

**José Wesley Lima Silva**

**Modelos estatísticos e técnicas de inteligência artificial para  
estimativa do volume de clones de *Eucalyptus* spp. com adição de  
variáveis climáticas**

**Recife**

**28 de fevereiro de 2020**



**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

**Modelos estatísticos e técnicas de inteligência artificial para estimativa do volume de clones de *Eucalyptus* spp. com adição de variáveis climáticas**

**Tese julgada adequada para obtenção do título de Doutor em Biometria e Estatística Aplicada defendida e aprovada por unanimidade em 28/02/2020 pela comissão examinadora.**

**Área de concentração: Biometria e Estatística Aplicada**

**Orientador: Prof. PhD. José Antônio Aleixo da Silva**

**Recife**

**28 de fevereiro de 2020**

Dados Internacionais de Catalogação na Publicação  
Universidade Federal Rural de Pernambuco  
Sistema Integrado de Bibliotecas  
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

---

- J83m Silva, José Wesley Lima  
Modelos estatísticos e técnicas de inteligência artificial para estimativa do volume de clones de *Eucalyptus* spp. com adição de variáveis climáticas/ José Wesley Lima Silva. – 2020.  
153 f. : il.
- Orientador: Prof. PhD. José Antônio Aleixo da Silva  
Inclui referências.
- Tese (Doutorado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Biometria e Estatística Aplicada, Recife, 2020.
1. Produção. 2. Mudanças Climáticas. 3. Modelos não Lineares Generalizados. 4. Redes Neurais Artificiais. 5. Aprendizagem de Máquinas. I. Silva, Jose Antonio Aleixo da, orient. II. Título

CDD 519.5

---

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

**Modelos estatísticos e técnicas de inteligência artificial para estimativa do volume de clones de *Eucalyptus* spp. com adição de variáveis climáticas**

José Wesley Lima Silva

Tese julgada adequada para obtenção do título de Doutor em Biometria e Estatística Aplicada defendida e aprovada por unanimidade em 28/02/2020 pela comissão examinadora.

Orientador:

---

**Prof. PhD. José Antônio Aleixo da Silva**  
Orientador

Banca examinadora:

---

**Prof. Dr. César Augusto Guimarães**  
**Finger**  
**Universidade Federal de Santa Maria**

---

**Prof. Dr. Rinaldo Luiz Caraciolo Ferreira**  
**Universidade Federal Rural de**  
**Pernambuco**

---

**Dr. Ouorou Ganni Mariel Guera**  
**Instituto Politécnico**  
**Nacional/CEPROBI/México**

---

**Prof. Dr. Fernando Henrique de Lima**  
**Gadelha**  
**Instituto Federal de Educação, Ciência**  
**e Tecnologia de Pernambuco**

*Dedico esse trabalho a minha  
família, porto seguro, na  
minha caminhada.*

# Agradecimentos

A Deus pela minha salvação e por me proporcionar oportunidades grandiosas, sem Ele nada seria possível.

Ao meu orientador Professor José Antônio Aleixo da Silva, pelo esforço, os ensinamentos que me passou de forma muito generosa, a confiança e disponibilidade durante todo o desenvolvimento da orientação.

À minha família, em especial minha mãe por se dispor mesmo sem condições a me ajudar e dar o melhor que uma mãe poderia dar, ao meu pai que é o meu orgulho e exemplo para todas as coisas, à minha vó querida. Ao meu irmão William e sua esposa Alanne pelas conversas de apoio que sempre me ajudaram.

Aos professores do programa de pós-graduação em Biometria e Estatística Aplicada pelos conhecimentos transmitidos e pela disposição em ajudar, ao Coordenador do programa Dr. Moacyr Cunha Filho.

Ao Secretário do programa Marco Antônio dos Santos por sua competência administrativa e pronto atendimento aos inúmeros pedidos burocráticos.

Ao chefe da Estação Experimental do Instituto Agrônomo de Pernambuco (IPA) - Araripina, Eng. Agrônomo José Alves Tavares pelo apoio durante toda a pesquisa. Também, todos os funcionários do IPA que contribuíram grandemente nos trabalhos de campo. Bem como, à Gerência de Meteorologia e Mudanças Climáticas da Agência Pernambucana de Águas e Clima (APAC) na pessoa do senhor Josafá Gomes por todo o apoio na disponibilização das bases de dados climáticos utilizados no estudo.

Aos colegas do DEINFO e do DCFL Glauce, Yara, Glaucia, Lêda, Géssyca, Gisele, Gabi, Augusto, Henrique, Tibério, Luan, Djailson, Júlio e Janilson pela companhia, amizade, troca de experiência e pelos momentos felizes que me proporcionaram.

Aos companheiros do trabalho em campo Yara, Géssyca, Raíssa, Marcella, Mayara, Aline Amorin, Aline Veloso, Sandra, Luan, Nathan, Leonardo, Pedro, Edinho e Robson, que tornaram o cansaço e os momentos difíceis bem mais fáceis de serem suportados.

À CAPES, pela concessão da bolsa de estudos.

À Universidade Federal Rural de Pernambuco, que me proporcionou cursar o doutorado em Biometria e Estatística Aplicada.

*“O temor do SENHOR  
é o princípio da sabedoria;  
revelam prudência todos os que o praticam.  
O seu louvor permanece para sempre.”  
(Bíblia Sagrada, Salmos 111, 10)*

# Resumo

O Polo Gesseiro do Araripe é o principal produtor de gesso do Brasil. Para o processamento do minério gipsita a principal fonte energética é a lenha proveniente da Caatinga que tem sido explorada geralmente de forma ilegal. Nos dias atuais os efeitos do desmatamento da Caatinga combinado com as mudanças climáticas tornou a oferta dos recursos energéticos escassa. Como forma de mitigar o efeito do desmatamento e diminuir a supressão da vegetação Caatinga, experimentos com florestas de rápido crescimento foram implantadas na região. Para proporcionar um desenvolvimento mais adequado do manejo dessas florestas se faz necessário o ajuste de modelos estatísticos locais mais eficientes e que levem em consideração a mudança nas variáveis do clima. Desta forma, objetivou-se com esta pesquisa ajustar modelos não lineares e algoritmos da aprendizagem de máquinas e avaliar se a inclusão de variáveis climáticas gera equações com maior precisão e poder de generalização na estimativa da produção e do crescimento volumétrico de clones de *Eucalyptus* em região de clima semiárido do Nordeste brasileiro. Os dados foram obtidos por meio de uma pesquisa realizada na Estação Experimental do IPA, município de Araripina - PE, com três clones de *Eucalyptus* spp. (C39, C41, C11) plantados nos espaçamentos 3 m x 3 m, 4 m x 2 m, 3 m x 2 m, 2 m x 2 m e 2 m, x 1 m, em experimento multivariado de medidas repetidas, implantado no ano de 2010. A cada seis meses foram realizadas medições de altura e diâmetro à altura do peito. Na idade de 96 meses o experimento foi cortado e o volume de madeira foi cubado rigorosamente pelo método de Smalian. A base de dados do período chuvoso (2002 à 2009) foi obtida de um experimento desenvolvido na mesma região com os mesmos tratos silviculturais. Foram empregados modelos não lineares e algoritmos da inteligência artificial para produção, crescimento e prognose do volume. Os ajustes foram avaliados com base no índice de ajuste de Schlaegel corrigido ( $IA_c$ ), índice de Furnival (IF), erro padrão de estimativa em porcentagem ( $S_{xy}\%$ ), raiz do erro quadrático médio (REQM), erro absoluto médio (EAM) e análise gráfica dos resíduos. A produção volumétrica se diferenciou entre clones e entre os diferentes níveis de adensamento. Os modelos simétricos e os modelos não lineares generalizados produziram equações com altos valores de  $IA_c$  e baixos valores de  $S_{xy}\%$  e com resíduos atendendo aos pressupostos da regressão. As técnicas de inteligência artificial se adequaram com eficiência a estimativa da produção e do crescimento em volume gerando estatísticas de boa qualidade. A inclusão das variáveis climáticas nos modelos de crescimento e prognose melhoraram as estimativas.

**Palavras-chaves:** Produção. Mudanças Climáticas. Modelos Simétricos. Modelos não Lineares Generalizados. Redes Neurais Artificiais. Aprendizado de Máquinas.



# Abstract

The Gypsum Pole of Araripe is the main plaster producer in Brazil. To dehydrate the gypsum mineral that becomes plaster, the main energy source is firewood from the Caatinga, which has been exploited generally in an illegal manner. Nowadays the effects of deforestation in the Caatinga combined with climate change have made the supply of energy resources scarce. As a way of mitigating the effect of deforestation and reducing the suppression of Caatinga vegetation, experiments with fast-growing forests were implanted in the region. In order to provide a more adequate development of the management of these forests, it is necessary to adjust local statistical models that are more efficient and that take into account changes in climate variables. Thus, the objective of this research was to adjust nonlinear models and machine learning algorithms and to evaluate whether the inclusion of climatic variables generates equations with greater precision and generalization power in the estimation of production and volumetric growth of *Eucalyptus* in a semi-arid climate region in Northeast Brazil. The data were obtained through a research carried out at the IPA Experimental Station, municipality of Araripina - PE, with three clones of *Eucalyptus* spp. (C39, C41, C11) planted in spacing 3 m x 3 m, 4 m x 2 m, 3 m x 2 m, 2 m x 2 m and 2 m, x 1 m, in a multivariate experiment of repeated measures, implemented in 2010. A height and diameter measurements were taken every six months. At the age of 96 months, the volume was determined by the Smalian method. The rainy season database (2002 to 2009) was obtained from an experiment developed in the same region with the same silvicultural treatments. Nonlinear models and artificial intelligence algorithms were used for volume production, growth and prognosis. The adjustments were evaluated based on the corrected Schlaegel adjustment index ( $IA_c$ ), Furnival index (IF), standard error of percentage estimate ( $S_{xy}$  %), root mean square error (RMSE), mean absolute error (MAE) and graphical analysis of the residues. Volumetric production differed among clones and between different levels of density. Symmetric models and generalized nonlinear models produced equations with high  $AI_c$  values and low values of  $S_{xy}$  % with residuals meeting the assumptions of the regression. The artificial intelligence techniques have efficiently adapted the estimate of production and growth in volume, generating statistics of good quality. The inclusion of climatic variables in the growth and prognosis models improves the estimates.

**Key-words:** Production. Climate change. Simetric Models. Generalized Nonlinear Models. Artificial Neural Network. Machine Learning.

# Lista de Figuras

Figura 1 – Mapa do Brasil com destaque para as cinco cidades que compõem o Polo Gesseiro do Araripe-PE. . . . .	5
Figura 2 – Mapa do Nordeste com destaque para a área experimental localizada no Instituto Agrônômico de Pernambuco (IPA) no município de Araripina-PE. . . . .	18
Figura 3 – Destaque das posições dos diâmetros medidos na fase de cubagem rigorosa em experimento com clones de <i>Eucalyptus</i> no Polo Gesseiro do Araripe-PE. . . . .	20
Figura 4 – Taxa média da sobrevivência dos clones de <i>Eucalyptus</i> sob diferentes espaçamentos, cultivados no Polo Gesseiro do Araripe-PE. . . . .	24
Figura 5 – Produção volumétrica por espaçamentos para clones de <i>Eucalyptus</i> em cinco diferentes espaçamentos plantados no Polo Gesseiro do Araripe-PE. . . . .	26
Figura 6 – Produção volumétrica por clones de <i>Eucalyptus</i> em cinco diferentes espaçamentos plantados no Polo Gesseiro do Araripe-PE. . . . .	27
Figura 7 – Análise gráfica dos resíduos das equações obtidas para estimar a idade técnica de corte (ITC) dos clones de <i>Eucalyptus</i> , cultivados no Polo Gesseiro do Araripe-PE. . . . .	30
Figura 8 – Idade técnica de corte (ITC), definida pela interseção das curvas de incremento médio anual (IMA) e incremento corrente anual (ICA), para clones de <i>Eucalyptus</i> em cinco espaçamentos, cultivados no Polo Gesseiro do Araripe-PE. . . . .	31
Figura 9 – Mapa do Nordeste destacando o Polo Gesseiro do Araripe com ênfase na área experimental localizada no Instituto Agrônômico de Pernambuco, município de Araripina-PE. . . . .	42
Figura 10 – Distribuição dos resíduos para as equações geradas pelos modelos clássicos de regressão. . . . .	54
Figura 11 – Análise dos resíduos para as equações ajustadas por meio dos modelos não lineares com distribuição simétrica dos erros. . . . .	56
Figura 12 – Distribuição dos resíduos para as equações geradas pelos modelos não lineares generalizados. . . . .	58
Figura 13 – Análise dos resíduos das redes neurais artificiais treinadas para estimar a produção em volume dos clones de <i>Eucalyptus</i> . . . . .	60

Figura 14 – Análise dos resíduos para os modelos de Florestas Aleatórias - FA ( <i>Random Forest</i> ) para estimar a produção em volume dos clones de <i>Eucalyptus</i> . . . . .	63
Figura 15 – Análise dos resíduos para algoritmos de Máquinas de Vetores de Suporte - MVS para estimar a produção em volume dos clones de <i>Eucalyptus</i> . . . . .	65
Figura 16 – Localização geográfica das áreas experimentais e estação meteorológica do estudo conduzido no Instituto Agronômico de Pernambuco, Araripina-PE. . . . .	78
Figura 17 – Precipitação histórica por década na Estação Experimental do IPA no município de Araripina-PE. . . . .	85
Figura 18 – Análise de tendência para a precipitação histórica acumulada anual, mensal e temperatura média histórica mensal no município de Araripina-PE. . . . .	87
Figura 19 – Precipitação acumulada anual, com destaque para os períodos experimentais, registrada na Estação Experimental do IPA no município de Araripina-PE. . . . .	88
Figura 20 – Temperatura histórica, com destaque para os períodos experimentais, registrada na Estação Experimental do IPA no município de Araripina-PE. . . . .	89
Figura 21 – Histogramas para os dados de precipitação e temperatura nos experimentos. . . . .	90
Figura 22 – Análise dos resíduos para as equações de crescimento nos dois períodos experimentais. Experimento 1 (esquerda) e Experimento 2 (direita). . . . .	92
Figura 23 – Curvas de crescimento em volume para os dois experimentos avaliados.	92
Figura 24 – Produtividade de clones de <i>Eucalyptus</i> em dois períodos climáticos no Polo Gesseiro do Araripe-PE. . . . .	93
Figura 25 – Localização da área experimental com clones de <i>Eucalyptus</i> inserida no município de Araripina-PE. . . . .	102
Figura 26 – Importância de variáveis climáticas na estimativa do volume obtidas via redes neurais artificiais. . . . .	109
Figura 27 – Análise dos resíduos para as equações clássicas de crescimento em volume de clones de <i>Eucalyptus</i> . . . . .	111
Figura 28 – Análise dos resíduos para as equações com adição de variáveis climáticas para estimativa do crescimento em volume de clones de <i>Eucalyptus</i> . . . . .	113

Figura 29 – Análise dos resíduos para os algoritmos de aprendizagem de máquinas treinados para estimar o crescimento em volume de clones de <i>Eucalyptus</i> . . . . .	115
Figura 30 – Análise dos resíduos para os algoritmos de aprendizagem de máquinas com adição de variáveis climáticas treinados para estimar o crescimento em volume de clones de <i>Eucalyptus</i> . . . . .	117
Figura 31 – Curvas de crescimento real e estimada e distribuição dos resíduos para as equações de prognose do volume de clones de <i>Eucalyptus</i> . . . . .	128
Figura 32 – Curvas de crescimento real e estimada e distribuição dos resíduos para as equações de prognose do volume de clones de <i>Eucalyptus</i> com adição de variáveis climáticas. . . . .	131
Figura 33 – Análise dos resíduos para os algoritmos de aprendizagem de máquinas treinados para prognose do volume de clones de <i>Eucalyptus</i> . . . . .	133
Figura 34 – Análise dos resíduos para os algoritmos de aprendizagem de máquinas com adição de variáveis climáticas treinados prognose do volume de clones de <i>Eucalyptus</i> . . . . .	135

## Lista de tabelas

Tabela 1 – Descrição dos clones e espaçamentos utilizados em experimento realizado no Polo Gesseiro do Araripe-PE. . . . .	19
Tabela 2 – Análise de variância para a sobrevivência entre os clones de <i>Eucalyptus</i> spp. sob diferentes espaçamentos avaliados no Polo Gesseiro do Araripe. . . . .	23
Tabela 3 – Análise de variância para a produção em volume ( $\text{m}^3 \text{ha}^{-1}$ ) dos clones de <i>Eucalyptus</i> nos diferentes espaçamentos testados. . . . .	25
Tabela 4 – Produção volumétrica e o incremento médio anual (IMA) para os três clones de <i>Eucalyptus</i> distribuídos nos cinco espaçamentos, 8 anos após o plantio, no Polo Gesseiro do Araripe-PE. . . . .	28
Tabela 5 – Equações para estimativa da idade técnica de corte (ITC), por espaçamentos e clones, em experimento realizado no Polo Gesseiro do Araripe - PE. . . . .	29
Tabela 6 – Modelos não lineares para estimativa do volume de madeira ( $\text{m}^3 \text{árvore}^{-1}$ ) para clones de <i>Eucalyptus</i> plantados no Polo Gesseiro do Araripe-PE. . . . .	43
Tabela 7 – Estatísticas de qualidade e coeficientes das equações não lineares para estimar o volume ( $\text{m}^3 \text{árvore}^{-1}$ ) dos clones de <i>Eucalyptus</i> plantados no Polo Gesseiro do Araripe-PE. . . . .	52
Tabela 8 – Estatísticas de qualidade e coeficientes das equações ajustadas via modelos não lineares com distribuição simétrica dos resíduos, para estimativa do volume ( $\text{m}^3 \text{árvore}^{-1}$ ) dos clones de <i>Eucalyptus</i> plantados no Polo Gesseiro. . . . .	55
Tabela 9 – Coeficientes e estatísticas de teste das equações obtidas via modelos não lineares generalizados, para estimativa do volume por árvore ( $\text{m}^3 \text{árvore}^{-1}$ ) dos clones de <i>Eucalyptus</i> plantados na Chapada do Araripe-PE. . . . .	57
Tabela 10 – Hiperparâmetros de treinamento da RNA para estimar a produção dos clones de <i>Eucalyptus</i> no Polo Gesseiro. . . . .	59
Tabela 11 – Configurações e estatística de qualidade para a RNA do tipo Perceptron de Múltiplas Camadas treinada para estimar o volume individual por $\text{m}^3$ dos clones de <i>Eucalyptus</i> cultivados no Polo Gesseiro do Araripe. . . . .	60

