hd e^{-e^{0,684343-0,04057(72)}} e^{e^{0,684343-0,04057I-1}}, \text{ em que } Hd \text{ é a altura}$$

dominante observada na idade I , em meses. A amplitude de alturas dominantes e de índices de sítio podem ser observadas na Figura 10.

Figura 10 - Curvas de índice de sítio para clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG, para idade-índice de 72 meses



4.1. Avaliação dos modelos

As estatísticas de precisão e de exatidão estimadas ao empregar o modelo de Clutter e redes neurais artificiais, com arquiteturas clássica e profunda, são apresentadas na Tabela 1. Nas Figuras 11 a 13 são apresentadas as análises gráficas utilizadas como complemento às estatísticas de validação. Observamos que a RNA profunda resultou em distribuições de resíduos ligeiramente melhores do que aquelas do modelo de Clutter e da RNA clássica (Figuras 11 e 12).

As equações do modelo de Clutter foram:

$$\ln B_2 = \ln B_1 \left(\frac{I_1}{I_2} \right) + 3,8250 \left(1 - \frac{I_1}{I_2} \right) - 0,0059 \left(1 - \frac{I_1}{I_2} \right) S_1 \quad e$$

$$\ln V_2 = 1,7197 - 17,7620 \left(\frac{1}{I_2} \right) + 0,0194 S_1 + 1,1380 \ln B_2$$

Tabela 1 - Resultados estatísticos das estimativas de produção obtidas com emprego de regressão (modelo de Clutter) e redes neurais artificiais clássica e profunda em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

Método	r	MDA (m ³ ha ⁻¹)	RQEM (%)	Bias (m ³ ha ⁻¹)	Bias (%)
Regressão (modelo de Clutter)	0,9311	22,08	13,15	-2,68	-1,172
Rede neural artificial clássica	0,9415	20,59	12,03	-1,37	-0,006
Rede neural artificial profunda	0,9330	22,47	13,28	-8,15	-0,036

Figura 11 - Dispersão das produções estimadas por regressão (modelo de Clutter) e preditas por RNA, em função das produções observadas em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

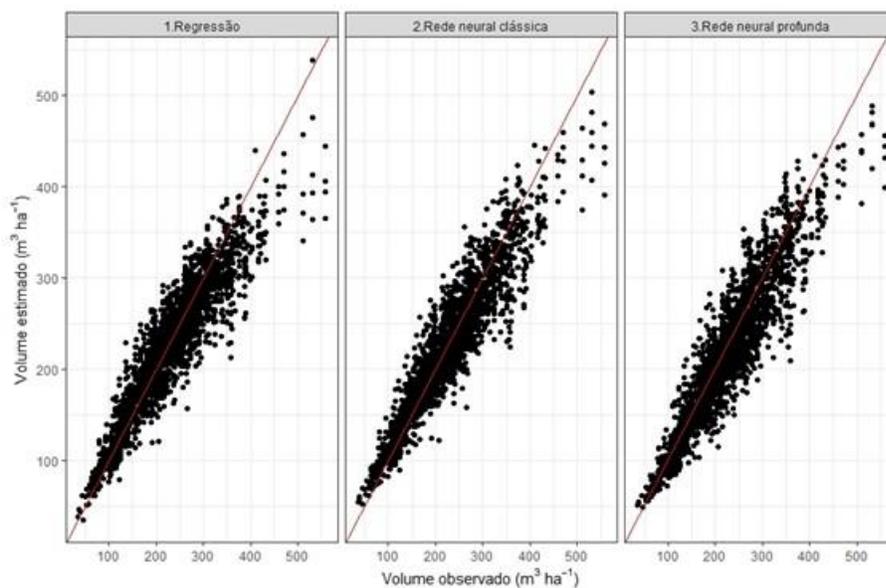


Figura 12 - Dispersão dos erros relativos percentuais em função das produções observadas em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

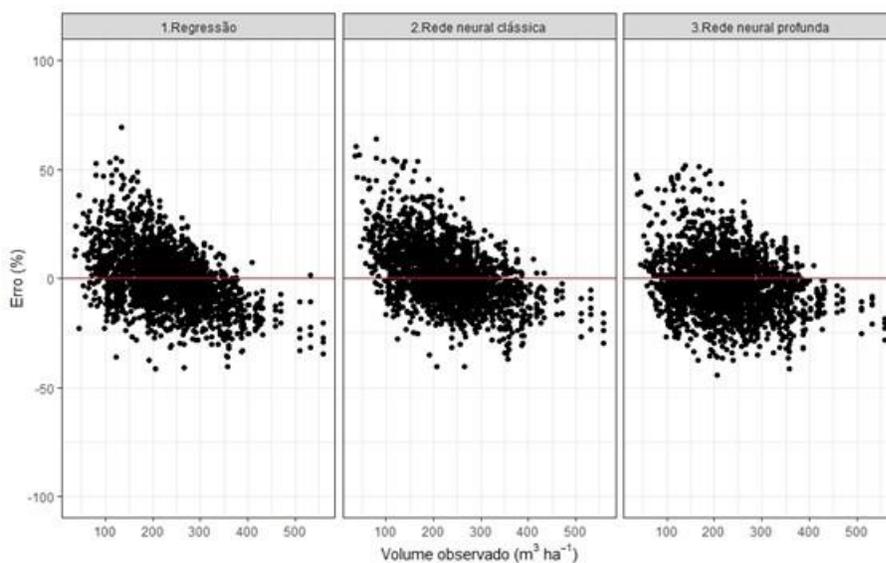
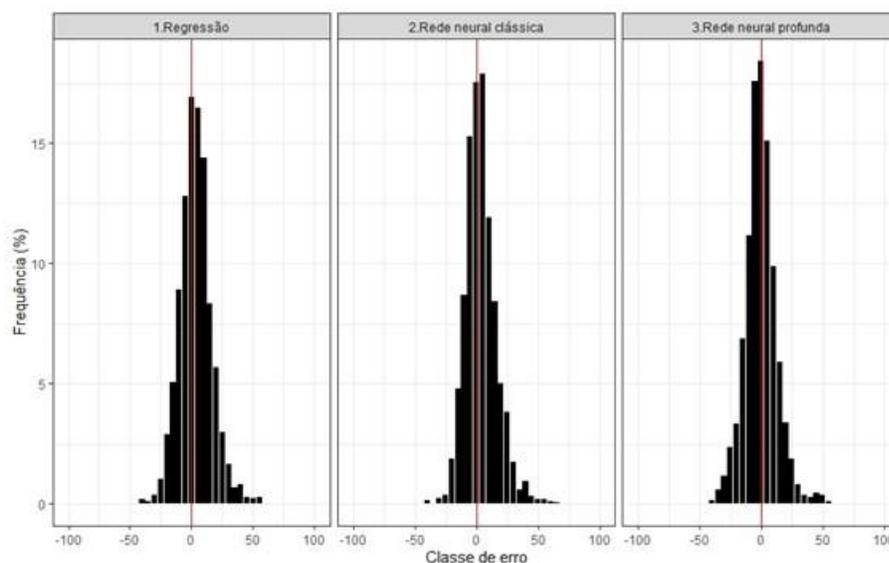


Figura 13 - Percentagem de casos por classe de erro relativo percentual em clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG



4.2. Avaliação das projeções volumétricas a partir de idade precoce

Para avaliar a exatidão das projeções volumétricas se fez uma comparação com os três métodos (Regressão, RNA clássica e RNA profundo), projetando desde uma idade precoce (2 anos) até as idades de projeção de 6 e 7 anos (Tabela 2).

Os três métodos utilizados apresentaram resultados estatísticos com uma mesma tendência em relação ao intervalo de projeção. Quanto maior o intervalo de projeção (diferenças entre as idades final e inicial), menor a exatidão. Porém, para idade inicial de cerca de 2 anos (primeira medição das parcelas permanentes) foram observadas algumas discrepâncias nos resultados obtidos pelos três métodos.

Para entender o comportamento das projeções, se fez os gráficos das tendências das estatísticas (Figura 14 al 17). Se observa que o RQEM% para cada idade de projeção (6 e 7 anos) ao longo das idades iniciais (Figura 14), a RNA apresentou maior acurácia em suas estatísticas resultando em menor variação dos valores ao longo das idades em relação à regressão. A exatidão de modo geral aumenta à medida em que a idade inicial aumenta.

Os três métodos utilizados neste estudo apresentam boas estimações para projeções volumétricas a partir de idade precoce, mas utilizando a inteligência artificial como o método de redes neurais artificiais, tende-se a diminuir o RQEM%, indicador importante nas projeções de volume, podendo-se considerar na toma das decisões e influenciar no plano de manejo. Uma comparação interessante é entre as RNA clássica e de aprendizagem profundo, esta última, o RQEM% é menor na idade inicial de 4 anos (Figura 14) e as bias% (Figura 17) apresenta uma tendência com valores próxima a zero, se aprecia esta influência quando a idade de projeção é de 7 anos.

Figura 14 - Tendência do RQEM% ao longo das idades iniciais para as idades de projeções volumétricas de 6 e 7 anos pelo método de regressão e redes neurais clássica e profunda em clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

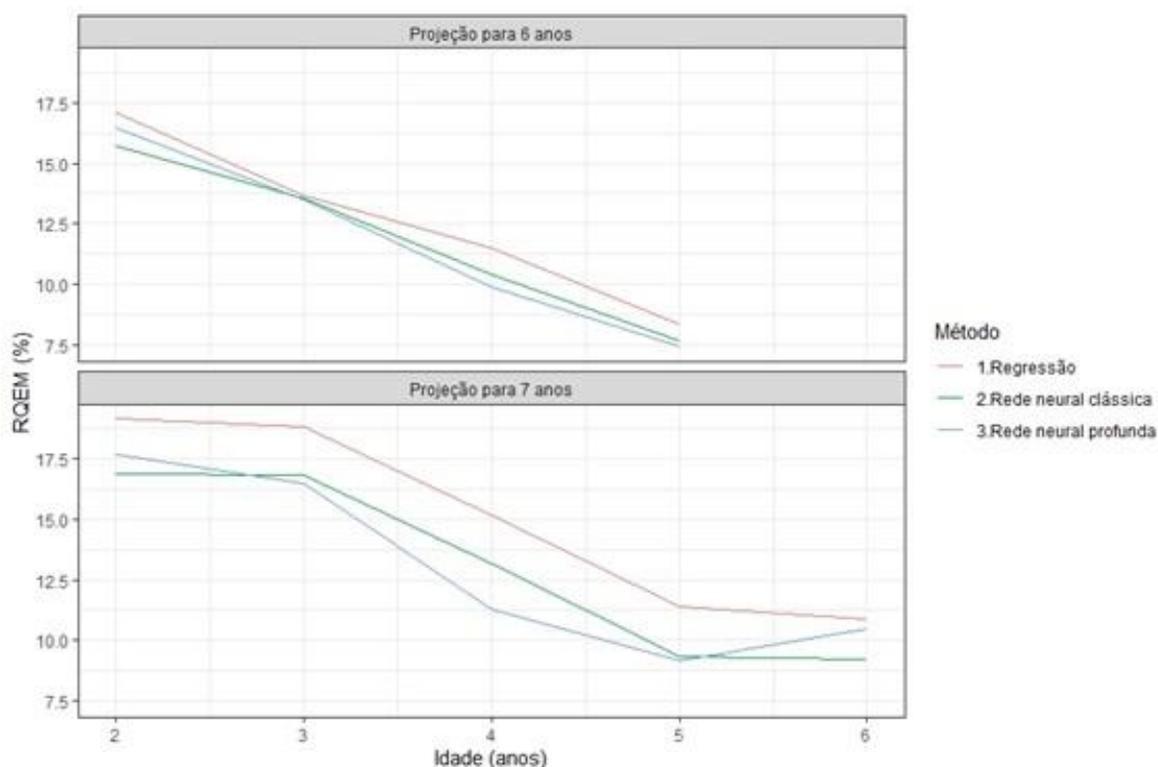


Tabela 2 - Resultados estatísticos pelo método de regressão e RNA clássica e profunda em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG.