Tabela 12 – Hiperparâmetros de treinamento para Florestas Aleatórias - FA para estimar o volume dos clones de <i>Eucalyptus</i> no Polo Gesseiro. . . . .	62
Tabela 13 – Configurações e estatística de qualidade para a técnica de Florestas Aleatórias - FA para estimar a produção dos clones de <i>Eucalyptus</i> no Polo Gesseiro. . . . .	62
Tabela 14 – Hiperparâmetros de treinamento para a técnica de Máquinas de Vetores de Suporte - MVS ( <i>Support Vector Machine</i> ) para estimar a produção dos clones de <i>Eucalyptus</i> no Polo Gesseiro. . . . .	64
Tabela 15 – Configurações e estatística de qualidade para a técnica de Máquinas de Vetores de Suporte - MVS ( <i>Support Vector Machine</i> ) para estimar a produção dos clones de <i>Eucalyptus</i> no Polo Gesseiro. . . . .	64
Tabela 16 – Resumo estatístico das melhores metodologias empregadas na estimativa da produção de clones de <i>Eucalyptus</i> no Polo Gesseiro do Araripe-PE. . . . .	67
Tabela 17 – Resultados dos testes de Pettitt, Mann-Kendall e da análise de regressão linear para variáveis do clima oriundas da estação experimental do Instituto Agrônomo de Pernambuco - IPA. . . . .	86
Tabela 18 – Análise estatística descritiva para as variáveis climáticas nos dois períodos de realização dos experimentos implantados no Polo Gesseiro do Araripe-PE. . . . .	89
Tabela 19 – Estatísticas para avaliar a qualidade das equações empregadas para comparação do crescimento nos dois períodos experimentais. . . . .	91
Tabela 20 – Modelos clássicos de crescimento para estimativa do volume em nível de povoamento. . . . .	103
Tabela 21 – Variáveis climáticas selecionadas pelo método <i>setwise</i> . . . . .	108
Tabela 22 – Variáveis selecionadas pelo valor de inflação da variância (VIF). . . . .	110
Tabela 23 – Estatísticas e coeficientes das equações para estimativa do crescimento em volume ( $m^3 ha^{-1}$ ) para os clones de <i>Eucalyptus</i> plantados no Polo Gesseiro do Araripe-PE. . . . .	110
Tabela 24 – Estatísticas e coeficientes das equações com adição de variáveis climáticas para estimativa do crescimento em volume ( $m^3 ha^{-1}$ ) para os clones de <i>Eucalyptus</i> plantados no Polo Gesseiro do Araripe-PE. . . . .	112
Tabela 25 – Resumo estatístico dos algoritmos de aprendizagem de máquinas para estimativa do crescimento em volume de clones de <i>Eucalyptus</i> spp. . . . .	114

Tabela 26 – Resumo estatístico dos algoritmos de aprendizagem de máquinas com adição de variáveis climáticas para estimativa do crescimento em volume de clones de <i>Eucalyptus</i> spp. . . . . .	116
Tabela 27 – Modelos de prognose ajustados para os clones de <i>Eucalyptus</i> plantados no Polo Gesseiro do Araripe-PE. . . . . .	125
Tabela 28 – Estatísticas e coeficientes das equações para prognose do volume ( $\text{m}^3 \text{ ha}^{-1}$ ) dos clones de <i>Eucalyptus</i> implantados no Polo Gesseiro do Araripe-PE. . . . . .	127
Tabela 29 – Estatísticas e coeficientes das equações com adição de variáveis climáticas para prognose do volume ( $\text{m}^3 \text{ ha}^{-1}$ ) para os clones de <i>Eucalyptus</i> plantados no Polo Gesseiro do Araripe-PE. . . . . .	130
Tabela 30 – Resumo estatístico dos algoritmos de aprendizagem de máquinas para prognose do volume de povoamentos de clones de <i>Eucalyptus</i> spp. . . . . .	132
Tabela 31 – Resumo estatístico dos algoritmos de aprendizagem de máquinas com adição de variáveis climáticas para prognose do volume de povoamentos de clones de <i>Eucalyptus</i> spp. . . . . .	134

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Objetivos</b>	<b>4</b>
2.1	Objetivo geral	4
2.2	Objetivos específicos	4
<b>3</b>	<b>Revisão de Literatura</b>	<b>5</b>
3.1	O Polo Gesseiro do Araripe	5
3.2	Florestas de rápido crescimento	6
3.3	Variáveis climáticas	7
3.3.1	Análise de tendência em séries climáticas	7
3.4	Modelos de regressão	9
3.5	Modelagem Florestal	10
3.6	Inteligência Artificial	11
3.6.1	Redes neurais artificiais (RNA)	12
3.6.2	Máquinas de vetores de suporte (MVS)	13
3.6.3	Florestas aleatórias (FA)	13
<b>4</b>	<b>Produção volumétrica de clones de <i>Eucalyptus</i> spp. sob diferentes espaçamentos em período de seca extrema no semiárido pernambucano</b>	<b>15</b>
4.1	Introdução	16
4.2	Material e Métodos	17
4.2.1	Descrição da área experimental	17
4.2.2	Planejamento do experimento	18
4.2.3	Sobrevivência	19
4.2.4	Variáveis dendrométricas e cubagem rigorosa	20
4.2.5	Idade técnica de corte	21
4.3	Resultados e Discussão	22
4.3.1	Análise da sobrevivência	22
4.3.2	Produção volumétrica	25
4.3.3	Incremento médio anual (IMA) e Idade técnica de corte (ITC)	28
4.4	Conclusões	33
	Referências Bibliográficas	35
<b>5</b>	<b>Modelos não lineares modificados e aprendizagem de máquinas para modelagem do volume de madeira em clones de <i>Eucalyptus</i></b>	<b>38</b>



5.1	Introdução . . . . .	39
5.2	Material e Métodos . . . . .	41
5.2.1	Descrição do experimento . . . . .	41
5.2.2	Modelos não lineares . . . . .	43
5.2.3	Algoritmos de aprendizagem de máquinas . . . . .	45
5.2.3.1	Redes neurais artificiais (RNA) . . . . .	45
5.2.3.2	Florestas aleatórias (FA) . . . . .	47
5.2.3.3	Máquinas de vetores de suporte (MVS) . . . . .	48
5.3	Estatísticas para avaliação dos modelos de regressão e de aprendizagem de máquinas . . . . .	49
5.4	Recursos computacionais para ajuste dos modelos . . . . .	51
5.5	Resultados e Discussão . . . . .	52
5.5.1	Modelos não lineares ajustados . . . . .	52
5.5.1.1	Modelos não lineares clássicos . . . . .	52
5.5.1.2	Modelos não lineares com erros simétricos . . . . .	55
5.5.1.3	Modelos não lineares generalizados . . . . .	57
5.5.2	Algoritmos de aprendizagem de máquinas . . . . .	59
5.5.2.1	Redes Neurais Artificiais . . . . .	59
5.5.2.2	Florestas Aleatórias . . . . .	61
5.5.2.3	Máquinas de Vetores de Suporte . . . . .	64
5.5.3	Comparação entre os modelos . . . . .	66
5.6	Conclusões . . . . .	68
	Referências Bibliográficas . . . . .	69
6	Análise do crescimento de clones de <i>Eucalyptus</i> spp. em período climático chuvoso e de seca extrema . . . . .	75
6.1	Introdução . . . . .	76
6.2	Material e Métodos . . . . .	78
6.2.1	Área experimental e planejamento do experimento . . . . .	78
6.2.1.1	Experimento 1 . . . . .	79
6.2.1.2	Experimento 2 . . . . .	79
6.2.2	Dendrometria e volume . . . . .	79
6.2.3	Dados climáticos . . . . .	80
6.2.4	Comparação da produtividade em dois períodos climáticos . . . . .	81
6.2.4.1	Análise de tendência e pontos de mudanças em séries climáticas . . . . .	81
6.2.4.2	Análise descritiva dos períodos climáticos e comparação da produtividade . . . . .	84

6.3	<b>Resultados e Discussão</b>	85
6.3.1	<b>Análise histórica das séries climáticas</b>	85
6.3.2	<b>Comparação das variáveis climáticas dos períodos experimentais</b>	88
6.3.3	<b>Crescimento em volume de clones <i>Eucalyptus</i> em condições climáticas distintas</b>	91
6.4	<b>Conclusão</b>	94
	<b>Referências Bibliográficas</b>	<b>95</b>
7	<b>Variáveis climáticas aplicadas em modelos não lineares e algoritmos de aprendizagem de máquinas para estimativa do crescimento volumétrico de clones de <i>Eucalyptus</i> spp.</b>	<b>98</b>
7.1	<b>Introdução</b>	100
7.2	<b>Material e Métodos</b>	101
7.2.1	<b>Descrição do experimento e variáveis dendrométricas</b>	101
7.2.2	<b>Variáveis climáticas</b>	102
7.2.3	<b>Modelos de crescimento clássicos</b>	103
7.2.4	<b>Estimativa do crescimento via algoritmos da aprendizagem de máquinas</b>	104
7.2.4.1	<b>Rede neural artificial</b>	104
7.2.4.2	<b>Floresta aleatória</b>	105
7.2.4.3	<b>Máquina de vetor de suporte</b>	106
7.2.5	<b>Seleção das variáveis climáticas</b>	107
7.2.6	<b>Avaliação das estimativas</b>	108
7.3	<b>Resultados e Discussão</b>	108
7.3.1	<b>Análise da importância das variáveis climáticas</b>	108
7.3.2	<b>Modelagem do crescimento clássica</b>	109
7.3.3	<b>Modelagem do crescimento com algoritmos de aprendizagem de máquinas</b>	114
7.4	<b>Conclusões</b>	117
	<b>Referências Bibliográficas</b>	<b>119</b>
8	<b>Modelos não lineares e algoritmos de aprendizagem de máquinas para prognose do volume de clones de <i>Eucalyptus</i> spp. com adição de variáveis climáticas</b>	<b>122</b>
8.1	<b>Introdução</b>	123
8.2	<b>Material e Métodos</b>	124
8.2.1	<b>Obtenção dos dados dendrométricos e climáticos</b>	124
8.2.2	<b>Modelos estatísticos para prognose</b>	125

8.2.3	Prognose do volume com algoritmos de aprendizagem de máquinas . . . . .	126
8.2.4	Avaliação das estimativas . . . . .	126
8.3	Resultados e Discussão . . . . .	127
8.3.1	Modelos estatísticos para prognose do volume . . . . .	127
8.3.2	Prognose do volume por meio de algoritmos de aprendizagem de máquinas . . . . .	132
8.4	Conclusão . . . . .	135
	Referências Bibliográficas . . . . .	136
9	Conclusão Geral . . . . .	139
	Referências Bibliográficas . . . . .	141

## CAPÍTULO 1

---

### Introdução

Os efeitos da mudança climática tem se intensificado nos últimos anos. O semiárido nordestino vem sofrendo com as mudanças no clima, principalmente com eventos extremos de seca. Com base em dados meteorológicos a década atual se configura umas das mais secas dos últimos 40 anos, com baixos índices de precipitação e aumento da temperatura mínima, média e máxima.

Estes efeitos, atrelado a exploração desordenada da vegetação Caatinga podem causar impactos ambientais irreparáveis na região semiárida. Inserido neste cenário, o Arranjo Produtivo Local do Polo Gesseiro do Araripe (denominado APL do Gesso) é o principal responsável por avanços no desmatamento em áreas de Caatinga ao seu entorno (ABREU, 2018), uma vez que a biomassa florestal é a fonte energética mais empregada pelas indústrias que beneficiam a gipsita para produção do gesso.

O APL do Gesso é responsável pela produção de 95% do gesso nacional. Com base em dados referentes ao ano de 2016, as cinco cidades que compõem o Polo Gesseiro do Araripe são responsáveis por quase 1,47% do PIB (Produto Interno Bruto) do setor industrial de Pernambuco e por 1,0% do PIB corrente do Estado (IBGE - Instituto Brasileiro de Geografia e Estatística, 2019). Com esta importância econômica e com todos os riscos ambientais envolvidos, é de fundamental importância desenvolver pesquisas que possam mitigar os efeitos da mudança climática e que proporcionem o desenvolvimento sustentável da região.

Desta forma, a implantação de florestas de rápido crescimento se torna uma opção para suprir a crescente demanda energética da região. Essas florestas podem ser implantadas em áreas onde a vegetação nativa já foi suprimida sem a necessidade de desmatar novas áreas. Florestas de clones de *Eucalyptus* foram implantadas na região e sua produtividade vem sendo estudada aos longo dos anos. Silva (2008/2009), Silva et al. (2013), Gadelha et al. (2015) e Gadelha, Silva e Ferreira (2018) desenvolveram pesquisas na região com o objetivo de descobrir quais os clones de *Eucalyptus* que apresentavam melhor adaptação às condições do local. Os autores citados indicam que o uso de *Eucalyptus* é uma opção economicamente e ambientalmente viável para a região.

Entretanto, com os efeitos da mudança climática incidindo sobre a ocorrência de eventos de seca extrema e elevando a temperatura média de diferentes biorregiões, a produtividade de florestas nativas e exóticas pode sofrer uma baixa significativa. Sendo assim, é importante desenvolver estudos que avaliem os possíveis efeitos das mudanças no clima sobre as florestas. Esses estudos envolvem desde a avaliação do material genético mais adequado, como o melhor arranjo espacial para plantação e as melhores técnicas de manejo e de prognose.

Para se realizar um planejamento florestal sustentado é necessário possuir eficiência em modelos de prognose. São nas técnicas de modelagem que se concentram a maior parte dos estudos de manejo, pois em geral, envolvem um grande número de variáveis populacionais, longos períodos de tempo, variações temporais e do mercado financeiro.

Os modelos empíricos são os mais aplicados no planejamento florestal, pois, precisam de um número pequeno de variáveis baseadas em mensurações e observações coletadas em campo (SUBRAMANIAN, 2016). No entanto, as pesquisas atuais se concentram em identificar os processos biológicos que mais influenciam o crescimento, o volume e a produção de uma floresta. Isto é possível com a adição de variáveis fisiológicas, gerando os modelos mecanísticos. Contudo, esse tipo de modelo se torna bastante complexo, pois, necessita de um grande número de variáveis de difícil obtenção (PENG et al., 2002). Com isto, surgem os modelos híbridos, que combinam modelos empíricos com variáveis climáticas. Assim, baseado na complexidade dos processos fisiológicos os modelos florestais podem ser classificados em fisiológicos simples (empíricos), fisiológicos complexos (mecanísticos) e híbridos empíricos-fisiológicos (JIN; HE; THOMPSON, 2016). O uso de modelos híbridos permite maior explicação da variabilidade encontrada no ambiente, aumentando a precisão da estimativa e o poder de generalização, logo serão os modelos mais utilizados no futuro (MAESTRI et al., 2013).

Além do desenvolvimento de novos modelos de regressão, o uso de algoritmos da aprendizagem de máquinas (Machine learning) vem ganhando bastante atenção na modelagem florestal. Segundo Chiarello et al. (2019), a partir de 2007 existe um aumento contínuo no uso de algoritmos da aprendizagem de máquinas aplicadas a problemas florestais. Ainda de acordo com os autores, do total de artigos com aprendizagem de máquinas 78% fazem uso de redes neurais artificiais (RNA). Entretanto, outros algoritmos como florestas aleatórias e máquinas de vetores de suporte, que possuem desempenho semelhante ou até superior, são pouco explorados.

Estudos que levam em consideração as variáveis climáticas combinadas com

algoritmos da aprendizagem de máquinas praticamente não são encontrados. O desenvolvimento destes modelos pode gerar decisões com grande valor de importância, aumentando o poder de generalização e precisão das estimativas.

## CAPÍTULO 2

---

### Objetivos

#### 2.1 Objetivo geral

Ajustar modelos não lineares e algoritmos da aprendizagem de máquinas e avaliar se a inclusão de variáveis climáticas gera equações com maior precisão e poder de generalização na estimativa da produção e do crescimento volumétrico de clones de *Eucalyptus* em região de clima semiárido do Nordeste brasileiro.

#### 2.2 Objetivos específicos

- Avaliar se a produção volumétrica de clones de *Eucalyptus* spp. é afetada por diferentes níveis de espaçamentos em povoamentos implantados em região semiárida na ocorrência de condições de seca extrema durante a rotação do plantio;
- Analisar se existe mudança nas séries climáticas da região do Araripe-PE e avaliar se afetam o crescimento dos clones de *Eucalyptus* em dois cenários climáticos;
- Ajustar modelos volumétricos não lineares simétricos, generalizados e avaliar se algoritmos da aprendizagem de máquinas, com destaque para as Redes Neurais Artificiais, Máquinas de Vetores de Suporte e Florestas Aleatórias são eficientes para predição do volume de clones de *Eucalyptus* spp.;
- Modelar o crescimento volumétrico em povoamentos de *Eucalyptus* spp. e avaliar a contribuição de variáveis climáticas ao desempenho dos modelos de crescimento clássicos e nos algoritmos de aprendizagem de máquinas;
- Avaliar o efeito da inclusão de variáveis climáticas na precisão de modelos de prognose clássicos e em algoritmos de aprendizagem de máquinas para prognose do volume.