Método	Idade de projeção (anos)	Idade inicial (ano)	r	MDA (m ³ ha ⁻¹)	RQEM (%)	Bias (m ³ ha ⁻¹)	Bias (%)
Regressão	6	2	0,7029	35,39	17,12	-14,62	-0,056
		3	0,8383	28,37	13,68	0,79	0,003
		4	0,9052	25,24	11,50	2,71	0,010
		5	0,9373	17,16	8,35	3,45	0,013
	7	2	0,5441	44,98	19,17	-16,96	-0,058
		3	0,8004	46,27	18,83	-16,89	-0,054
		4	0,8939	38,95	15,16	-23,57	-0,072
		5	0,9164	28,01	11,38	-6,81	-0,022
		6	0,9277	23,77	10,85	-8,65	-0,029
Rede neural clássica	6	2	0,7336	32,21	15,70	-7,77	-0,030
		3	0,8276	28,34	13,52	-2,06	-0,008
		4	0,9064	21,89	10,42	-4,61	-0,017
		5	0,9397	16,10	7,66	-2,41	-0,009
	7	2	0,6534	39,03	16,88	-11,52	-0,039
		3	0,8053	42,03	16,82	-8,91	-0,029
		4	0,9035	35,55	13,13	-25,12	-0,077
		5	0,9199	22,19	9,32	-7,11	-0,023
		6	0,9429	19,57	9,23	-6,38	-0,022
Rede neural profunda	6	2	0,7347	34,37	16,49	-13,37	-0,052
		3	0,8374	27,52	13,48	-6,54	-0,025
		4	0,9119	21,49	9,90	2,60	0,009
		5	0,9451	15,18	7,40	0,69	0,003
	7	2	0,5800	41,41	17,68	-4,54	-0,015
		3	0,8178	41,58	16,50	-9,79	-0,031
		4	0,8962	29,88	11,29	-10,49	-0,032
		5	0,9207	21,17	9,16	7,77	0,025
		6	0,9276	21,59	10,50	-8,92	-0,030

Figura 15 - Tendência do Coeficiente de correlação de Pearson ao longo das idades iniciais para as idades de projeções volumétricas de 6 e 7 anos pelo método de regressão e redes neurais clássica e profunda em clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

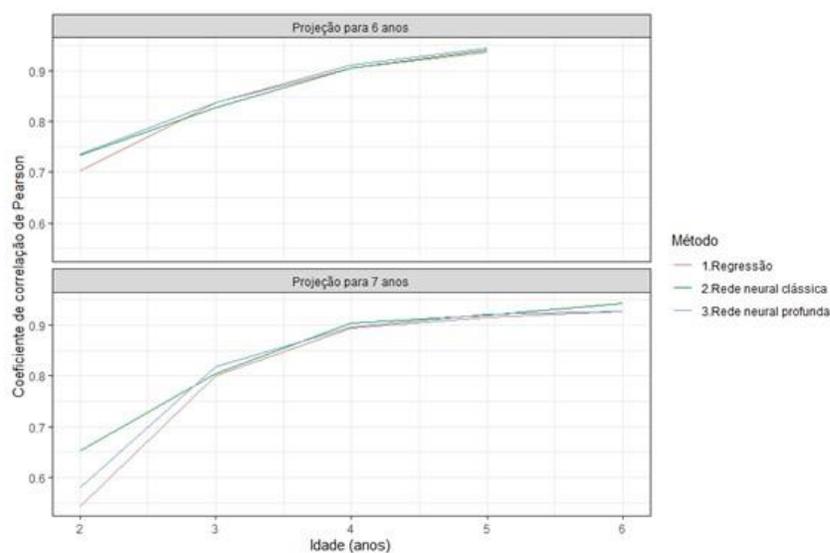


Figura 16 - Tendência do MDA ao longo das idades iniciais para as idades de projeções volumétricas de 6 e 7 anos pelo método de regressão e redes neurais clássica e profunda em clones de híbridos de *Eucalyptus urophylla* X *Eucalyptus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG