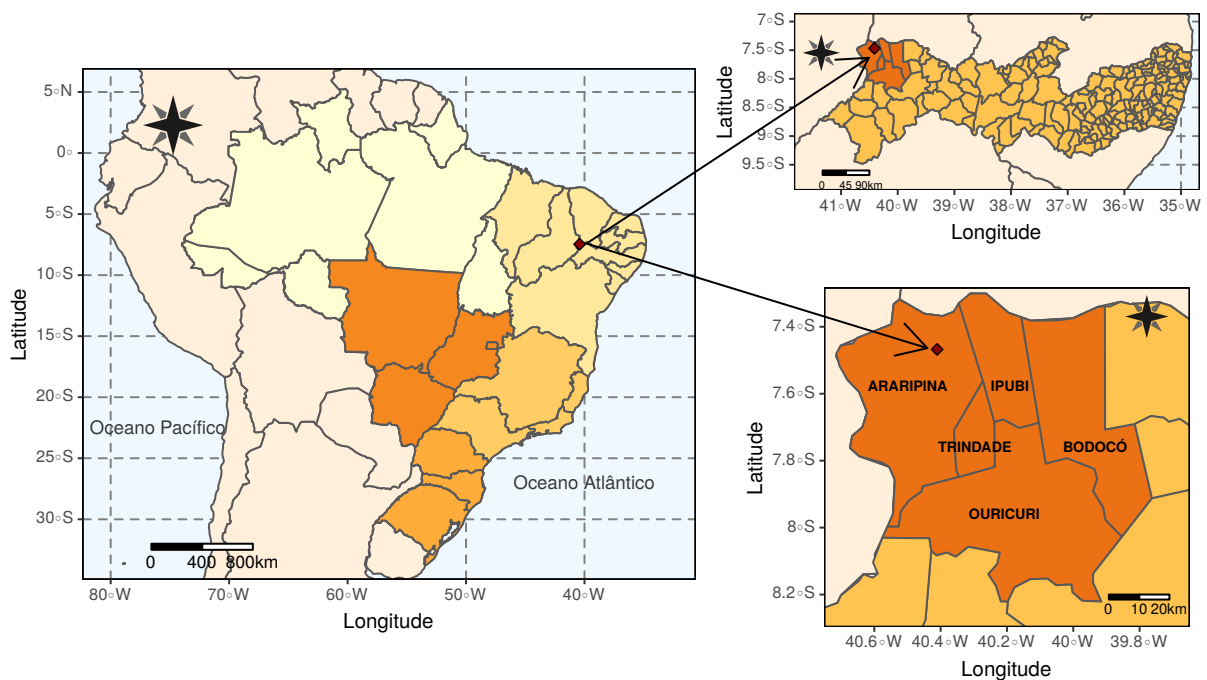
## CAPÍTULO 3

### Revisão de Literatura

#### 3.1 O Polo Gesseiro do Araripe

A região denominada de Polo Gesseiro do Araripe abrange as cidades de Ouricuri, Bodocó, Ipubi, Trindade e Araripina (Figura 1). Está localizado no semiárido do estado de Pernambuco, possui uma área de 11.792 Km<sup>2</sup> o que compreende 18% da área total do estado (SILVA et al., 2018). Estas cidades compõem o Arranjo Produtivo Local, conhecido como APL do gesso, juntas, são responsáveis por 95% da produção nacional de gesso (LINS, 2018). Segundo Abreu (2018), o Polo Gesseiro do Araripe possui 21 minas e 128 calcinadoras, gerando cerca de 45 mil empregos diretos e 45 mil indiretos.

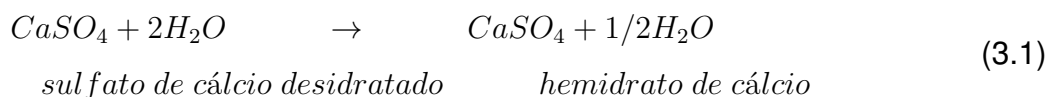
Figura 1 – Mapa do Brasil com destaque para as cinco cidades que compõem o Polo Gesseiro do Araripe-PE.



A extração do minério gipsita ( $\text{CaSO}_4 + 2\text{H}_2\text{O}$ ), matéria prima do gesso, ocorre na Bacia Sedimentar do Araripe desde a década de 60 (BRASIL, 2009). Segundo Campos



et al. (2018), a calcinação da gipsita para formação do gesso ocorre por meio das reações químicas dadas abaixo:



A forma mais comum para que esta desidratação ocorra é por aumento na temperatura, geralmente a gipsita é submetida a temperaturas entre 110 e 900 °C a depender do tipo de gesso que se pretende obter. A gipsita triturada é posta em um forno aquecido (entre 110 e 150 °C) e transita no seu interior até uma extremidade de saída, sendo este o processo de desidratação que gera o gesso do tipo Beta ( $\beta$ ), o mais utilizado na construção civil (CAMPOS et al., 2018).

A alta demanda por fontes energéticas para o processo da calcinação é o que gera os maiores problemas ambientais no Polo Gesseiro do Araripe. A principal fonte energética utilizada pelas indústrias do gesso é a lenha, o que gera grande pressão antrópica sobre os recursos florestais da região (FONTENELE et al., 2018). Gadelha, Silva e Ferreira (2018) estimaram que a demanda por lenha no Polo é de 652.680 m<sup>3</sup> por ano. Pesquisas realizadas por Riegelhaupt e Pareyn (2010) apontam que uma proporção de 80% da madeira consumida pelas indústrias do gesso vem de origem ilegal.

### 3.2 Florestas de rápido crescimento

As florestas plantadas para fins comerciais ocupam uma área de 264 milhões de hectares no globo terrestre. Deste total, 61% se encontram na China, Índia e Estados Unidos (SCOLFORO et al., 2016). No Brasil a área total de florestas plantadas é de 9,9 milhões de hectares, desses, 75,2% correspondem à florestas de eucaliptos (IBGE - Instituto Brasileiro de Geografia e Estatística, 2017). A madeira brasileira é empregada principalmente para produção de celulose, papel, painéis de madeira, pisos laminados, produtos sólidos e para fins energéticos (IBÁ, 2019). No ano de 2017 o setor de árvores plantadas com a finalidade da produção para a indústria de papel e celulose gerou um valor de R\$ 5,1 bilhões (IBGE - Instituto Brasileiro de Geografia e Estatística, 2017). Apesar de todo o desenvolvimento das florestas plantadas, no semiárido nordestino a alocação dos plantios florestais com espécies de rápido crescimento ainda é insipiente.

Com todos os avanços tecnológicos que ocorreram no desenvolvimento de clones mais resistentes e por as espécies de *Eucalyptus* já apresentarem características de fácil adaptação para diferentes condições ambientais (CASTRO et al., 2016), se

tornam um opção interessante para suprir a demanda energética do Polo Gesseiro. Segundo Gadelha, Silva e Ferreira (2018) e Fontenele et al. (2018) florestas de *Eucalyptus* são economicamente rentáveis e capazes de suprir a demanda por lenha das indústrias de gesso.

### 3.3 Variáveis climáticas

As mudanças na média das séries históricas de variáveis climáticas (em escala global ou regional), tais como, precipitação, temperatura, nebulosidade, velocidade do vento, entre outras, definem o termo *Mudança climática*. Alguns fatores, já bem conhecidos, que contribuem para aumentar o efeito dessa mudança são: o uso de combustíveis fósseis, industrialização, urbanização, mau uso da terra etc, todos estes fatores acabam afetando a composição atmosférica, o derretimento do gelo polar, poluição do ar, água e solo; e outras inúmeras consequências (BONAN; DONEY, 2018). Estas consequências, que também fazem parte do efeito estufa, causaram impactos drásticos no planeta obrigando as populações do globo a mudarem sua forma de viver (LOBELL; SCHLENKER; COSTA-ROBERTS, 2011).

Os avanços científicos nas formas de observações da mudança climática fornecem um melhor entendimento da variabilidade do sistema climático da Terra e sua provável resposta às influências humanas e naturais (MOSS et al., 2010). Ainda segundo os autores as implicações da mudança climática para o ambiente e sociedade também dependem de como a humanidade responde com mudanças (adaptações) na tecnologia, economia, política e estilo de vida. Diante destas inúmeras incertezas, para a sociedade e para o meio ambiente, faz-se necessário o desenvolvimento de estudos que busquem explicar o comportamento de florestas em diferentes condições climáticas, principalmente, para as regiões semiáridas, que sofrerão ainda mais com efeitos extremos do clima. Segundo Lacerda et al. (2014/2015) estudos indicam que para Pernambuco existe uma tendência de diminuição das chuvas, podendo acarretar modificações na composição e distribuição das florestas, na incidência de pragas e principalmente na produtividade (OLIVEIRA; BORROZZINO, 2018).

#### 3.3.1 Análise de tendência em séries climáticas

Segundo Gonçalves e Back (2018) o interesse em entender a mudança climática e o que ela pode causar ao ambiente natural reforça a necessidade de se desenvolver pesquisas em análise de tendência das séries temporais. Métodos estatísticos vem sendo largamente utilizados para detectar a existência de tendências em variáveis de

precipitação e temperatura (COSTA; FERREIRA; CORDEIRO, 2015). No geral esses testes se dividem em paramétricos e não paramétricos. Os testes não paramétricos possuem a vantagem de não possuir exigências em relação a amostra e são robustos a presença de *outliers*, já os paramétricos precisam que os dados da amostra sejam independentes e quase todos são definidos para uma distribuição normal (GONÇALVES; BACK, 2018).

Um dos testes mais utilizados para análise de tendências é o de Mann-Kendall. A principal justificativa para tal acontecimento se dá pela robustez do teste à falta de normalidade e a não estacionariedade da série (BLAIN, 2010). Este teste foi empregado por diversos autores em estudos para detectar tendências nas variáveis de precipitação (mensal e anual), temperatura (mínima, média e máxima) e incêndios naturais (COSTA; FERREIRA; CORDEIRO, 2015; RODRIGUES; JIMÉNEZ; RIVA, 2016; GONÇALVES; BACK, 2018). Outra metodologia bastante empregada para detecção de tendências é a análise de regressão. Back (2001) e Dumas (2013) trabalhando com séries de temperatura e precipitação fizeram uso desta metodologia para estimar a taxa de mudança da tendência.

Além do conhecimento geral da existência de tendência em uma série temporal ( $X$ ), é importante destacar em qual ponto de  $X$  essa mudança ocorreu. Desta forma, um ponto de mudança pode ser definido da seguinte forma: seja uma série temporal qualquer, expressa pelas variáveis aleatórias  $X_1, X_2, X_3, \dots, X_N$ , esta série pode apresentar um ponto de mudança em  $\lambda$  se a série  $X_n$  para  $n = 1, 2, 3, \dots, \lambda$  possuir uma função de distribuição  $F_1(x)$  e a série  $X_n$  para  $n = \lambda + 1, \lambda + 2, \lambda + 3, \dots, N$  possuir uma função de distribuição  $F_2(x)$  e  $F_1(x) \neq F_2(x)$  (PETTITT, 1979). Desta forma, pode-se concluir que  $\lambda$  é um ponto de mudança para a tendência da série.

Dentre os testes de detecção de pontos de mudanças citados na literatura científica, o teste de Pettitt é um dos que mais se destaca, sendo utilizado para diversos tipos de fenômenos temporais, tais como, detecção de mudanças em séries de precipitação, temperatura do ar, vazão hidrológica, incêndios florestais naturais e outros (PEZZATTI et al., 2013; ULIANA et al., 2015; SANTOS et al., 2016). Este é um teste não paramétrico, ou seja, robusto a falta de normalidade e uma das principais vantagens do método é sua robustez à presença de *outliers* que podem aparecer na série.

Outra metodologia que nos últimos anos ganhou bastante destaque, foi a descrita por Chen e Gupta (2000). A metodologia *change point* é uma técnica paramétrica capaz de detectar múltiplos pontos de mudanças em uma série temporal. Medicina, genética e economia são algumas das áreas em que Chen e Gupta (2012) destacam o

uso deste teste, outros autores já fizeram uso em cosmologia, meteorologia e ciências ambientais (JIMÉNEZ-RUANO; MIMBRERO; FERNÁNDEZ, 2017; TCHORBADJIEFF; ANGELOV, 2017).

### 3.4 Modelos de regressão

A análise de regressão é uma das técnicas mais empregadas da estatística (MONTGOMERY; PECK; VINING, 2013). Em diversos fenômenos ambientais seu uso é extremamente recorrente. A análise de regressão é empregada em, economia, física, ciências biológicas, engenharia, química, ciências sociais e manejo florestal (MONTGOMERY; PECK; VINING, 2013).

O desenvolvimento de modelos de regressão para uso no manejo florestal está relacionado com a identificação das características do crescimento e da produção melhor para entender qual o tempo de corte e exploração da área, o momento indicado de obter novas áreas para implantar novos plantios, etc (SILVA, 2016). Conhecer o estoque de madeira presente e futuro se torna uma tarefa menos complexa com o uso de modelos de crescimento e produção (SALES et al., 2015). Os modelos mais empregados no manejo florestal são os lineares e não lineares denominados de empíricos. Esses modelos consideram as relações entre as variáveis preditoras e a resposta de forma estatística não levando em consideração os processos biológicos e fisiológicos do sistema. No entanto, é importante desenvolver modelos mais robustos para gerar estimativas que possam ser mais generalizadas.

Considerando que dados florestais apresentam problemas quanto a homogeneidade da variância e autocorrelação dos resíduos (CALEGARIO et al., 2005), tem-se cada vez mais buscado metodologias estatísticas mais robustas (HESS et al., 2015).

Uma das formas que mais tem sido empregada na estatística é o uso dos modelos lineares generalizados, esses foram introduzidos por Nelder e Wedderburn (1972), abrindo um leque para o uso de outras distribuições da variável resposta, desde que a distribuição pertença à família exponencial. O uso destes modelos permitem uma maior adequação entre a relação funcional da média da variável preditiva com o preditor linear  $\eta$  (GUERRA, 2016). Os autores também desenvolveram um processo iterativo para estimar os parâmetros. Entretanto, a estrutura linear não representa bem o crescimento biológico das florestas, o que torna esses modelos pouco usuais nas ciências florestais. Os modelos lineares generalizados foram mais empregados em modelagem hipsométrica (CALEGARIO et al., 2005; HESS et al., 2015).

Para contornar a natureza não linear do crescimento florestal os modelos não

lineares generalizados podem ser empregados. Esses, são uma generalização da metodologia linear para casos não lineares. Esses modelos já foram reportados em ajustes para diversos problemas (TURNER; FIRTH, 2015; CURRIE, 2016).

Outra metodologia que tem sido empregada é os modelos com distribuição simétrica do erros. Esses são apresentados conforme Cysneiros, Paula e Galea (2005). Seja  $y$  uma variável aleatória definida nos  $\mathbb{R}$ , a qual possui um parâmetro de locação  $\mu$  e de escala  $\delta > 0$ , com a seguinte função densidade de probabilidade,

$$f(y; \mu, \delta) = \frac{1}{\sqrt{\delta}} \cdot g \left[ \frac{(y - \mu)^2}{\delta} \right], \quad (3.2)$$

para qualquer função  $g(\cdot)$  com as seguintes propriedades;  $g(w) > 0$ , para  $w > 0$  e  $\int_0^\infty w^{-\frac{1}{2}} g(w) dw = 1$ . Desta forma, tem-se que  $g(w)$  é uma função geradora de densidades e  $y$  é uma variável aleatória simétrica,  $y \sim S(\mu, \delta)$ .

Ainda segundo Cysneiros, Paula e Galea (2005) as principais distribuições que fazem parte da classe simétrica são: a normal, normal contaminada, logística tipos I, II e generalizada, exponencial potência, Kotz, Kotz generalizada, t-Student e t-Student generalizada.

Trabalhos com aplicação dessa metodologia já foram reportados em algumas pesquisas (GADELHA, 2014; FILHO et al., 2012; SALES et al., 2015).

### 3.5 Modelagem florestal incluindo variáveis climáticas

A seleção de modelos de crescimento adequados para estimar as variáveis de uma floresta é de fundamental importância para a tomada de decisão no setor florestal (BURKHART; TOMÉ, 2012). Nos últimos anos, muito se tem avançado nesta linha de pesquisa.

O crescimento de uma floresta se refere ao aumento das dimensões de uma ou de várias árvores em um período de tempo. A produção se refere a suas dimensões totais ao final de um período (BARRAZA, 2014). De acordo com Bague (2014) e Barrio-Anta et al. (2012), o crescimento e a produção de uma floresta podem ser estimados por: curvas de índice de sítio (Site Index), modelos de crescimento em diâmetro e área basimétrica, funções de altura (hipsométricas) e funções de crescimento do volume (volumétricas).

A modelagem de crescimento relacionada com variáveis climáticas é possível apenas quando existe uma forte adequação das mensurações realizadas em um inven-

tário florestal com as variáveis do clima, em relação ao espaço-tempo (SCOLFORO et al., 2016), ou seja, é necessário que a distância entre as estações meteorológicas utilizadas e as parcelas do inventário sejam pequenas.

Para incluir as variáveis climáticas nos modelos florestais existem diferentes formas. Segundo Maestri (2003), em modelos de índices de sítio se pode relacionar o par altura-idade com as variáveis climáticas, o que também pode ser feito para modelos de crescimento em altura dominante. A formulação desse modelo pode ser realizada de diferentes formas, a metodologia das diferenças algébricas generalizadas (Generalized Algebraic Difference Approach - GADA) proposta por Cieszewski e Bailey (2000) é uma das mais utilizadas. Essa metodologia é uma generalização do método das diferenças algébricas (Algebraic Difference Approach - ADA) proposto por Bailey e Clutter (1974). A principal vantagem do método GADA é que a equação de base pode ser expandida de acordo com várias teorias de crescimento, como a taxa de crescimento e assíntota, desta forma, mais de um parâmetro de cada modelo depende da qualidade do local, permitindo que as famílias de curvas obtidas sejam mais flexíveis (CIESZEWSKI, 2001; BURKHART; TOMÉ, 2012). Bem como, também é possível incluir essas variáveis no modelo de área basal (FERREIRA, 2009).

Em modelos de crescimento volumétrico a adição de variáveis climáticas pode ser realizada por meio de uma modificação no parâmetro assintótico (MAESTRI, 2003). A vantagem de incluir um fator modificador na assíntota é a preservação da estrutura não linear dos modelos (SCOLFORO et al., 2013). Diversos trabalhos foram reportados com a aplicação de modelos com adição de variáveis climáticas, sendo os resultados satisfatórios (ZAMIN et al., 2013; MACHADO et al., 2014; ELLI et al., 2017).