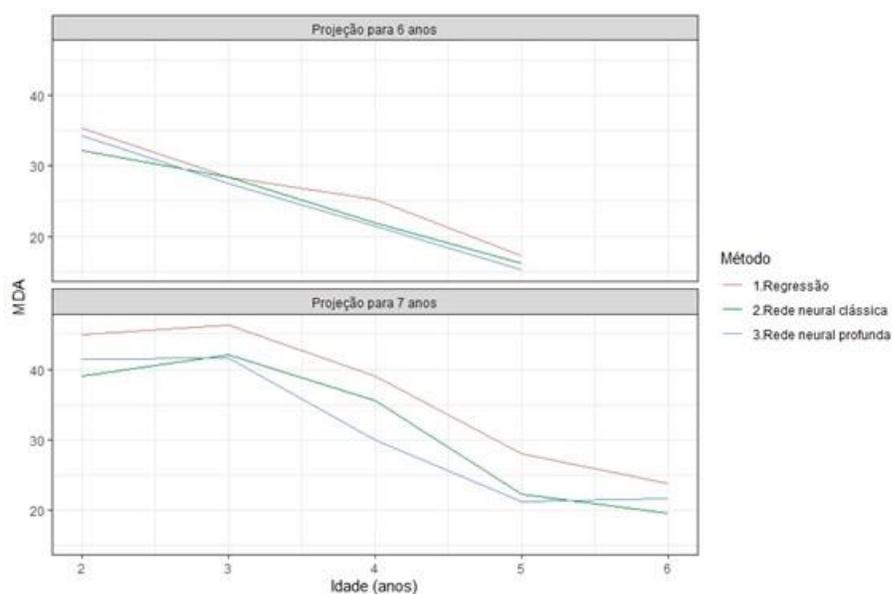
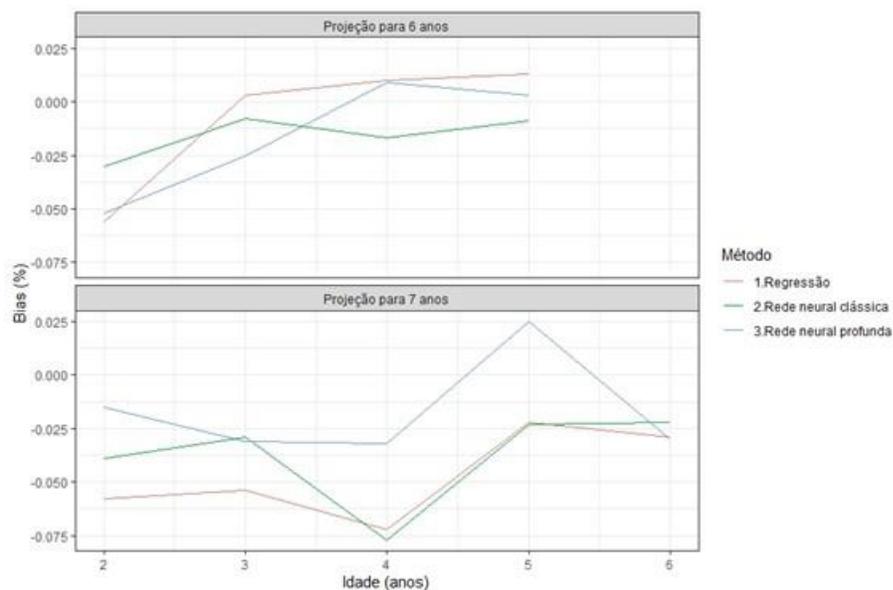


Figura 17 - Tendência do bias% ao logo das idades iniciais para as idades de projeções volumétricas de 6 e 7 anos pelo método de regressão e redes neurais clássica e profunda em clones de híbridos de *Eucalyptus urophylla* X *Eucalytus grandis* nas regiões de Cocais, Sabinópolis e Virginópolis, MG



5. DISCUSSÃO

Neste estudo se comparou regressão (modelo de Clutter) e redes neurais artificiais para modelagem de crescimento e produção de povoamentos de eucalipto. Os resultados mostraram que as três técnicas (ou métodos) avaliadas resultam em estimativas com exatidão satisfatória quando comparadas a estudos similares e que tanto as redes clássicas quanto as de aprendizagem profundo (*deep learning*) resultam em estimativas semelhantes àquelas obtidas com o modelo de Clutter. Porém, nenhuma dessas alternativas resolve o problema da inexatidão de projeções feitas a partir de idade muito precoce (2 a 3 anos).

O emprego de RNA clássicas com arquitetura de multicamadas, com uma única camada oculta e neurônios definidos, treinadas com algoritmo de *Resilient Propagation*, têm

resolvido diversos problemas em manejo florestal, na maioria das vezes com maior exatidão em comparação ao uso de modelos de regressão.

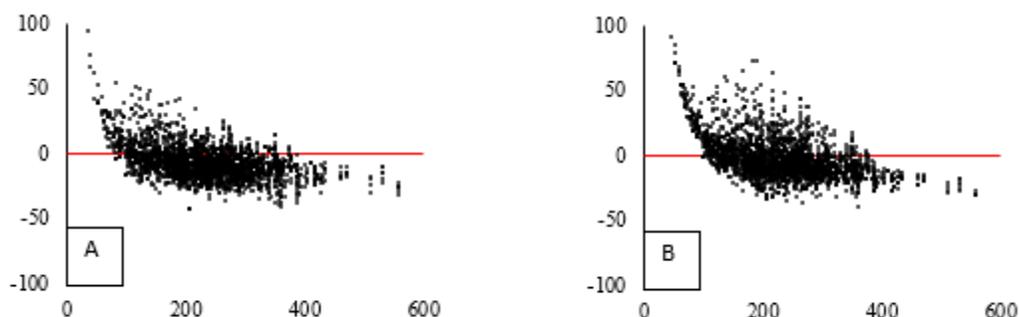
De modo geral as três técnicas resultaram em estimativas com exatidão semelhante ao que tem sido obtido em pesquisas similares (GÖRGENS et al., 2009; DA SILVA et al., 2009; BINOTI et al., 2013; GÖRGENS et al., 2014; BINOTI et al., 2014b; VENDRUSCOLO et al., 2015; LEAL et al., 2015; CAMPOS et al., 2016; DOS REIS MARTINS et al., 2016; RUZA et al., 2017; ALCÂNTRA et al., 2018; CASAS et al., 2018; DÁVILA VEGA, 2018; ROCHA et al., 2018; DA CUNHA NETO et al., 2019; MEDEIROS et al., 2019; TAVARES JUNIOR et al., 2019; FREITAS et al., 2020). A maior exatidão é explicada, em parte, pela inclusão de variáveis categóricas nas RNA e ajuste das regressões sem a estratificação para tais variáveis categóricas. No entanto muitos desafios ainda precisam ser resolvidos. Neste estudo, aprofundamos o aprendizado de uma RNA e utilizando as mesmas variáveis para cada método.

Ainda mais diferenças entre esses métodos podem ser observadas em suas distribuições de resíduos (Figura 12) e em sua frequência a relação com a classe de erro (Figura 13), na qual se observa claramente que o RNA profundo em que sua distribuição se desdobra de maneira robusta e seu histograma atende ao pressuposto de normalidade.

Na RNA profundo, o número de neurônios foi aumentado proporcionalmente de 2: 1 (200-100 neurônios) para apenas uma camada oculta a mais do que um RNA clássico. Embora seja verdade que as estatísticas quantitativas entre esses dois RNAs testados neste estudo sejam semelhantes, o RNA profundo possui uma estrutura hierárquica que os torna particularmente bem adaptados à aprendizagem de hierarquias de conhecimento (NIELSEN, 2015). Essa evidência pode ser observada quando se analisou seus gráficos de distribuição resíduos e histograma de frequência em relação à classe de erro, ou seja, para analisar as variáveis quantitativas não vastas para tomar decisões de avaliação em uma rede profunda, mas também seu comportamento dos gráficos de erro.

Vários testes também foram realizados, utilizando a mesma configuração, o número de camadas ocultas foi aumentado, porém, com mais camadas, os resultados não foram satisfatórios (Figura 18).

Figura 18 - Distribuição de resíduos testados com 3 camadas ocultas e número de neurônios de 200-100-50 (A) e com 5 camadas ocultas com 200 neurônios cada camada (B) numa configuração de RNA profunda



Aumentar as camadas ocultas aumenta a complexidade do modelo, causando erros de superestimação. Portanto, nossos dados são resolvidos aumentando uma camada oculta, encontrando uma proporcionalidade do número de neurônios e um algoritmo capaz de convergir uma grande quantidade de pesos em um tempo razoável, em nosso estudo foi o *ADADELTA* (ZEILER, 2012).

Para o treino da rede profunda, os dados foram padronizados os valores de entrada e da saída, invés de normalizados como o treino da rede clássica, isto foi para a consideração das faixas das variações dinâmicas na utilização da função de ativação *Tanh* na camada oculta. Esta função é uma função logística redimensionada e deslocada; sua simetria em torno de 0 permite que o algoritmo de treinamento convirja mais rápido (CANDEL et al., 2016).

O uso de tipos das normalizações ou padronizações é critério do autor que deve ser em função das ativações usadas. Por exemplo; não se descarta na camada oculta o uso da função *Tanh* em estudos de modelagem de crescimento e produção, pois esta função tem muitas vantagens (LECUN et al, 2012) em relação a outras, mas força uma antissimetria em torno de 0 que está ausente em neurônios biológicos (GLOROT et al., 2011). A tangente hiperbólica tem um estado estacionário em 0 e, portanto, é preferida do ponto de vista de otimização (LECUN et al., 1998). Por isso, outra forma de garantir seu uso é normalizar o domínio com o princípio dos segmentos proporcionais *teorema de Tales*, $X' = (2(X - X_{min}) / (X_{max} - X_{min}) - 1)$, em que o conjunto dos valores X é convertido para um domínio

proporcional entre -1 e 1, representado a própria faixa das variações dinâmicas dessa função (SILVA et al., 2016).

As projeções volumétricas para idades de projeção de 6 e 7 anos foram analisadas estatisticamente sendo observado que RNA apresenta superioridade em sua avaliação estatística, principalmente quando a projeção é para 7 anos. Porém, ela ainda é tendenciosa para projetar a partir de uma idade muito precoce, de cerca de 2 a 3 anos.

Modelos que preveem a produção futura ajudam a melhorar nossa compreensão do crescimento das árvores nas plantações florestais (LHOTKA, 2017). É recomendável realizar a projeção em idades próximas, pois intervalos mais longos podem causar uma superestimação da produção (WEISKITTEL et al., 2016). Em nosso estudo, confirma-se o que tem sido observado nos trabalhos de modelagem de crescimento: menor exatidão quando a projeção é feita a partir de uma idade precoce (cerca de 2 anos), independentemente do modelo utilizado. A baixa exatidão mencionada ocorre pelo fato da variância de Hd e de V ser relativamente pequena em idades próximas a dois anos. Também, é muito provável que a quantidade de dados não foi suficiente, dificultando a aprendizagem da rede profunda.

Considerando este estudo, a técnica de RNA para projeções volumétricas se destaca, mas mesmo assim, há desafios pela frente na resolução de projeções volumétricas desde idades iniciais. Usando o RNA profundo, foi observado que o RQEM% diminui mais do que com os outros dois métodos, a partir de uma idade inicial de 4 anos. Para todos os casos as projeções são desenvolvidas com idades próximas à idade de projeção, ou seja, projetando de 5 a 6 anos, observa-se bom desempenho, o mesmo acontece quando a idade de projeção é de 7 anos.

Resultados semelhantes foram obtidos por Dávila Vega (2018), que projetou a produção de povoamentos de eucalipto para idades 6 e 7, desde cerca de 2 anos, com técnicas de redes neurais artificiais. A autora evidenciou a dificuldade de se projetar a produção a idades muito precoces. Finalmente, cabe ressaltar que neste estudo não foram utilizadas variáveis categóricas nas redes neurais. O uso de tais variáveis de certo aumenta a exatidão, mas, conforme já observado em diversos estudos, ainda não resolveu a inexatidão comum nas projeções a partir de idades precoces.

6. CONCLUSÕES

O uso de redes neurais profundas não diminui a inexatidão e insegurança das projeções de produção de povoamentos de eucalipto feitas a partir de uma idade de cerca de 2 anos, em comparação à projeções feitas com RNA clássicas ou com o modelo de Clutter.