### **3.6 Modelagem florestal e aprendizagem de máquinas**

Algoritmos de aprendizagem de máquinas aplicados ao manejo florestal são cada vez mais frequentes, principalmente em estudos de modelagem. As Redes Neurais Artificiais (RNAs) se fazem presentes em pesquisas de diversas naturezas, tais como, estimativa do crescimento volumétrico, estimativa da produção volumétrica, estimativa da altura, etc (NIETO et al., 2012a; BINOTI et al., 2014; MAROHASY; ABBOT, 2015; DIAMANTOPOULOU et al., 2015; ÖZÇELIK et al., 2017). Recentemente, os algoritmos de Máquinas de Vetores de Suporte (Support Vector Machines) e Florestas Aleatórias (Random Forest) também estão sendo aplicados em problemas de modelagem florestal (SCHIKOWSKI et al., 2018; ABREU, 2019). Apesar de na área florestal as aplicações destes algoritmos estarem restritas ao que está implementado em *softwares*

prontos, tem-se verificado resultados semelhantes ou superiores aos da análise de regressão (ÖZÇELİK et al., 2017; SCHIKOWSKI et al., 2018).

### 3.6.1 Redes neurais artificiais (RNA)

Uma RNA é um modelo que está em uso crescente nos últimos anos e possui como ideia base o funcionamento do cérebro humano (VALENÇA, 2005). Em outras palavras, uma rede neural é um processador constituído por unidades de processamento interconectadas, chamado de neurônio, capazes de calcular funções, e estão distribuídos em camadas; camada de entrada, intermediárias e de saída, as quais estão conectadas entre si por pesos que possuem o poder de memorizar o conhecimento experimental (VALENÇA, 2010; BINOTI et al., 2015).

Segundo Valença (2010), a primeira proposta de um neurônio matemático com a finalidade de representar um neurônio biológico foi dada em 1943, por McCulloch e Pitts. Em 1962 Frank Rosenblatt propôs o modelo de rede neural mais simples, conhecido como *Perceptron*, o qual pode apresentar diferentes números de neurônios em uma camada de entrada ligados por pesos sinápticos, há, apenas, um neurônio na camada de saída. A principal característica desta rede é a capacidade de aprendizagem e generalização, mediante a apresentação de um exemplo. Posteriormente foi desenvolvida a rede *Adalaine*, com uma regra de aprendizagem baseada no método dos mínimos quadrados, que busca minimizar o erro médio quadrado por meio do gradiente descendente, assim, determinando o conjunto de pesos ótimos (HAYKIN, 2004).

Só com o surgimento da rede *Perceptron Multicamadas (Multilayer Perceptron - MLP)*, como uma generalização do *Perceptron*, por meio da adição de ao menos uma camada escondida, as RNAs passaram a ser capazes de aproximar qualquer função contínua (HAYKIN, 2004; VALENÇA, 2005). Este é o modelo de rede mais utilizado em sistemas de previsão de séries temporais e na estimação de variáveis, bem como, na maioria dos problemas de aproximações de funções (SOARES et al., 2011; BHERING et al., 2015).

Para o treinamento de uma rede MLP, pode-se utilizar diferentes algoritmos e funções de ativação. O algoritmo *backpropagation* foi o primeiro desenvolvido para treinar redes do tipo MLP, sendo uma generalização da técnica do gradiente descendente, utilizada para o treinamento da rede *Adalaine*. Algumas desvantagens para o algoritmo são, ficar preso a mínimos locais e tempo de convergência, mesmo assim, é o mais utilizado por está implementado na maioria dos *softwares* disponíveis (SURATGAR;

TAVAKOLI; HOSEINABADI, 2005).

### **3.6.2 Máquinas de vetores de suporte (MVS)**

Em paralelo com as RNA's, o método de Vetores de Suporte (Support Vector, SV) também constitui uma ferramenta de grande importância na inteligência artificial. Este método é parte integrante da aprendizagem de máquinas por meio das Máquinas de Vetores de Suporte (MVS, Support Vector Machines). As MVS's estão fundamentadas na teoria de aprendizagem estatística, propostas por Vapnik (1998).

Atualmente, a generalização das MVS's para casos de regressão vem ganhando destaque na aplicação em fenômenos ambientais, generalização essa que só foi possível pelo desenvolvimento do método de regressão de vetores de suporte (Support Vector Regression, SVR). Para prever novos valores, a SVR necessita apenas de um conjunto de dados de informação e de sua função de custo. De forma prática, esta metodologia busca a minimização do risco estrutural, tentando minimizar um limite superior do erro de generalização, diferentemente das RNA's que buscam a minimização do erro em cada exemplo apresentado (LORENA; CARVALHO, 2007).

Alguns pesquisadores fizeram uso desta ferramenta e constataram que é possível encontrar resultados superiores em comparação com a análise de regressão (NIETO et al., 2012b; GARCÍA-GUTIÉRREZ et al., 2015; GÖRGENS; MONTAGHI; RODRIGUEZ, 2015).

### **3.6.3 Florestas aleatórias (FA)**

Árvore de decisão é uma das técnicas de aprendizagem de máquinas mais utilizadas para inferências indutivas. Este método pode ser utilizado tanto para classificação como para regressão (JAMES et al., 2015). Na regressão o método de florestas aleatórias (Random forest) é mais robusto, pois utiliza blocos de árvores para construir modelos de predição mais poderosos (JAMES et al., 2015). As florestas aleatórias foram desenvolvidas por Breiman (2001) e, consistem em um método estatístico não paramétrico que usa uma abordagem baseada na amostragem do banco de dados para criar um conjunto de árvores de classificação ou regressão (GENUER; POGGI; TULEAU-MALOT, 2015).

Alguns dos pontos positivos desta metodologia são: possui alta precisão para generalizar dados de variáveis categóricas ou contínuas, é insensível a dados ruidosos e ao sobre ajustamento, é capaz de lidar com dados desbalanceados e o algoritmo



é capaz de identificar as variáveis mais importantes (ADAM et al., 2014; GENUER; POGGI; TULEAU-MALOT, 2015).

Modelos florestais já foram desenvolvidos com essa metodologia, para prever o volume de *Eucalyptus* com uso de variáveis de sensoriamento remoto (GÖRGENS; MONTAGHI; RODRIGUEZ, 2015; ABREU, 2019; REIS et al., 2019). No entanto, modelos criados a partir dos algoritmos de aprendizagem de máquinas com adição de variáveis climáticas são pouco desenvolvidos.

## CAPÍTULO 4

---

### **Produção volumétrica de clones de *Eucalyptus* spp. sob diferentes espaçamentos em período de seca extrema no semiárido pernambucano**

#### **Resumo**

Nesta pesquisa se observou a adaptação de clones de *Eucalyptus* spp. na Chapada do Araripe-PE. Esta região possui uma alta demanda por fontes energéticas alternativas devido ao Polo Gesseiro manter sua matriz energética basicamente da exploração da vegetação Caatinga. Sendo assim, como forma de aumentar os ganhos em produtividade volumétrica em plantios de clones de *Eucalyptus* spp. é importante entender quais níveis de espaçamentos proporcionam melhor competição entre as árvores. Desta forma, objetivou-se avaliar se a produção volumétrica de clones de *Eucalyptus* spp. é afetada por diferentes níveis de espaçamentos em povoamentos implantados em condições de seca extrema na Chapada do Araripe-PE. O experimento foi desenvolvido na Estação Experimental do Instituto Agronômico de Pernambuco (IPA) no município de Araripina-PE. Foram dispostos três clones de *Eucalyptus* (C11, C39 e C41) com cinco níveis de espaçamentos (2 m x 1 m, 2 m x 2 m, 3 m x 2 m, 3 m x 3 m e 4 m x 2 m) em delineamento inteiramente aleatório com arranjo fatorial (3 x 5) com 4 repetições. A taxa de sobrevivência do experimento foi superior a 94%. A maior produtividade em volume foi obtida com o clone C39 no arranjo espacial 2 m x 1 m. A condição de seca severa afetou negativamente a produtividade, mesmo assim, o clone C39 apresentou valores de IMA acima da média registrada para as espécies nativas da região.

**Palavras-chave:** Polo Gesseiro do Araripe-PE, volume de madeira, incremento médio anual, idade técnica de corte.

#### **Abstract**

In this research the adaptation of *Eucalyptus* clones in the Chapada of Araripe-PE was observed. This region has a high demand for alternative energy sources due to the Gypsum Pole maintaining its energy matrix basically from the exploitation of

Caatinga vegetation. Therefore, to increase the gains in volumetric productivity in planting *Eucalyptus* spp clones, it is important to understand which spacing levels provide the best competition among the trees. Thus, the objective was to evaluate the volumetric productivity of *Eucalyptus* spp. clones are affected by different spacing levels in stands implanted under severe weather conditions in Chapada of Araripe-PE. The experiment was carried out at the Experimental Station of the Pernambuco Agronomic Institute (IPA) in the municipality of Araripina-PE. Three *Eucalyptus* clones (C11, C39 and C41) with five spacing levels (2 m x 1 m, 2 m x 2 m, 3 m x 2 m, 3 m x 3 m and 4 m x 2 m) were arranged in a completely randomized design with factorial arrangement (3 x 5) with 4 repetitions. The survival rate of the experiment was over 94%. The highest volume productivity was obtained with clone C39 in the 2 m x 1 m spatial arrangement. The condition of severe drought negatively affected productivity, even so, clone C39 presented values of IMA above the average registered for the native species of the region.

**Key-words:** Gypsum Pole of Araripe-PE, wood volume, mean annual increment, technical rotation age.

#### 4.1 Introdução

Florestas plantadas de eucaliptos são bastante desenvolvidas nas regiões Sudeste e Centro-Oeste, correspondendo a mais de 60% da área total plantada no Brasil. Por outro lado, na região Nordeste os plantios se concentram basicamente no estado da Bahia, compreendendo apenas 11% da área total brasileira (IBÁ, 2019). As características edafoclimáticas do semiárido nordestino são um fator limitante para o desenvolvimento dessas florestas. Nos últimos anos, vários estudos foram realizados com o objetivo de melhorar a compreensão sobre a inserção do gênero *Eucalyptus* em regiões semiáridas com elevado estresse hídrico (DRUMOND; LIMA; SANTOS, 2003; SILVA, 2008/2009; GADELHA; SILVA; FERREIRA, 2018).

Os primeiros estudos realizados no semiárido nordestino concentraram esforços em identificar as espécies e/ou clones de *Eucalyptus* que apresentam adaptações aos fatores edafoclimáticos característicos da região (ALVES et al., 2007; FONTENELE et al., 2018), nos quais foram confirmadas a viabilidade de plantio para clones provenientes das espécies *E. tereticornis* e *E. urophylla*. Desta forma, os novos estudos na região são voltados para as práticas silviculturais envolvendo diferentes densidade populacionais.

A determinação do arranjo espacial do povoamento que proporciona maiores rendimentos volumétricos é um fator importante, pois a madeira proveniente de florestas

plantadas na região é empregada principalmente como fonte energética local (BARROS et al., 2010; SANTANA, 2016). Drumond et al. (2012) em estudo realizado na Chapada do Araripe, semiárido pernambucano, testaram cinco níveis de espaçamento (3 m x 2 m; 3 m x 2,5; 3 m x 3 m; 3 m x 3,5 m e 3 m x 4 m) e concluíram que o espaçamento 3 m x 2 m apresentou maior incremento em volume comparado aos demais níveis. Silva et al. (2013) recomendam testar diferentes arranjos espaciais para avaliar o desempenho produtivo de florestas de eucaliptos na região semiárida. Pouca atenção foi dada para espaçamentos com menor área disponível para o desenvolvimento da planta. O nível do espaçamento em povoamentos florestais acarretam diferenças na distribuição diamétrica, na altura e na taxa de sobrevivência (CARON et al., 2015). As taxas de incrementos médio e corrente do volume, bem como a idade técnica de corte do plantio também podem ser afetadas em casos de pequeno ciclo de rotação. Para a finalidade energética, menores níveis de espaçamentos podem proporcionar um aumento na produtividade em volume.

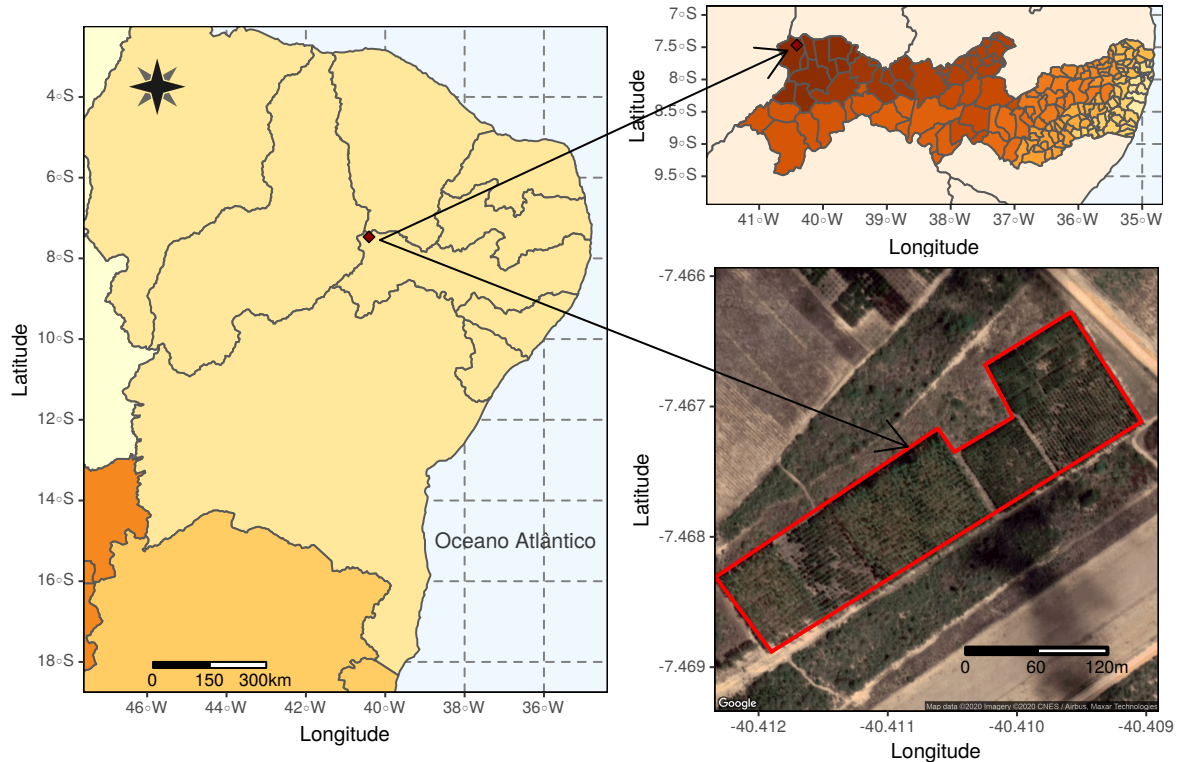
O grande volume de plantios de *Eucalyptus* estão concentrados no Sudeste, Centro-Oeste e litoral da Bahia (IBÁ, 2019) principalmente por estas regiões apresentarem precipitação anual acumulada superior a 1.000 mm. Porém, Alves et al. (2007), Drumond et al. (2012) e Silva et al. (2013) já evidenciaram a adaptação de clones de eucaliptos no semiárido pernambucano em período com precipitação média acumulada de 740 mm. No entanto, devido o semiárido nordestino apresentar ciclos de seca severa cabe investigar o desenvolvimento de clones do gênero nessas condições. Desta forma, objetivou-se avaliar se a produção volumétrica de clones de *Eucalyptus* spp. é afetada por diferentes níveis de espaçamentos em povoamentos implantados na região semiárida na ocorrência de condições de seca extrema durante a rotação do plantio.

## **4.2 Material e Métodos**

### **4.2.1 Descrição da área experimental**

O experimento foi instalado no mês de março de 2010 em uma área de 7,5 hectares na Estação Experimental do Instituto Agrônomo de Pernambuco (IPA) no município de Araripina (coordenadas 07° 29' 00" S e 40° 36' 00" W, altitude de 816 metros), inserida no semiárido pernambucano (Figura 2).

Figura 2 – Mapa do Nordeste com destaque para a área experimental localizada no Instituto Agronômico de Pernambuco (IPA) no município de Araripina-PE.



Segundo classificação de Köppen, o clima da região é do tipo BSh, semiárido quente, caracterizado por escassez de chuvas com bastante irregularidade em sua distribuição, a precipitação média anual é de 740 mm, apresenta índices de elevada evaporação e temperatura média do ar alta (25 °C). A vegetação característica desse tipo de clima é xerófila (Caatinga). O solo é do tipo LVAd1, Latossolos Vermelho - Amarelo Distróficos (SANTOS et al., 2011).

#### 4.2.2 Planejamento do experimento

A área experimental corresponde a uma área de agricultura abandonada em estágio de desertificação. A vegetação rasteira existente na área foi retirada. Com base na análise química do solo se fez a adubação e correção da acidez. Para melhorar a retenção de água no solo foi utilizado um polímero hidro retentor (GADELHA et al., 2015).

Para o plantio das mudas foram abertas covas com dimensões de 30 cm x 20 cm. O experimento foi composto por três clones (C11 - Clone de *Eucalyptus brassiana*, C39 - Clone de *Eucalyptus urophylla* e C41 - Clone de *Eucalyptus urophylla*) distribuídos em cinco diferentes espaçamentos (2 m x 1 m, 2 m x 2 m, 3 m x 2 m, 3 m x 3 m e 4

m x 2 m) e quatro repetições, disposto em um delineamento inteiramente casualizado com arranjo fatorial (3 x 5). Cada parcela foi composta por um total de 49 árvores, desprezando-se o efeito de bordadura, foram utilizadas 25 árvores. Os dados foram coletados em março de 2018, compreendendo um povoamento de 8 anos ou 96 meses de idade.

Tabela 1 – Descrição dos clones e espaçamentos utilizados em experimento realizado no Polo Gesseiro do Araripe-PE.