Se comprovam o que vem sendo observado por outros autores, que é a maior exatidão proporcionada pelas redes neurais artificiais em comparação aos modelos de crescimento e produção, na maioria das vezes, porém, com a dificuldade de se projetar com eficiência estatística a partir de qualquer idade inicial.

Se conclui que ainda é melhor evitar projeções da produção de povoamentos de eucalipto a partir de idade próxima a dois anos. E ainda, que é preciso investir mais em pesquisa visando obter maior assertividade e exatidão na prognose a partir de idades precoces.

7. REFERENCIAS BIBLIOGRÁFICAS

AGGARWAL, C. C. **Neural networks and deep learning**. Springer, 2018.

ALCÂNTRA, A. E. M.; DE ALBUQUERQUE SANTOS, A. C.; DA SILVA, M. L. M.; BINOTI, D. H. B.; SOARES, C. P. B.; GLERIANI, J. M.; LEITE, H. G. Use of artificial neural networks to assess yield projection and average production of eucalyptus stands. **African Journal of Agricultural Research**, v. 13, n. 42, p. 2285-2297, 2018.

ARAÚJO JÚNIOR, C. A.; SOUZA, P. D. D.; ASSIS, A. L. D.; CABACINHA, C. D.; LEITE, H. G.; SOARES, C. P. B.; SILVA, A. A. L.; CASTRO, R. V. O. Artificial neural networks, quantile regression, and linear regression for site index prediction in the presence of outliers. **Pesquisa Agropecuária Brasileira**, v. 54, 2019.

BENGIO, Y. **Learning deep architectures for AI**. Now Publishers Inc, 2009.

BENGIO, Y.; COURVILLE, A.; VINCENT, P. Representation learning: A review and new perspectives. **IEEE transactions on pattern analysis and machine intelligence**, v. 35, n. 8, p. 1798-1828, 2013.

BENGIO, Y.; LAMBLIN, P.; POPOVICI, D.; LAROCHELLE, H. Greedy layer-wise training of deep networks. **Advances in neural information processing systems**, v. 19, p. 153, 2007.

BHARATENDRA, R. **Advanced deep learning with R become an expert at designing, building, and improving advanced neural network models using R**. B3 2PB, UK. Birmingham, 2019.

BINOTI, D. H. B.; DA SILVA B. M. L. M.; LEITE, H. G. Configuração de redes neurais artificiais para estimação do volume de árvores. **Revista Ciência da Madeira (Brazilian Journal of Wood Science)**, v. 5, n. 1, p. 10-12953/2177-6830. v05n01a06, 2014a.

BINOTI, M. L. M. D. S.; BINOTI, D. H. B.; LEITE, H. G.; GARCIA, S. L. R.; FERREIRA, M. Z.; RODE, R.; DA SILVA, A. A. L. Redes neurais artificiais para estimação do volume de árvores. **Revista Árvore**, v. 38, n. 2, p. 283-288, 2014b.

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, 2012a.

BINOTI, M. L. M. D. S. **Use of Artificial Neural Networks in Measurement and Forest Management**. 2012. 138 f. Tese (Doutorado em Manejo Florestal; Meio Ambiente e Conservação da Natureza; Silvicultura; Tecnologia e Utilização de) - Universidade Federal de Viçosa, Viçosa, 2012b.

BINOTI, M. L. M. S.; BINOTI, D.H.B.; LEITE, H.G. Application of artificial neural networks to estimate the height of even-aged stands of eucalyptus. **Revista Árvore**, v. 37, n. 4, p. 639-645, 2013.

BULLINARIA, J. A. **Introduction to Neural Computation** **Introduction to Neural Computation**, 2014. Disponível em: <<http://www.cs.bham.ac.uk/~jxb/inc.html>>. Acesso em: 25 jan. 2018

CAMPOS, B.P.F.; DA SILVA, G.F.; BINOTI, D.H.B.; DE MENDONÇA, A.R.; LEITE, H.G. Predição da altura total de árvores em plantios de diferentes espécies por meio de redes neurais artificiais. **Brazilian Journal of Forest Research/Pesquisa Florestal Brasileira**, v. 36, n. 88, p. 375-385, 2016.

CAMPOS, J.C.C.; LEITE, H.G. **Mensuração florestal: perguntas e respostas**. 5.ed. atual. e ampl. Viçosa: Ed. da UFV, 2017. 636p.

CANDEL, A.; PARMAR, V.; LEDELL, E.; ARORA, A. Deep learning with H2O. **H2O. ai Inc**, 2016.

CASAS, G. G.; VILLANUEVA, J. R. B.; LEITE, H. G.; DE ALBUQUERQUE SANTOS, A. C.; JÚNIOR, C. A. A.; DE ALCÁNTARA, A. E. M. Regulación de la producción forestal con pronósticos de modelación Clutter y redes neuronales artificiales en plantaciones de eucalipto del Estado de Minas Gerais–Brasil. **Revista Científica UNTRM: Ciencias Naturales e Ingeniería**, v. 1, n. 1, p. 55-59, 2018.

CHAGAS, E. T. O. **Aprendizaje profundo y sus aplicaciones hoy**. **Revista Científica Multidisciplinar Núcleo do Conhecimento**. Ed. 05, v. 04, p. 05-26, 2019

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. 333p.

DA CUNHA NETO, E.M.; BEZERRA, J.C.F.; DE MIRANDA, L.C.; DO MAR, A.L.; VAZ, M.M., DA SILVA MELO, M.R.; DA ROCHA, J. E. C. Modelo de kozak e redes neurais artificiais na estimativa do afilamento do fuste de eucalipto. **Revista de engenharia e tecnologia**, v. 11, n. 3, p. 150-157, 2019.