Clones	Espaçamentos	Área da Parcela (m <sup>2</sup> )	Área Útil (m <sup>2</sup> )	Plantas por Hectare
<b>C11 - Clone de <i>Eucalyptus brassiana</i></b>	2m x 1m	98	50	5000
	2m x 2m	196	100	2500
	3m x 2m	294	150	1667
	4m x 2m	392	200	1250
	3m x 3m	441	225	1111
<b>C39 - Clone de <i>Eucalyptus urophylla</i></b>	2m x 1m	98	50	5000
	2m x 2m	196	100	2500
	3m x 2m	294	150	1667
	4m x 2m	392	200	1250
	3m x 3m	441	225	1111
<b>C41 - Clone de <i>Eucalyptus urophylla</i></b>	2m x 1m	98	50	5000
	2m x 2m	196	100	2500
	3m x 2m	294	150	1667
	4m x 2m	392	200	1250
	3m x 3m	441	225	1111

Todos os clones foram obtidos por cruzamento natural.

Fonte: Adaptado de (GADELHA, 2014).

#### 4.2.3 Sobrevivência

A análise da sobrevivência foi realizada com base nas árvores da área útil de cada parcela na idade de 96 meses. A porcentagem de sobrevivência foi calculada por meio da proporção do número de plantas vivas em relação ao total de plantas. Para verificar diferenças nos tratamentos foi empregada a análise de variância (ANOVA). Constatada diferenças pela ANOVA foi empregado o teste de Tukey, conforme Hsu (1996), para comparação dos fatores (quando não houve diferenças significativas para a interação, em caso de significância para a interação se procedeu com os cálculos dos desdobramentos dos graus de liberdade), ao nível de 5% de significância.

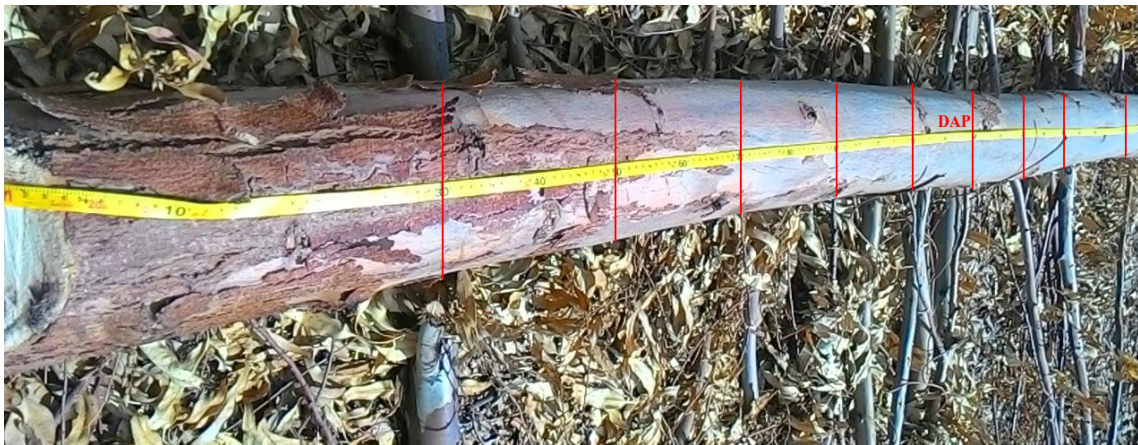
Todos os procedimentos das análises de variância realizados nesta pesquisa foram obtidos por meio do *Software R* (R Core Team, 2016), bem como, os cálculos

dos pós testes. Neste caso, para o teste de Tukey foi utilizado o pacote *agricolae* (MENDIBURU, 2017).

#### 4.2.4 Variáveis dendrométricas e cubagem rigorosa

Após a derrubada das árvores em campo, foram mensurados com uso de uma fita métrica: a altura total do fuste (Ht) e a altura comercial do fuste (Hf, considerando o fuste como a porção compreendida entre a altura de corte (15 cm) até a altura onde a árvore atinge um diâmetro mínimo de três centímetros). Utilizando uma suta foram tomadas medidas de diâmetro ao longo do fuste nos seguintes intervalos: 0,30 m; 0,50 m; 0,70 m; 0,90 m; 1,10 m; 1,30 m; 1,50 m; 1,70 m; 1,90 m; 2,30 m, a partir destes intervalos foram obtidas medidas a cada um metro (Figura 3).

Figura 3 – Destaque das posições dos diâmetros medidos na fase de cubagem rigorosa em experimento com clones de *Eucalyptus* no Polo Gesseiro do Araripe-PE.



Ainda com auxílio da suta foi medido o diâmetro a altura do peito (DAP, por convenção 1,30 m em relação ao solo). Com as medidas dos diâmetros seccionais foi possível determinar o volume total de madeira por árvore. Para tal, utilizou-se a fórmula de Smalian (SILVA; NETO, 1979), onde inicialmente foi calculado o volume para todas as seções do fuste;

$$v_i = \frac{g_i + g_{i+1}}{2} \cdot l_i \quad (4.1)$$

em que,

$v_i$  = é o volume da seção  $i$  ( $m^3$ );

$g_i$  = área basimétrica no início da seção  $i$  ( $m^2$ );

$g_{i+1}$  = área basimétrica no fim da seção  $i$  ( $m^2$ );

$l_i$  = comprimento da seção  $i$ .

$g_i$  é calculado por meio da seguinte fórmula;

$$g_i = \frac{\pi \cdot d_i^2}{4} \quad (4.2)$$

e  $d_i$  é o diâmetro da seção na posição que se pretende calcular.

Desta forma, o volume do fuste  $V_j$ , em (m<sup>3</sup>), é dado pelo somatório do volume de todas as seções;

$$V_j = \sum_{i=1}^n v_i \quad (4.3)$$

O volume por hectare foi obtido por meio da multiplicação do volume médio da parcela, pelo número de árvores em um hectare (de acordo com o espaçamento) e pela taxa de sobrevivência. Posteriormente, procedeu-se com a análise de variância (ANOVA). Existindo diferenças significativas nas análises, foi aplicado o teste de comparações múltiplas de Tukey (quando não houve diferenças significativas para a interação, em caso de significância para a interação se procedeu com os cálculos dos desdobramentos dos graus de liberdade), ao nível de 5% de significância.

#### 4.2.5 Idade técnica de corte

A idade de máxima produtividade volumétrica ou idade técnica de corte (ITC) é dada pela idade em que o incremento médio anual (IMA) é igual ao incremento corrente anual (ICA). O IMA pode ser obtido por meio da seguinte expressão:

$$IMA = \frac{V_i}{I} \quad (4.4)$$

Em que,  $I$  é a idade de interesse.

Os gráficos que apresentam as relações entre os incrementos médio e corrente, foram desenvolvidos considerando que a produção volumétrica pode ser expressa pelo seguinte modelo de Schumacher (CAMPOS; LEITE, 2017):

$$V_i = e^{(\beta_0 + \beta_1 \frac{1}{I})} + \epsilon_i \quad (4.5)$$



As letras gregas,  $\beta_0$  e  $\beta_1$  são parâmetros do modelo e  $\epsilon_i$  é o erro de estimativa que segue  $N(0, \sigma)$ . Desta forma, o ICA foi obtido por meio da derivada do modelo em função da idade I, como segue:

$$ICA = \frac{dV_i}{dI} = \frac{d e^{(\beta_0 + \beta_1 \frac{1}{I})}}{dI} = \left( \frac{-\beta_1}{I^2} \right) \cdot e^{(\beta_0 + \beta_1 \frac{1}{I})} \quad (4.6)$$

Considerando que a idade em que IMA = ICA é a idade de máxima produtividade, a ITC foi obtida por meio de:

$$\begin{aligned} IMA &= ICA \\ \frac{e^{(\beta_0 + \beta_1 \frac{1}{I})}}{I} &= \left( \frac{-\beta_1}{I^2} \right) \cdot e^{(\beta_0 + \beta_1 \frac{1}{I})} \\ I &= -\beta_1 \quad \therefore \quad ITC = -\beta_1. \end{aligned} \quad (4.7)$$

Para o desenvolvimento das curvas de incrementos será considerado o teste de Tukey. Não havendo diferenças significativas para a interação espaçamento x clone, as curvas serão desenvolvidas em grupos com base nas diferenças de cada fator. Logo, o clone C39 será renomeado de C1 e os clones C11 e C41 serão renomeados de C2, o espaçamento 2 m x 1 m receberá a nomenclatura E1, os espaçamentos 2 m x 2 m, 3 m x 2 m e 3 m x 3 m serão o E2 e o 4 m x 2 m será nomeado de E3. Desta forma, as curvas de incrementos serão desenvolvidas para cada novo tratamento, assim denominado: Tratamento E1C1, E1C2, E2C1, E2C2, E3C1 e E3C2.

A qualidade de ajuste das equações desenvolvidas foi avaliada por meio do Índice de Ajuste de Schlaegel corrigido ( $IA_c$ ) e erro padrão de estimativa em porcentagem  $S_{xy}\%$ , os resíduos foram submetidos ao teste de normalidade de Shapiro-Wilk (SW) e análise gráfica. Todos os procedimentos das análises de variância e o cálculo dos pós testes, bem como ajuste dos modelos desenvolvidos nesta pesquisa foram realizados por meio do *software* R (R Core Team, 2016).

### 4.3 Resultados e Discussão

#### 4.3.1 Análise da sobrevivência

Os resultados da análise de variância para a sobrevivência dos clones de *Eucalyptus* plantados no Polo Gesseiro do Araripe são apresentados na Tabela 2.

Considerando a taxa de sobrevivência para os efeitos clones e espaçamentos na idade de 96 meses não foi constatado diferenças significativas por meio da análise de

Tabela 2 – Análise de variância para a sobrevivência entre os clones de *Eucalyptus* spp. sob diferentes espaçamentos avaliados no Polo Gesseiro do Ararape.

Fonte de Variação	GL	Soma de Quadrados	Quadrado Médio	F
Clone	2	25,6	12,8	1,4257 <sup>ns</sup>
Espaçamento	4	9,6	2,4	0,2673 <sup>ns</sup>
Clone x Espaçamento	8	150,4	18,800	2,0941 <sup>ns</sup>
Resíduos	45	404,0	8,9778	
Total	59	589,6	CV = 3,06%	

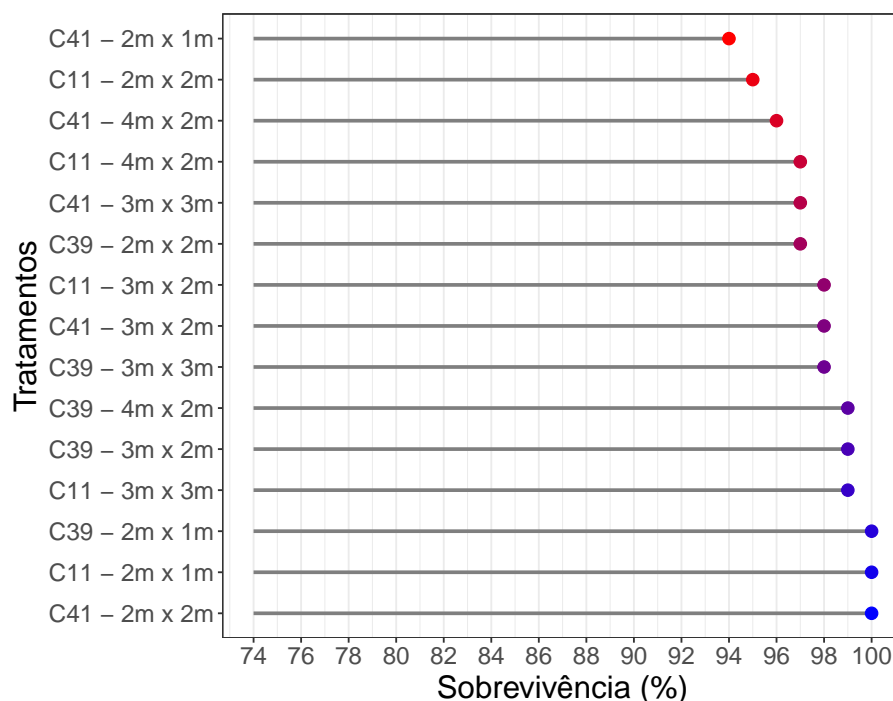
GL: Graus de liberdade; F: teste F; CV: Coeficiente de variação; <sup>ns</sup> Não significativo ao nível de 5% de significância.

variância, bem como, a interação desses efeitos não foi significativa ao nível de 5% de significância.

Os diferentes arranjos espaciais do povoamento podem acarretar em mudanças nos níveis de competição por recursos e espaços. Entretanto, no experimento não foi observado o aumento da competição em espaçamentos mais adensados, o que poderia diminuir as taxas de sobrevivência. O contrário foi observado por (SILVEIRA; REINER; SMANIOTTO, 2014), em que, espaçamentos menos adensados geraram menor competição por luz, água e nutrientes.

As taxas de sobrevivência por clones e espaçamentos são apresentadas na Figura 4. Os clones C39, C11 e C41 nos espaçamentos 2 m x 1 m e 2 m x 2 m mostraram as maiores taxas.

Figura 4 – Taxa média da sobrevivência dos clones de *Eucalyptus* sob diferentes espaçamentos, cultivados no Polo Gesseiro do Araripe-PE.



Magalhães et al. (2007) trabalhando com clones de *Eucalyptus*, com 9,4 anos de idade, nos espaçamentos 3 m x 2 m, 6 m x 2 m, 6 m x 3 m e 6 m x 4 m, na região Noroeste de Minas Gerais, verificaram a existência do aumento na taxa de sobrevivência em espaçamentos menos adensados (6 m x 4 m). A ausência de diferenças nas taxas de sobrevivência em relação aos níveis de espaçamento deste experimento podem está relacionadas com o elevado estresse hídrico durante o período experimental.

Segundo Rodrigues, Jiménez e Riva (2016), um dos fatores que podem elevar os valores da taxa de sobrevivência, principalmente em florestas monoespecíficas, está relacionado com o poder de adaptação da espécie ou do clone às condições edafoclimáticas da região estudada. As elevadas taxas de sobrevivência encontradas nesta pesquisa corroboram esta informação, pois os clones utilizados no experimento são os que apresentaram o melhor desempenho de produtividade, considerando, um total de 15 clones testados em experimento realizado no período de 2002 à 2009 na mesma região, conforme trabalhos realizados por Silva et al. (2013) e Gadelha, Silva e Ferreira (2018).

### 4.3.2 Produção volumétrica

A produção volumétrica e o IMA de um povoamento florestal são algumas das variáveis de maior interesse nos estudos com florestas energéticas e de alta produtividade. Diferenças significativas foram encontradas para os fatores clone e espaçamento, já a interação entre os fatores não se mostrou significativa (Tabela 3).

Tabela 3 – Análise de variância para a produção em volume ( $\text{m}^3 \text{ha}^{-1}$ ) dos clones de *Eucalyptus* nos diferentes espaçamentos testados.

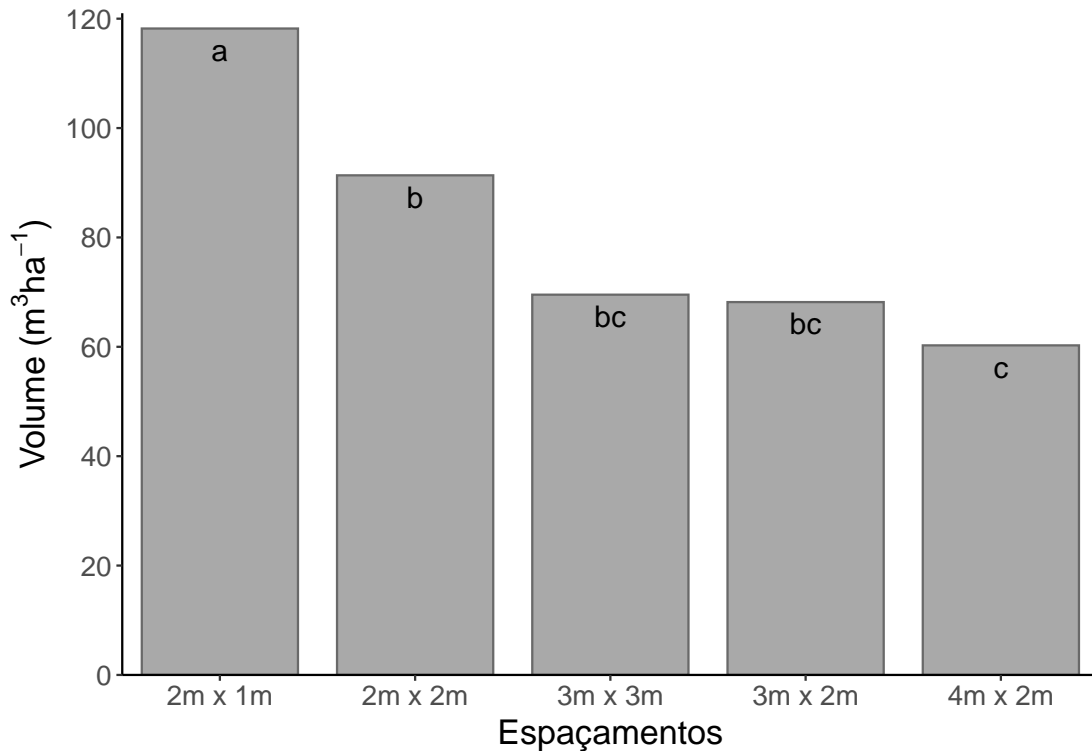
Fonte de Variação	GL	Soma de Quadrados	Quadrado Médio	F
Clone	2	4058,0	2028,8	4,5044**
Espaçamento	4	26590,0	6647,5	14,7592**
Clone x Espaçamento	8	4508,0	563,5	1,2510 <sup>ns</sup>
Resíduos	45	20268,0	450,4	
Total	59	55423,0	CV = 26,04%	

GL: Graus de liberdade; F: Teste F; CV: Coeficiente de variação; \*\*Significativo ao nível de 1% de significância; <sup>ns</sup> Não significativo ao nível de 5% de significância.

Os efeitos dos fatores são independentes, isto é, a produção dos clones independe do nível de espaçamento e vice-versa. Dessa forma, as comparações foram realizadas ao nível de cada fator. Para a produção em volume de madeira o espaçamento 2 m x 1 m apresentou valores médios superior, estatisticamente, aos demais arranjos espaciais (Figura 5). Esta mesma relação da produção nos diferentes níveis de espaçamentos foi encontrada por Gadelha (2014) aos 42 meses de idade, isto indica que mesmo o experimento possuindo variações na taxa de incremento por espaçamento e ao longo do tempo esses valores não foram suficientes para modificar a relação produção/adensamento.

Pode-se constatar que, para os espaçamentos mais adensados, a produção foi superior aos demais espaçamentos, ou seja, possivelmente existe uma relação positiva do adensamento populacional com aumento na produção da floresta. Esta mesma relação foi identificada por Gadelha (2014), trabalhando com o mesmo experimento na idade de 42 meses, bem como foi constatada por Sereghetti et al. (2015) em experimento com um clone de *E. urophylla* x *E. grandis* sob diferentes arranjos espaciais e também evidenciada por Silveira, Reiner e Smaniotta (2014) e Moulin et al. (2017), que realizaram pesquisas com clones de *Eucalyptus* em diferentes idades. Contrastando com os resultados obtidos por Gadelha (2014) aos 42 meses de idade fica evidente que, para o povoamento implantado no ambiente semiárido pernambucano, a relação adensamento/produção se mantém inalterada após a estabilização da floresta. Entre-

Figura 5 – Produção volumétrica por espaçamentos para clones de *Eucalyptus* em cinco diferentes espaçamentos plantados no Polo Gesseiro do Araripe-PE.



\*Médias seguidas pelas mesmas letras não diferem estatisticamente entre si, ao nível de 5% de significância, pelo teste de Tukey.

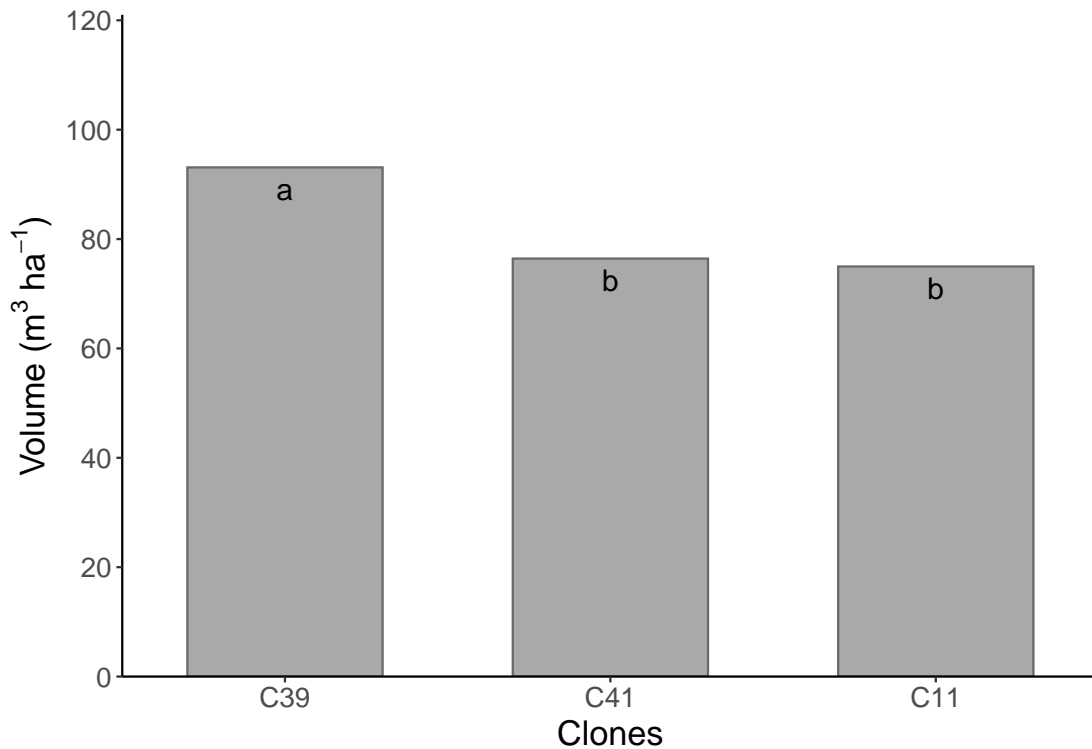
tanto, alguns autores destacam que esta relação pode ser alterada, pois plantas com maior área útil tendem a possuir uma reserva de nutrientes superior à florestas mais adensadas, principalmente quando o povoamento é mantido por períodos extensos (SILVEIRA; REINER; SMANIOTTO, 2014).

O principal fator que afeta a produção por hectare é o maior número de árvores por área plantada, já que o diâmetro é menor em áreas mais densas (DRUMOND et al., 2012; FERREIRA et al., 2016). Porém, é necessário destacar que apesar de plantios densos apresentarem maior volume de madeira por hectare, necessitam de custos maiores devido a imobilização de capital e para o combate de doenças e pragas, se comparado a povoamentos com arranjos espaciais menos densos (FILHO et al., 2018). O que acarreta perdas na viabilidade econômica do plantio.

Com relação a produção de madeira por clones, o C39 foi mais produtivo (Figura 6).

A produção destes clones ficou abaixo do encontrado em outros estudos. Em experimento realizado no período de 2002 à 2009 na mesma região, para um povoamento

Figura 6 – Produção volumétrica por clones de *Eucalyptus* em cinco diferentes espaçamentos plantados no Polo Gesseiro do Araripe-PE.



\*Médias seguidas pelas mesmas letras não diferem estatisticamente entre si ao nível de 5% de significância, pelo teste de Tukey.

mento de 7,5 anos de idade, Silva et al. (2013) determinaram a produção de 222,55 m<sup>3</sup> ha<sup>-1</sup> para o mesmo clone C39. Entretanto, Fontenele et al. (2018) para o sistema de talhadia, iniciado em 2009 e colhido em 2015, determinaram uma produção média de 72,99 m<sup>3</sup> ha<sup>-1</sup> para o clone C39, em ambos os casos o espaçamento utilizado foi o 3 m x 2 m. Estas diferenças de produção certamente foram causadas por fatores climáticos, pois durante o período de rotação a pluviosidade média foi de aproximadamente de 500 mm ano<sup>-1</sup>.

As médias da produção volumétrica e o incremento médio anual para cada clone por espaçamento são dadas na Tabela 4.

Tabela 4 – Produção volumétrica e o incremento médio anual (IMA) para os três clones de *Eucalyptus* distribuídos nos cinco espaçamentos, 8 anos após o plantio, no Polo Gesseiro do Araripe-PE.

Clones	Espaçamentos	Produção/Produtividade	
		Volume (m <sup>3</sup> ha <sup>-1</sup> )	IMA (m <sup>3</sup> ha <sup>-1</sup> ano <sup>-1</sup> )
C39	2 m x 1 m	127,32	15,92
C41	2 m x 1 m	114,63	14,33
C11	2 m x 1 m	112,66	14,08
C39	3 m x 2 m	100,11	12,51
C41	2 m x 2 m	96,38	12,05
C11	2 m x 2 m	89,74	11,22
C39	2 m x 2 m	87,95	10,99
C39	3 m x 3 m	83,11	10,39
C39	4 m x 2 m	67,04	8,38
C11	3 m x 3 m	66,46	8,31
C11	4 m x 2 m	59,35	7,42
C41	3 m x 3 m	59,03	7,38
C41	3 m x 2 m	57,71	7,21
C41	4 m x 2 m	54,40	6,80
C11	3 m x 2 m	46,72	5,84

#### 4.3.3 Incremento médio anual (IMA) e Idade técnica de corte (ITC)

Outra variável de fundamental importância para determinar a viabilidade econômica de uma floresta plantada é o incremento médio anual do povoamento. As médias de incremento dos clones por espaçamento são apresentadas na Tabela 4. O clone C39 no espaçamento 2 m x 1 m foi o tratamento com maior IMA, aproximadamente três vezes superior ao menos produtivo. De forma geral, estes valores estão bem abaixo do esperado para o plantio, porém são compreensíveis se for levado em consideração o efeito do período de seca extrema durante a realização da pesquisa.

Considerando o experimento realizado na mesma região, Silva et al. (2013) determinaram o IMA para os clones C39, C41 e C11, de 29,68, 22,85 e 15,65 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup>, respectivamente, cultivados em espaçamento 3 m x 2 m. Seguindo a mesma escala de incrementos Drumond et al. (2012) trabalhando na Chapada do Araripe-PE

com dois clones híbridos de *E. brassiana* x *E. urophylla* e *E. grandis* x *E. camaldulensis* nos arranjos espaciais, 3 m x 2 m, 3 m x 2,5 m, 3 m x 3 m, 3 m x 3,5 m e 3 m x 4 m, determinaram IMA para os respectivos clones de 24,7 e 20,8 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup>. Estes valores são consideravelmente superiores aos determinados nesta pesquisa. Todavia, Fontenele et al. (2018) registraram baixos valores de IMA, , 13,03 a 2,58 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup> para 15 clones de *Eucalyptus* no regime de talhadia, plantados na mesma região em período de seca severa. Assim, fica evidente que anomalias climáticas afetam severamente a produtividade de florestas plantadas na região semiárida.

Para que o povoamento florestal seja mais rentável se faz necessário determinar a idade ótima de colheita ou idade técnica de corte (ITC), esta foi determinada para cada tratamento conforme nomenclatura discutida na metodologia. As equações determinadas e as estatísticas dos ajustes são apresentadas na Tabela 5.

Tabela 5 – Equações para estimativa da idade técnica de corte (ITC), por espaçamentos e clones, em experimento realizado no Polo Gesseiro do Araripe - PE.

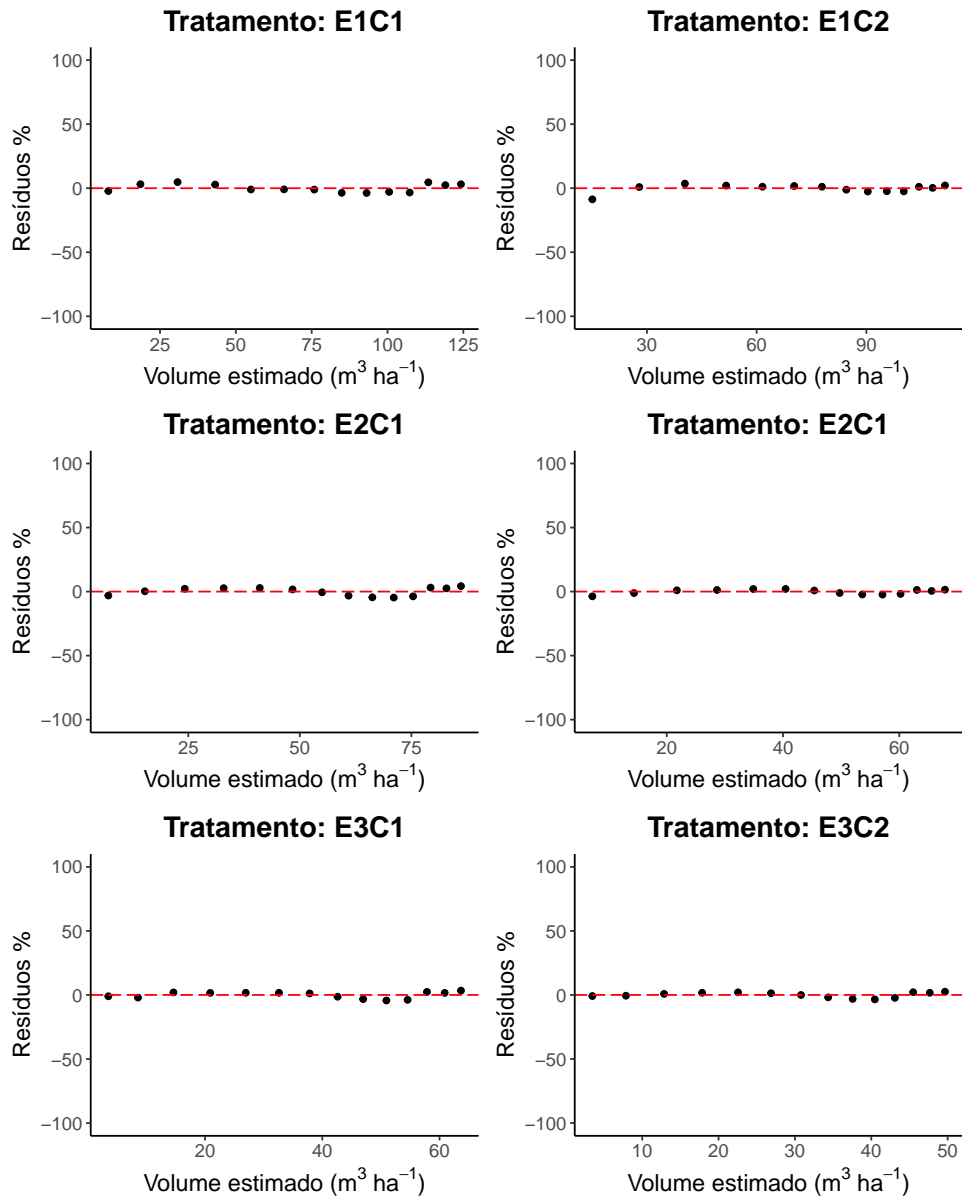
Tratamento	Equações	IA <sub>c</sub>	S <sub>xy</sub> (%)	SW (Valor p)	ITC (meses)
E1C1	$\hat{V} = e^{(5,4561-60,8717 \cdot \frac{1}{T})}$	0,9930	4,4447	0,06	60,87
E1C2	$\hat{V} = e^{(5,1755-44,3153 \cdot \frac{1}{T})}$	0,9904	4,3490	0,02	44,31
E2C1	$\hat{V} = e^{(5,0327-55,3847 \cdot \frac{1}{T})}$	0,9850	6,2435	0,06	55,38
E2C2	$\hat{V} = e^{(4,7348-49,6731 \cdot \frac{1}{T})}$	0,9912	4,5102	0,12	49,67
E3C1	$\hat{V} = e^{(4,8220-64,1821 \cdot \frac{1}{T})}$	0,9839	7,1229	0,07	64,18
E3C2	$\hat{V} = e^{(4,5191-58,9460 \cdot \frac{1}{T})}$	0,9819	7,1245	0,17	58,94

Todos os coeficientes foram significativos ao nível de 1% de probabilidade de erro pelo teste T. SW = Teste de normalidade dos resíduos de Shapiro-Wilk.

Considerando as condições de ajustes dos modelos, as estatísticas de teste foram satisfatórias, com elevados valores de índice de ajuste de Schlaegel corrigido (IA<sub>c</sub>) e baixo erro padrão da estimativa (S<sub>xy</sub>%), o que demonstra a boa precisão das equações. Em relação a distribuição dos resíduos todas as equações apresentaram desvios tendendo a normalidade, segundo o teste de Shapiro - Wilk com 5% de probabilidade de erro. A análise gráfica dos resíduos para as equações obtidas é evidenciada na Figura 7.



Figura 7 – Análise gráfica dos resíduos das equações obtidas para estimar a idade técnica de corte (ITC) dos clones de *Eucalyptus*, cultivados no Polo Gesseiro do Araripe-PE.

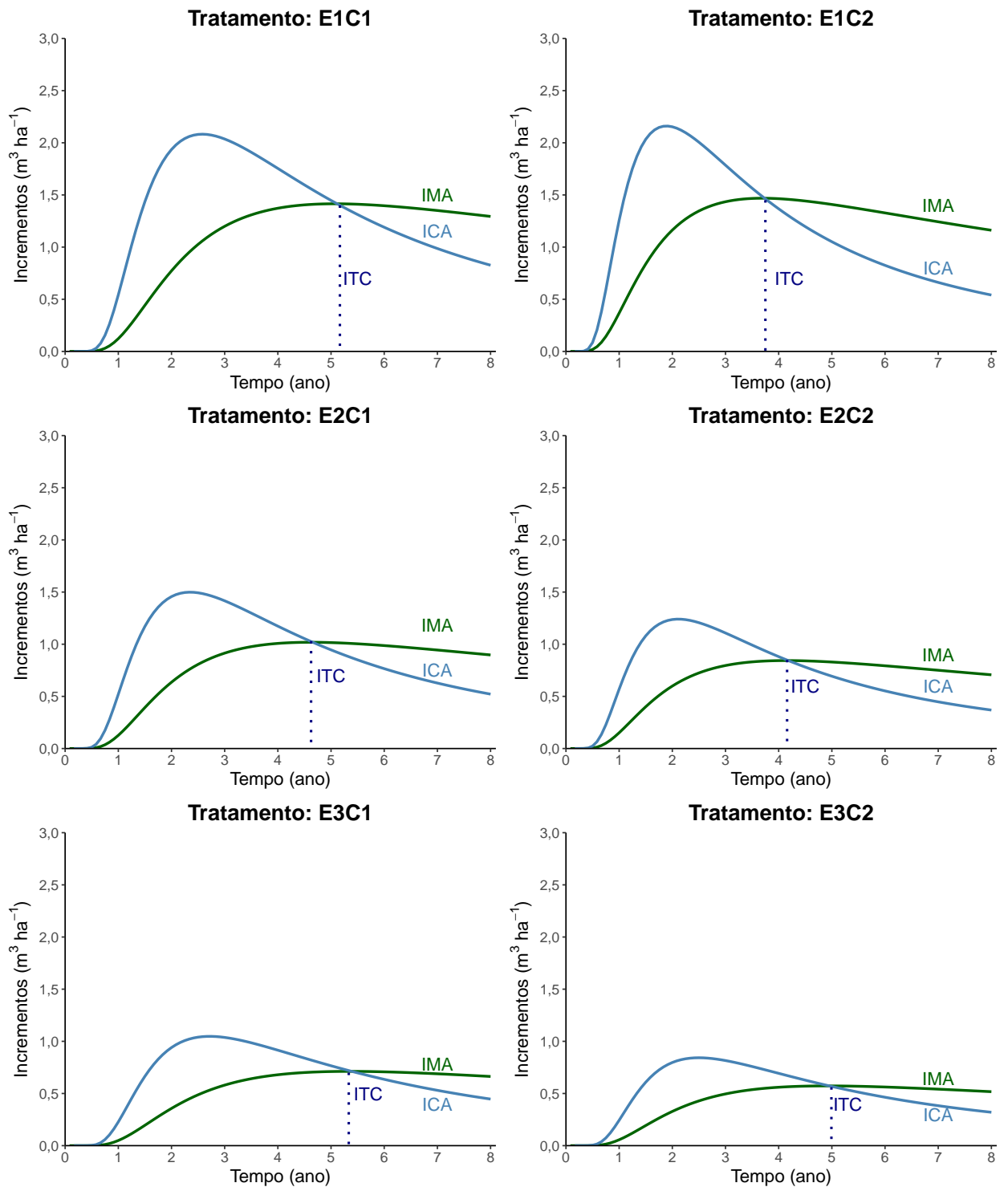


Fontenele et al. (2018) também encontraram bons ajustes para o modelo de Schumacher trabalhando com clones de *Eucalyptus* em segunda rotação no Polo Gesseiro. Segundo Guera et al. (2019a), o modelo de Schumacher permitiu melhores ajustes do crescimento para determinação da ITC em relação aos modelos de Chapman-Richards e Silva-Bailey para plantações de *Pinus caribaea* var. *caribaea*, o que justifica o emprego deste modelo para estimativa da ITC.