DA SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Ajuste do modelo de Schumacher e Hall e aplicação de redes neurais artificiais para estimar volume de árvores de eucalipto. **Revista Árvore**, v. 33, n. 6, p. 1133-1139, 2009.

DÁVILA VEGA, A. E. **Modelagem de povoamentos de eucalipto no nordeste da Bahia utilizando regressão e redes neurais artificiais**. 2018. 52f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa. 2018.

DENG, L.; YU, D. Deep learning: methods and applications. **Foundations and trends in signal processing**, v. 7, n. 3–4, p. 197-387, 2014.

DETTMERS, T. **Aprendizagem profunda em poucas palavras: Historia e Treinamento**. 2015. Disponível em: <https://devblogs.nvidia.com/deep-learning-nutshell-history-training/>

DOS REIS MARTINS, E.; DA SILVA BINOTI, M. L. M.; LEITE, H. G.; BINOTI, D. H. B.; DUTRA, G.C. Configuração de redes neurais artificiais para estimação do afilamento do fuste de árvores de eucalipto. **Revista Brasileira de Ciências Agrárias**, v. 11, n. 1, p. 33-38, 2016.

FAN, H.; CAO, Z.; JIANG, Y.; YIN, Q.; DOUDOU, C. Learning deep face representation. **arXiv preprint arXiv:1403.2802**, 2014.

FERREIRA, M. P. et al. Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. **Forest Ecology and Management**, v. 475, p. 118397, 2020.

FREITAS, E. C. S.; DE PAIVA, H. N.; NEVES, J. C. L.; MARCATTI, G. E.; LEITE, H. G. Modeling of eucalyptus productivity with artificial neural networks. **Industrial Crops and Products**, v. 146, p. 112149, 2020.

GAVILÁN-ACUÑA, G.; OLMEDO, G. F.; MENA-QUIJADA, P.; GUEVARA, M.; BARRÍA-KNOPF, B.; WATT, M. S. Reducing the Uncertainty of Radiata Pine Site Index Maps Using an Spatial Ensemble of Machine Learning Models. **Forests**, v. 12, n. 1, p. 77, 2021.

GLOROT, X.; BORDES, A.; BENGIO, Y. Deep sparse rectifier neural networks. In: **Proceedings of the fourteenth international conference on artificial intelligence and statistics**. JMLR Workshop and Conference Proceedings, 2011. p. 315-323.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A.; BENGIO, Y. **Deep learning**. Cambridge: MIT press, 2016.

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

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

HAMDI, Z. M; BRANDMEIER, M.; STRAUB, C. Forest damage assessment using deep learning on high resolution remote sensing data. **Remote Sensing**, v. 11, n. 17, p. 1976, 2019.

HATTIE, J. What are the attributes of excellent teachers. **Teachers make a difference: What is the research evidence**, p. 3-26, 2002.

HAYKIN, S.; ENGEL, P. M. **Redes neurais: princípios e prática**. Porto Alegre: Bookman, 2001.

HINTON, G. E.; OSINDERO, S.; TEH, Y. W. A fast learning algorithm for deep belief nets. **Neural computation**, v. 18, n. 7, p. 1527-1554, 2006b.

HINTON, G. E.; SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. **science**, v. 313, n. 5786, p. 504-507, 2006a.

HODGKIN, A. L.; HUXLEY, A. F. Currents carried by sodium and potassium ions through the membrane of the giant axon of Loligo. **The Journal of physiology**, v. 116, n. 4, p. 449, 1952.

IVAKHNENKO, A. G.; LAPA, V. G. **Cybernetics and forecasting techniques**. 1967.

IVAKHNENKO, A. G.; LAPA, V. G. **Cybernetic predicting devices**. Purdue Univ Lafayette Ind School of Electrical Engineering, 1966.

KENTSCH, S.; LOPEZ CACERES, M. L.; SERRANO, D.; ROURE, F.; DIEZ, Y. Computer vision and deep learning techniques for the analysis of drone-acquired forest images, a transfer learning study. **Remote Sensing**, v. 12, n. 8, p. 1287, 2020.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, v. 25, p. 1097-1105, 2012.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. **Communications of the ACM**, v. 60, n. 6, p. 84-90, 2017.

- LEAL, F.A.; MIGUEL, E.P.; MATRICARDI, E.A.T.; PEREIRA, R.S. Redes neurais artificiais na estimativa de volume em um plantio de eucalipto em função de fotografias hemisféricas e número de árvores. **Revista Brasileira de Biometria**, v. 33, n. 2, p. 233-249, 2015.
- LECUN, Y. A.; BOTTOU, L.; ORR, G. B.; MÜLLER, K. R. Efficient backprop. In: **Neural networks: Tricks of the trade**. Springer, Berlin, Heidelberg, 2012. p. 9-48.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, v. 521, n. 7553, p. 436-444, 2015.
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278-2324, 1998.
- 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.
- LEIJNEN, S.; VEEN, F. The Neural Network Zoo. In: **Multidisciplinary Digital Publishing Institute Proceedings**. 2020. p. 9.
- LEITE, R. V. et al. Estimating stem volume in eucalyptus plantations using airborne LiDAR: A comparison of area-and individual tree-based approaches. **Remote Sensing**, v. 12, n. 9, p. 1513, 2020.
- LHOTKA, J. M. Examining growth relationships in Quercus stands: An application of individual-tree models developed from long-term thinning experiments. **Forest ecology and management**, v. 385, p. 65-77, 2017.
- LOPES, L. S. S. et al. Uso de regressão e redes neurais artificiais na estimativa do volume de Khaya ivorensis. **Revista Ciência da Madeira (Brazilian Journal of Wood Science)**, v. 11, n. 2, 2020.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, n. 4, p. 115-133, 1943.

MEDEIROS, R. A.; DE PAIVA, H. N.; SOARES, Á. A. V.; MARCATTI, G. E.; TAKIZAWA, F. H.; DOMICIANO, C. A. R.; LEITE, H. G. Productive potential of *Tectona grandis* in Midwest Brazil. **Advances in Forestry Science**, v. 6, n. 4, p. 803-809, 2019.

NAIR, V.; HINTON G., E. Rectified linear units improve restricted boltzmann machines. In: **ICML**. 2010.

NIELSEN, M. A. **Neural networks and deep learning**. San Francisco, CA: Determination press, 2015.

NOROUZI, M. **Convolutional Restricted Boltzmann Machines for Feature Learning**. Master of Science—Canada: Simon Fraser University, 2009.

QUANTITATIVE MICRO SOFTWARE. **EViews 6 Command Reference**. Irvine CA, USA, 2007.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria: R Foundation for Statistical Computing, 2020.

RANZATO, M. A.; HUANG, F. J.; BOUREAU, Y. L.; LECUN, Y. Unsupervised learning of invariant feature hierarchies with applications to object recognition. In: **2007 IEEE conference on computer vision and pattern recognition**. IEEE, 2007. p. 1-8.

RIEDMILLER, M.; BRAUN, H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In: **IEEE international conference on neural networks**. IEEE, 1993. p. 586-591.

ROCHA, S. J. S. S.; TORRES, C. M. M. E.; JACOVINE, L. A. G.; LEITE, H. G.; GELCER, E. M.; NEVES, K.; SCHETTINI, B. L. S.; VILLANOVA, P. H.; SILVA, L. F.; REIS, L. P.; ZANUNCIO, J. C. Artificial neural networks: Modeling tree survival and mortality in the Atlantic Forest biome in Brazil. **Science of the total environment**, v. 645, p. 655-661, 2018.

RUZA, M.S.; DALLA CORTE, A.P.; HENTZ, A.M.K.; SANQUETTA, C.R.; SILVA, C.A.; SCHOENINGER, E.R. Inventário de Sobrevivência de povoamento de *Eucalyptus* com uso de Redes Neurais Artificiais em Fotografias obtidas por VANTs. **Advances in Forestry Science**, v. 4, n. 1, p. 83-88, 2017.

SANTANA, M. *Deep Learning: do Conceito às Aplicações*. 2018. Disponível em: <https://medium.com/data-hackers/deep-learning-do-conceito-%C3%A0s->

SILVA, I., N.; SPATTI, D., H.; FLAUZINO, R., A. **Redes Neurais Artificiais para engenharia e ciências aplicadas curso prático**. 2ª edição revisada e ampliada. São Paulo: Artliber, 2016. 431p.

SILVA, S.; DE OLIVEIRA NETO, S. N.; LEITE, H. G.; DE ALCÂNTARA, A. E. M.; DE OLIVEIRA NETO, R. R.; DE SOUZA, G. S. A. Productivity estimate using regression and artificial neural networks in small familiar areas with agrosilvopastoral systems. *Agroforestry Systems*, v. 94, n. 6, p. 2081-2097, 2020.

SOCHER, R.; MUNDRA, R. S. **CS 224D: Deep Learning for NLP1**. 2016.

SOUZA, G. S. A. et al. Multi-sensor prediction of Eucalyptus stand volume: A support vector approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 156, p. 135-146, 2019.

SUN, Y.; LIU, Y.; WANG, G.; ZHANG, H. Deep learning for plant identification in natural environment. *Computational intelligence and neuroscience*, v. 2017, 2017.

TAIGMAN, Y.; YANG, M.; RANZATO, M. A.; WOLF, L.. Deepface: Closing the gap to human-level performance in face verification. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2014. p. 1701-1708.

TAVARES JUNIOR, I.S.; ROCHA, J.E.C.; EBLING, A.A.; CHAVE, A.S.; ZANUNCIO, J. C.; FARIAS, A. A.; LEITE, H. G. Artificial neural networks and linear regression reduce sample intensity to predict the commercial volume of eucalyptus clones. *Forests*, v. 10, n. 3, p. 268, 2019.

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

WEISKITTEL, A.; KUEHNE, C.; MCTAGUE, J. P.; OPPENHEIMER, M. Development and evaluation of an individual tree growth and yield model for the mixed species forest of the Adirondacks Region of New York, USA. **Forest Ecosystems**, 2016, vol. 3, n. 1, p. 26.

WICKHAM, H. **ggplot2: Elegant Graphics for Data Analysis**. Springer-Verlag New York, 2016.

XI, Z.; HOPKINSON, C.; ROOD, S. B.; PEDDLE, D. R. See the forest and the trees: Effective machine and deep learning algorithms for wood filtering and tree species classification from terrestrial laser scanning. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 168, p. 1-16, 2020.

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.

YANG, Y. **Learning Hierarchical Representations For Video Analysis Using Deep Learning**. Doctor of Philosophy (Ph.D.)—Orlando, Florida: University of Central Florida, 2013.

ZEILER, Matthew D. Adadelta: an adaptive learning rate method. **arXiv preprint arXiv:1212.5701**, 2012.

CONCLUSÕES GERAIS

A hipótese H_{o1} : a exatidão das estimativas do modelo de Clutter independe da forma de organização dos dados das parcelas permanentes é rejeitada;

A hipótese H_{o2} : a exatidão das estimativas do modelo de Buckman independe da forma de organização dos dados das parcelas permanentes é rejeitada;

A hipótese H_{o3} : a inexatidão de produções obtidas com projeções feitas para idade de 6 e 7 anos em povoamentos de eucalipto, a partir de idade precoce (cerca de 2 anos), independe do método utilizado: regressão, redes neurais artificiais ou redes profundas não foi rejeitada;

Banco de dados para ajuste do modelo de Buckman devem conter os dados pareados mês a mês.