Considerando as curvas de incremento médio anual (IMA) e incremento corrente anual (ICA) é possível destacar que o ponto de incremento máximo ocorreu entre 1,5 e

2,5 anos de idade para os seis tratamentos (Figura 8).

Figura 8 – Idade técnica de corte (ITC), definida pela interseção das curvas de incremento médio anual (IMA) e incremento corrente anual (ICA), para clones de *Eucalyptus* em cinco espaçamentos, cultivados no Polo Gesseiro do Araripe-PE.



O clone C39 (C1) o que apresentou os maiores pontos de máximo. Esse período coincide com o intervalo em que o experimento recebeu o maior volume de precipitação, embora muito baixo quando comparado com outros anos.

A ITC variou de 3,7 a 5,3 anos. O clone C39 (C1) apresentou a maior ITC em comparação com os demais, o mesmo ocorreu para o tratamento E3 que corresponde ao espaçamento 4 m x 2 m.

As ITC's determinadas nesse experimento foram inferiores às encontradas por Silva et al. (2013), que registraram idades de 6,5 à 7 anos, considerando um experimento com 15 clones de *Eucalyptus* na mesma região. A principal razão para estas diferenças são os fatores climáticos extremos que ocorreram no período experimental, outro fator que pode ter influenciado, são os diferentes arranjos espaciais instalados. Fontenele et al. (2018) também registraram baixos valores de ITC, 5 anos para os clones C39, C41 e C11, em experimento desenvolvido de 2009 à 2015 na mesma região em regime de talhadia. Estes autores também atribuem os baixos valores da ITC aos fatores climáticos.

Diante do exposto é importante discutir a produtividade de clones de *Eucalyptus* em relação às espécies nativas e exóticas da Caatinga, vegetação predominante na região. A comparação entre os IMA's é um bom ponto de partida para esta análise. Barros et al. (2010) trabalhando na área do Polo Gesseiro do Araripe-PE instalaram parcelas com espécies nativas e exóticas para efeito de comparação com os clones de *Eucalyptus*. O experimento foi conduzido no período de 2002 a 2008, configurando um povoamento de 6,5 anos de idade. Os incrementos determinados para as espécies, Sabiá (*Mimosa caesalpinifolia*), Jurema (*Mimosa tenuiflora*), Angico (*Anadenanthera columbrina*), Ipê 1 (*Tabebuia sp. 1*) e Algaroba (*Prosopis juliflora*) foram de 8,04, 6,58, 5,64, 1,28 e 0,24 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup>, respectivamente. Vale ressaltar que esse experimento foi conduzido em uma época com condições climáticas consideradas normais para a região, mesmo assim, os IMA's apresentaram grande variação, isto indica que as espécies possuem diferentes pontos de máximo incremento. Considerando o clone mais produtivo neste experimento (C39 no espaçamento 2 m x 1 m) o IMA foi de 16,5943 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup>, duas vezes acima ao encontrado para o Sabiá.

Comparando com a Caatinga em Planos de Manejo Florestal (PMF) com seu IMA máximo de 5 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup>, com uma rotação técnica de 15 anos, comumente usada na região (MEUNIER; FERREIRA; SILVA, 2018), este tempo implica em praticamente duas rotações com eucaliptos. CARVALHO et al. (2020) analisando PMF no Agreste do estado do Rio Grande do Norte encontraram IMA de aproximadamente 3 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup>. Já, Riegelhaupt, Pareyn e Bacalini (2010) trabalhando em sete áreas de Caatinga no

Rio Grande do Norte e Ceará com PMF's em diferentes idades para o ciclo de corte (período de corte variando entre 5 e 20 anos), encontraram valores de IMA de 0,3 até 11 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup>. É importante destacar que os autores encontraram valores IMA acima de 5 m<sup>3</sup> ha<sup>-1</sup> ano<sup>-1</sup> apenas em duas fazendas localizadas a menos de 35 km do litoral do estado do Ceará, com precipitação acumulada média anual acima de 1200 mm, o que não reflete as condições da região semiárida. Isto resulta que aproximadamente cada hectare plantado com eucaliptos equivale a aproximadamente entre cinco e seis hectares de Caatinga. Se levar em consideração o rendimento nos fornos de calcinação da gipsita em que para uma tonelada de gesso produzida se usa 0,151m<sup>3</sup> de madeira de Caatinga e 0,100 m<sup>3</sup> de madeira de clones de eucaliptos (FONTENELE et al., 2018), esta relação passa a ser acima de 8 vezes, implicando que além de ser uma fonte de energia viável para indústria do gesso, um fator muito importante pode ser considerado que é a possibilidade de se diminuir a devastação da vegetação da Caatinga pelo oferecimento de madeira de eucaliptos em curto espaço de tempo. Segundo Gadelha, Silva e Ferreira (2018) aos 42 meses de idade os três clones testados no experimento apresentavam viabilidade econômica. Sendo que para alguns espaçamentos os autores destacam que o valor presente líquido só foi positivo a uma taxa de juros de 4% a.a. ou em idades superiores aos 42 meses, entretanto a maior parte dos espaçamentos possuem valor presente líquido positivo a taxas de juros de até 8% a.a., apesar destas considerações se pode destacar a viabilidade econômica do experimento em condições de seca extrema.

Mesmo em condições de seca severa e atípica, levando em consideração que foram seis anos com chuvas abaixo da normal climatológica para a região semiárida, os experimentos com clones de *Eucalyptus* conduzidos na Chapada do Araripe - PE evidenciam a possibilidade de cultivo, com rendimentos volumétricos acima da média apresentada por espécies nativas.

#### 4.4 Conclusões

- Os clones testados estão aptos à região semiárida, mesmo em diferentes arranjos espaciais, apresentando taxas de sobrevivências acima de 94%.
- O clone C39 é o mais produtivo em termos volumétricos.
- O fator nível de espaçamento afeta mais fortemente a produtividade. Para os três clones a maior produtividade foi obtida com o espaçamento 2 m x 1 m.

- A idade técnica de corte (ITC) é de aproximadamente 5 anos para o clone mais produtivo independente do nível de adensamento.
- A condição de seca severa afetou negativamente a produtividade, mesmo assim, o clone C39 apresentou valores de IMA acima da média registrada para as espécies nativas da região.

## Referências Bibliográficas

- ALVES, A. M. C. et al. Quantificação da produção de biomassa em clones de eucaliptos com 4, 5 anos, no Polo Gesseiro do Araripe-Pe. **Revista de Ciências Agrárias Amazonian Journal of Agricultural and Environmental Sciences**, v. 48, n. 1, p. 161–174, 2007.
- BARROS, B. C. et al. Volumetria e sobrevivência de espécies nativas e exóticas no Polo Gesseiro do Araripe, PE. **Ciência Florestal**, v. 20, n. 4, p. 641–647, 2010.
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 5. ed. Viçosa, MG: Editora UFV, 2017. 636 p.
- CARVALHO, A. C. et al. Produção de energia da madeira de espécies da Caatinga aliada ao manejo florestal sustentável. v. 48, n. 126, p. 1–13, 2020.
- CARON, B. O. et al. Quantificação da biomassa florestal em plantios de curta rotação com diferentes espaçamentos/quantification of forest biomass in short rotation plantations with different spacings. **Comunicata Scientiae**, v. 6, n. 1, p. 106–112, 2015.
- DRUMOND, M.; LIMA, P.; SANTOS, R. A. V. Comportamento de algumas espécies/procedências de *Eucalyptus* no município de Lagoa Grande-PE. **Brasil Florestal**, n. 78, p. 75–80, 2003.
- DRUMOND, M. A. et al. Performance of two hybrid clones of *Eucalyptus* planted under five spacings in the Araripe plateau, Pernambuco, Brazil. **Acta Horticulturae**, n. 959, p. 167–171, 2012.
- FERREIRA, A. D. et al. Influence of spatial arrangements on silvicultural characteristics of three *Eucalyptus* clones at integrated crop-livestock-forest system. **African Journal of Agricultural Research**, v. 11, n. 19, p. 1734–1742, 2016.
- FILHO, J. I. P. et al. Avaliação econômica de projetos de florestamento com eucalyptus grandis hill ex maiden implantado sob diferentes espaçamentos. v. 28, n. 1, p. 240–248, 2018.
- FONTENELE, N. d. M. et al. Volumetric and economic evaluation of *Eucalyptus* spp. clones in high forest, coppice and native vegetation systems at the Gypsum Pole of Araripe-PE. **Scientia Forestalis**, v. 46, n. 117, p. 77–86, 2018.

- GADELHA, F. H. L. **Desempenho silvicultural e avaliação econômica de clones híbridos de eucaliptos plantados em diferentes regimes de manejo para fins energéticos**. 147 p. Tese (Ciências Florestais) — Universidade Federal Rural de Pernambuco, Departamento de Ciências Florestais, Recife, 2014.
- GADELHA, F. H. L.; SILVA, J. A. A.; FERREIRA, R. L. C. Viabilidade econômica do cultivo de clones de *Eucalyptus* em função do espaçamento e do sistema silvicultural, em Araripina-PE. **Anais da Academia Pernambucana de Ciência Agrônoma**, v. 15, n. 2, p. 105–123, 2018.
- GADELHA, F. H. L. et al. Produtividade de clones de eucaliptos em diferentes sistemas de manejo para fins energéticos. **Pesquisa Florestal Brasileira**, v. 35, n. 83, p. 263–270, 2015.
- GUERA, O. G. M. et al. Modelos matemáticos para la determinación del turno óptimo en plantaciones forestales. **Madera y Bosques**, v. 25, n. 1, p. 1–19, 2019.
- HSU, J. **Multiple comparisons: theory and methods**. [S.I.]: Chapman and Hall/CRC, 1996. 234 p.
- IBÁ. **Indústria brasbrasileira de árvores**. São Paulo, 2019. 80 p.
- MAGALHÃES, W. M. et al. Desempenho silvicultural de clones e espécies/procedências de *Eucalyptus* na região noroeste de Minas Gerais. **Cerne**, v. 13, n. 4, 2007.
- MENDIBURU, F. d. **agricolae: Statistical Procedures for Agricultural Research**. [S.I.], 2017. R package version 1.2-8. Disponível em: <<https://CRAN.R-project.org/package=agricolae>>.
- MEUNIER, I. M. J.; FERREIRA, R. L. C.; SILVA, J. A. A. O licenciamento de Planos de Manejo Florestal da Caatinga assegura sua sustentabilidade? **Pesquisa Florestal Brasileira**, v. 38, p. 1–7, 2018.
- MOULIN, J. C. et al. Efeito do espaçamento, idade e irrigação no volume e densidade básica do eucalipto. **Floresta e Ambiente**, v. 24, p. 1–10, 2017.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Disponível em: <<https://www.R-project.org/>>.
- RIEGELHAUPT, E.; PAREYN, F. G. C.; BACALINI, P. O manejo florestal na caatinga: resultados da experimentação. In: GARIGLIO, M. A. et al. (Ed.). **Uso sustentável e conservação dos recursos florestais da Caatinga**. Brasília, DF: Ministério do Meio Ambiente, Serviço Florestal Brasileiro, 2010. p. 256–275.

- RODRIGUES, M.; JIMÉNEZ, A.; RIVA, J. d. I. Analysis of recent spatial–temporal evolution of human driving factors of wildfires in Spain. **Natural Hazards**, Springer, v. 84, n. 3, p. 2049–2070, 2016.
- SANTANA, O. A. Minimum age for clear-cutting native species with energetic potential in the Brazilian semi-arid region. **Canadian Journal of Forest Research**, v. 47, n. 3, p. 411–417, 2016.
- SANTOS, H. et al. O novo mapa de solos do Brasil: legenda atualizada. **Embrapa Solos-Documentos**, Rio de Janeiro: Embrapa Solos, p. 67, 2011.
- SEREGHETTI, G. C. et al. Efeito do espaçamento no crescimento e na densidade básica da madeira de *Eucalyptus urophylla* x *Eucalyptus grandis* em florestas de ciclo curto. **Energia na Agricultura**, v. 30, n. 3, p. 257–262, 2015.
- SILVA, J. A. A. et al. Produtividade volumétrica de clones de *Eucalyptus* spp. no Polo Gesseiro do Araripe, Pernambuco. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 10, p. 240–260, 2013.
- SILVA, J. A. A.; NETO, F. d. P. **Princípios básicos de dendrometria**. [S.l.]: Imprensa Universitária da UFRPE, 1979. 198 p.
- SILVEIRA, E. R.; REINER, D. A.; SMANIOTTO, J. R. Efeito do espaçamento de plantio na produção de madeira e serapilheira de *Eucalyptus dunnii* na região sudoeste do Paraná. **Revista Técnico-Científica**, v. 1, n. 2, p. 1–9, 2014.



## CAPÍTULO 5

---

### **Modelos não lineares modificados e aprendizagem de máquinas para modelagem do volume de madeira em clones de *Eucalyptus***

#### **Resumo**

A estimativa precisa do volume de árvores em povoamentos florestais é de fundamental importância para o desenvolvimento de planos de manejo florestal sustentável. O desenvolvimento de modelos florestais com boa precisão são necessários para que o agente manejador tome as decisões corretas sobre as intervenções no plano de manejo. Desta forma, objetivou-se com esta pesquisa ajustar e avaliar o desempenho dos modelos de regressão não linear simétricos, generalizados e os algoritmos de aprendizagem de máquinas Redes Neurais Artificiais, Máquinas de Vetores de Suporte e Florestas Aleatórias para estimativa do volume individual de clones de *Eucalyptus* em Região Semiárida. Sendo assim, foi desenvolvido um experimento com clones de *Eucalyptus* no município de Araripina-PE, onde aos 8 anos de idade foram mensurados o volume de madeira  $m^3$  e o diâmetro à altura do peito (DAP em cm) e altura (Ht em m). Foram ajustados os modelos de Schumacher e Hall, Silva e Bailey, Spurr, e Brody na forma normal, na forma com distribuição simétrica dos resíduos e na forma com estrutura generalizada. Os algoritmos da aprendizagem de máquinas treinados foram: redes neurais artificiais do tipo perceptron de múltiplas camadas, máquinas de vetores de suporte e floresta aleatória. Os critérios de avaliação da qualidade de ajuste foram: índice de ajuste de Schlaegel corrigido ( $IA_c$ ), erro padrão de estimativa ( $S_{xy}\%$ ), raiz do erro quadrático médio (REQM), erro absoluto médio (EAM) e análise gráfica dos resíduos. A performance das equações/modelos foram comparadas a partir do conjunto de base de teste que não foi empregado no ajuste/treinamento. Não foi constatada diferenças significativas no conjunto de dados de teste entre as estimativas do volume realizadas por meio dos modelos de regressão não linear, regressão não linear com erros simétricos, regressão não linear com estrutura generalizada, redes neurais artificiais, florestas aleatórias, máquinas de vetores de suporte e o volume de madeira real.

**Palavras-chave:** Polo Gesseiro do Araripe-PE, modelos simétricos, modelos não lineares

res generalizados, redes neurais artificiais, máquinas de vetores de suporte, florestas aleatórias.

## Abstract

The precise estimate of the volume of trees in forest stands is of fundamental importance for the development of sustainable forest management plans. The development of forest models with good precision is necessary for the management agent to make the correct decisions about interventions in the management plan. Thus, the objective of this research was to adjust and evaluate the performance of symmetric, generalized nonlinear regression models and the machine learning algorithms Artificial Neural Networks, Support Vector Machines and Random Forests to estimate the individual volume of *Eucalyptus* clones in Semiarid Region. Therefore, an experiment was developed with *Eucalyptus* clones in the municipality of Araripina-PE, where at 8 years of age the volume of wood  $m^3$  and the diameter at breast height (DBH in cm) and height (Ht in m). The models of Schumacher and Hall, Silva and Bailey, Spurr, and Brody were adjusted in the normal form, in the symmetrically distributed form of the residues and in the form with generalized structure. The trained machine learning algorithms were: multilayer perceptron artificial neural networks, support vector machines and random forest. The criteria for assessing the quality of fit were: corrected Schlaegel adjustment index ( $AI_c$ ), standard error of estimate ( $S_{xy}$  %), root of the mean square error (RMSE), error mean absolute value (EMA) and graphical analysis of residues. The performance of the equations/models were compared from the test base set that was not used in the adjustment/training. No significant differences were found in the test data set between volume estimates made using nonlinear regression models, nonlinear regression with symmetric errors, nonlinear regression with generalized structure, artificial neural networks, random forests, vector machines of support and the real volume of wood.

**Key-words:** Gypsum Pole of Araripe-PE, simetric models, non linear generalized models, artificial neural network, support vector machines, random forest.

## 5.1 Introdução

A produção em volume é uma das informações mais importantes para determinar o sucesso de um plano de manejo florestal (MENDONÇA et al., 2017). Logo, a estimativa precisa da produção de povoamentos florestais é fundamental para o de-

envolvimento de planos de manejos eficientes. Para o estudo da dinâmica florestal as técnicas de modelagem constituem uma das mais importantes metodologias (GUERA et al., 2019b). São nestas técnicas que se concentram a maior parte dos estudos de manejo, pois em geral, envolvem um grande número de variáveis.

O uso dos modelos de crescimento e produção se tornou popular a partir da década de 1930, dando início a era da estatística do crescimento ou modelagem da produtividade (SHIFLEY et al., 2017). Os primeiros modelos desenvolvidos são basicamente empíricos e estes usam relações estatísticas (análise de regressão) entre variáveis preditoras (diâmetro à altura do peito, altura, idade, etc.) e a variável resposta (SHIFLEY et al., 2017; FABRIKA; VALENT; MERGANIČOVÁ, 2019). As estimativas realizadas com as equações obtidas por meio dos modelos empíricos são formuladas geralmente para árvores individuais. Estas equações são empregadas em diversas aplicações de várias regiões espalhadas por todo o mundo (SCHNEIDER et al., 2016).

Os modelos empíricos para produção florestal podem apresentar estrutura linear ou não linear, sendo os não lineares mais difundidos, pois apresentam melhores ajustes aos fenômenos biológicos (SILVA et al., 2016/2017). Para o emprego correto destes modelos os pressupostos da análise de regressão devem ser atendidos, tais pressupostos são apresentados em Montgomery, Peck e Vining (2013). No entanto, é comum que em algumas bases de dados de fenômenos florestais tanto a heterogeneidade da variância residual quanto a autocorrelação dos resíduos violem as pressuposições da teoria da análise de regressão (CALEGARIO et al., 2005; HESS et al., 2015). Bem como, para alguns casos existe falta de aderência a pressuposição de normalidade dos resíduos.

Há muitos anos os modelos não lineares com suposição da distribuição normal para os resíduos e para a variável resposta são os mais empregados no estudo de fenômenos florestais (HESS et al., 2015). No entanto, Filho et al. (2012) afirmam que as estimativas dos coeficientes para equações com tais violações são sensíveis à observações extremas. Diante desta problemática novas formas de modelagem estão em constante desenvolvimento.

Algumas das metodologias empregadas na modelagem do crescimento e produção diante de violações dos pressupostos da regressão são: modelos não lineares com distribuição simétrica para os resíduos (FILHO et al., 2012; SALES et al., 2015), para estes modelos é considerada para os erros a classe de distribuições simétricas que possuem caudas mais pesadas que a distribuição normal; modelos lineares generalizados (HESS et al., 2015; WEBER et al., 2018), nestes modelos é dado a opção para a distribuição da variável resposta pertencer a família exponencial, bem como, permite

que a relação entre a média da variável resposta e o preditor linear  $\eta$  seja mais flexível; por fim, os modelos não lineares generalizados, estes permitem que a estrutura não linear seja expandida para a família exponencial de distribuições.

Outras formas de modelagem que podem ser empregadas diante desta problemática e que tem crescido nos últimos anos são os algoritmos de aprendizagem de máquinas. Dentre os principais algoritmos para estimativa da produção em volume estão: As redes neurais artificiais - RNA (ÖZÇELİK et al., 2010; SANQUETTA et al., 2018; GUERA et al., 2019b), uma RNA é definida como um processador constituído por unidades de processamento interconectadas, chamado de neurônio, capazes de calcular funções, e estão distribuídos em camadas, as quais estão conectadas entre si por pesos que possuem o poder de memorizar o conhecimento experimental (VALENÇA, 2010; BINOTI et al., 2015); As máquinas de vetores de suporte - MVS (REIS et al., 2018; ABREU, 2019), as MVS's estão fundamentadas na teoria de aprendizagem estatística, propostas por Vapnik (1998), essa metodologia busca a minimização do risco estrutural, tentando minimizar um limite superior do erro de generalização, diferentemente das RNAs que buscam a minimização do erro em cada exemplo apresentado (LORENA; CARVALHO, 2007); e por fim, as florestas aleatórias - FA (REIS et al., 2018; REIS et al., 2019), as FA's são consideradas um método estatístico não paramétrico que usa uma abordagem baseada na amostragem do banco de dados para criar um conjunto de árvores de classificação ou regressão (GENUER; POGGI; TULEAU-MALOT, 2015).

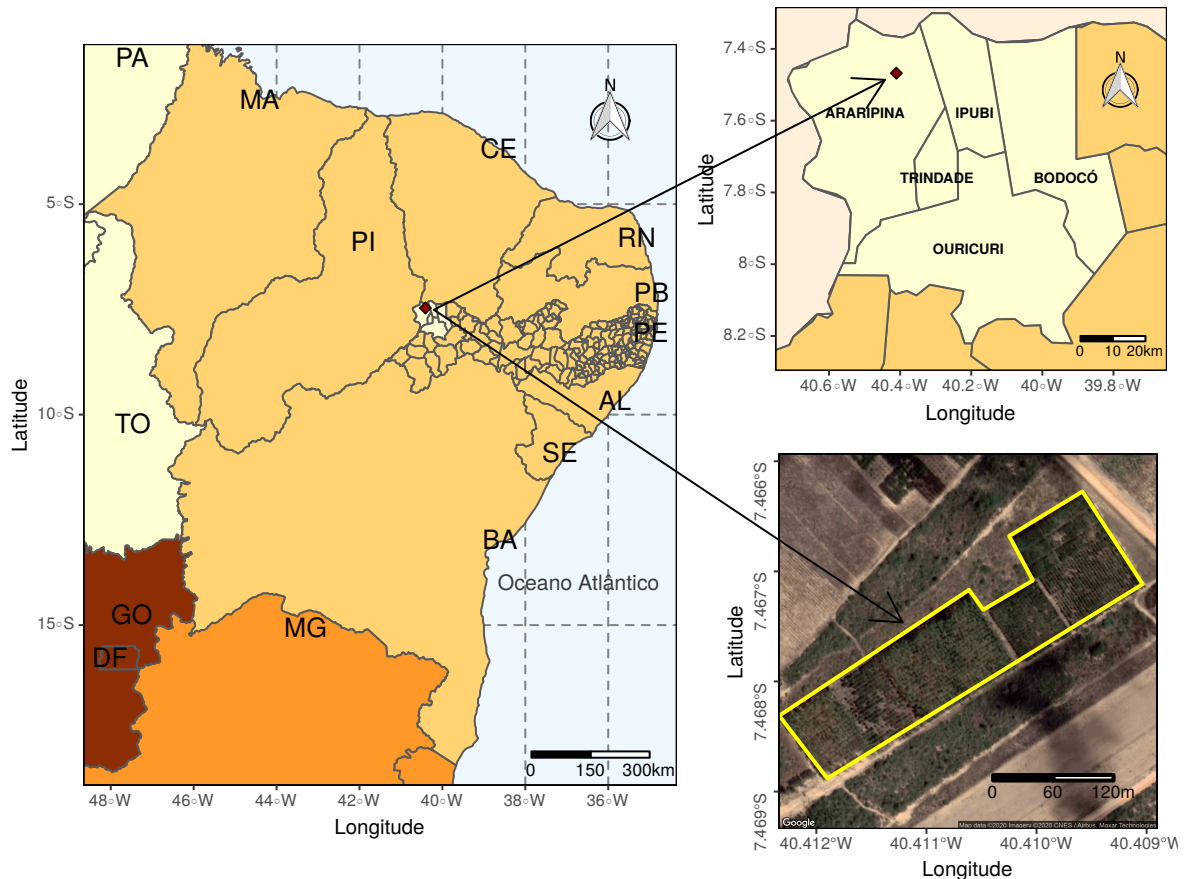
Diante do exposto, objetivou-se com esta pesquisa ajustar e avaliar o desempenho dos modelos de regressão não linear simétricos, generalizados e os algoritmos de aprendizagem de máquinas Redes Neurais Artificiais, Máquinas de Vetores de Suporte e Florestas Aleatórias para estimativa do volume individual de clones de *Eucalyptus* em Região Semiárida.

## **5.2 Material e Métodos**

### **5.2.1 Descrição do experimento**

O experimento foi desenvolvido no município de Araripina-PE (coordenadas 07° 29' 00" S e 40° 36' 00" W, altitude de 816 metros) na Estação Experimental do Instituto Agrônomo de Pernambuco (IPA) região semiárida do Nordeste brasileiro (Figura 9). A precipitação média anual da região é de 740 mm e possui elevados índices de evaporação e a temperatura média do ar de 25 °C. Segundo classificação de Köppen, o clima da região é do tipo BSh, semiárido quente, caracterizado por longos períodos de seca e chuvas irregulares.

Figura 9 – Mapa do Nordeste destacando o Polo Gesseiro do Araripe com ênfase na área experimental localizada no Instituto Agrônômico de Pernambuco, município de Araripina-PE.



A instalação do experimento ocorreu no mês de março de 2010 em uma área de 7,5 hectares, neste foram plantados três clones de *Eucalyptus*, sendo dois provenientes de *Eucalyptus urophylla* (C39 e C41) e um de *Eucalyptus brassiana* (C11), considerando cinco níveis de arranjos espaciais, estes são; 2 m x 1 m, 2 m x 2 m, 3 m x 2 m, 3 m x 3 m e 4 m x 2 m. O delineamento experimental utilizado foi o inteiramente aleatório em arranjo fatorial (3 x 5) com 4 repetições. No mês de março de 2018 foi realizada a colheita do experimento por meio da derrubada de todas as árvores. Foram tomadas medidas do diâmetro à altura do peito (DAP) e altura total (Ht), bem como foram tomadas medidas de diâmetro ao longo do fuste nos intervalos: 0,30 m; 0,50 m; 0,70 m; 0,90 m; 1,10 m; 1,30 m; 1,50 m; 1,70 m; 1,90 m; 2,30 m, a partir de 2,30 m foram obtidas medidas a cada um metro. Os valores dos diâmetros foram utilizados para cálculo do volume de madeira da árvore via cubagem rigorosa por meio da metodologia de Smalian, conforme SILVA e Neto (1979).

Conforme análise de variância e aplicação do teste de médias para os dados do experimento, realizada no Capítulo 4, o conjunto de dados para a modelagem foi composto pelos clones C41 e C11 por não diferirem estatisticamente entre si e pelos espaçamentos 2 m x 2 m, 3 m x 2 m e 3 m x 3 m também por não apresentarem diferenças significativas, o que totalizou uma amostra com n igual a 1204 árvores.

As variáveis empregadas nos modelos foram a variável resposta volume ( $m^3$ ) e as variáveis explicativas DAP (cm) e Ht (m). Para os algoritmos de aprendizagem de máquinas também será realizada uma tentativa com a adição de mais variáveis explicativas Área vital, definida pela área que cada planta possui para o seu desenvolvimento (o espaçamento 2 m x 2 m corresponde há uma área vital de 4  $m^2$ ), e o tipo de Clone.

Com a finalidade de garantir as mesmas condições para a comparação entre os modelos não lineares ajustados e os algoritmos de aprendizagem de máquinas treinados a base de dados foi dividida em base de treino (80%) e teste (20%).

### 5.2.2 Modelos não lineares

Os modelos não lineares, conforme Silva et al. (2013) e Campos e Leite (2017), utilizados para estimar o volume em madeira por árvore individual do plantio de clones de *Eucalyptus* no Polo Gesseiro do Araripe são descritos na Tabela 6.

Tabela 6 – Modelos não lineares para estimativa do volume de madeira ( $m^3$  árvore<sup>-1</sup>) para clones de *Eucalyptus* plantados no Polo Gesseiro do Araripe-PE.

Nº	Modelos	Autores
M1	$v_i = \beta_0 \cdot DAP_i^{\beta_1} \cdot Ht_i^{\beta_2} \cdot \epsilon_i$	Schumacher e Hall
M2	$v_i = \beta_0 \cdot (DAP^2 \cdot Ht)^{\beta_1} \cdot \epsilon_i$	Spurr
M3	$v_i = \beta_0 \cdot e^{\beta_1 \cdot (\beta_2^{DAP_i} + \beta_3^{Ht_i})} + \epsilon_i$	Silva e Bailey
M4	$v_i = \beta_0 - \beta_1 \cdot e^{(-\beta_2 DAP_i - \beta_3 Ht_i)} + \epsilon_i$	Brody

Em que,  $v_i$  é o volume por árvore em ( $m^3$ );  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  são os parâmetros dos modelos,  $DAP_i$  é o diâmetro à altura do peito em (cm),  $Ht_i$  é a altura total em (m) e  $\epsilon_i$  é uma componente do erro aleatório e  $\epsilon_i \sim N(0, \sigma^2)$ .

Estes modelos foram ajustados por meio do método dos mínimos quadrados ponderados utilizando o algoritmo iterativo de Gauss-Newton para obtenção dos coeficientes.

A heterocedasticidade dos resíduos é frequente em modelos florestais, para tentar corrigir o desvio há este pressuposto será realizada uma transformação na matriz de variância dos resíduos, testou-se as seguintes modificações:

- $E(\epsilon_i^2) = \sigma^2 \cdot \sqrt{(DAP^2 \cdot Ht)}$
- $E(\epsilon_i^2) = \sigma^2 \cdot \frac{1}{(DAP^2 \cdot Ht)^{0.5}}$
- $E(\epsilon_i^2) = \sigma^2 \cdot \frac{1}{(DAP^2 \cdot Ht)^2}$

É comum que os resíduos de modelos florestais se desviem do pressuposto de normalidade, para tais casos será empregada a metodologia de modelos simétricos, a qual atribui distribuições com caudas mais pesadas aos resíduos. Cysneiros, Paula e Galea (2005) definem um modelo de regressão com erros simétricos da seguinte forma: Sejam  $\epsilon_1, \epsilon_2, \dots, \epsilon_n$  variáveis aleatórias independentes com a seguinte função densidade

$$f_{\epsilon_i}(\epsilon) = \frac{1}{\sqrt{\delta}} \cdot g\left[\frac{(\epsilon)^2}{\delta}\right] \quad (5.1)$$

sendo que,  $\epsilon \in \mathbb{R}$  e  $g(\cdot)$  tem as propriedades de uma função densidade de probabilidade. O modelo simétrico não linear geral foi definido como:

$$v_i = \mu_i(\beta; \mathbf{x}_i) + \epsilon_i \quad (5.2)$$

Em que,  $\mu_i = \mu_i(\beta; \mathbf{x}_i)$  é parte sistemática dos modelos não lineares empregados neste estudo. Sendo  $\beta = (\beta_1, \dots, \beta_p)^T$  uma matriz de parâmetros diferenciável,  $v_i = (v_1, \dots, v_n)^T$  o vetor contendo a variável resposta volume e  $\mathbf{x}_i = (x_{iDAP}, x_{iHt})^T$  uma matriz com as variáveis explicativas  $DAP$  e  $Ht$  e  $\epsilon_i \sim S(o, \delta)$ . Assim, a densidade de  $v_i$  foi expressa por

$$f_{v_i}(v_i) = \frac{1}{\sqrt{\delta}} \cdot g[(w_i)] \quad (5.3)$$

sendo,  $w_i = \frac{(v_i - \mu_i)^2}{\delta}$  e  $v_i \sim S(o, \delta)$ . Desta forma, tem-se que o modelo dado em 5.2 e 5.3 é um modelo de regressão não linear simétrico. Para a estimativa de parâmetros deste tipo de modelo foi utilizado o logaritmo da função de verossimilhança de  $\theta = (\theta^T, \delta)^T$  que foi calculado pela expressão geral dada por

$$L(\theta) = -\frac{n}{2} \log(\delta) + \sum_{i=1}^n \log[g(w_i)]. \quad (5.4)$$

Assim, por meio de processos iterativos foram estimados os parâmetros dos modelos e do valor de  $\delta$ . Foram empregadas para  $L(\theta)$  e  $\mu_i$  as seguintes distribuições simétricas; t-Student, t-Student Generalizada, Logística I, Logística II e Exponencial Potência.

Mesmo se atribuindo distribuições de probabilidade com caudas mais pesadas a falta de aderência dos resíduos pode persistir. Para esses casos se fará uso dos modelos não lineares generalizados. Para o modelo de Schumacher sua representação generalizada é dada por:

$$v = \beta_0 \cdot DAP_i^{\beta_1} \cdot Ht_i^{\beta_2} \quad (5.5)$$

E duas funções:

- uma função de ligação que descreve como a média,  $E(V) = \mu$ , depende do preditor não linear

$$g(\mu) = v \quad (5.6)$$

- uma função da variância que descreve como a variância,  $Var(V)$  depende da média

$$Var(V) = \psi V(\mu) \quad (5.7)$$

Em que  $\psi$  é um parâmetro de dispersão. Dessa forma, a distribuição considerada não precisa necessariamente ser a normal podendo assumir qualquer distribuição contínua da família exponencial. Para a modelagem do volume de madeira foram consideradas as distribuições Gama e Normal Inversa com as funções de ligação canônica, inversa e Log. A estimativa dos parâmetros foi realizada pelo método da máxima verossimilhança com o algoritmo iterativo de Newton.

### 5.2.3 Algoritmos de aprendizagem de máquinas

#### 5.2.3.1 Redes neurais artificiais (RNA)

Na implementação de modelos de RNA's diversas configurações precisam ser ajustadas, tais como, a topologia da rede, as características dos neurônios, os hiperparâmetros, a base de dados, as regras de aprendizagem, entre outros.



Para desenvolver o sistema de treino dos algoritmos de aprendizagem de máquinas a base de dados já dividida em em treino (80%) e teste (20%) foi normalizada com a finalidade de diminuir a amplitude de valores extremos e aumentar a velocidade na convergência da rede. As variáveis explicativas empregadas para estimar o volume de madeira por árvore em clones de *Eucalyptus* foram o DAP e a Ht. Para o treinamento da rede foi utilizado a validação cruzada com o método *k-fold*. Sendo assim, o conjunto de treinamento foi dividido em *5-fold*, ou seja, o conjunto foi dividido em 5 partes iguais e o treinamento foi realizado com  $5 - 1$  partes, ficando uma parte para validação, este procedimento foi repetido 5 (*k*) vezes, deixando sempre um dos subconjuntos de fora.

De forma geral uma RNA é formada por camadas (entrada, intermediária e saída), todas estas camadas são compostas por neurônios. Cada neurônio possui um peso sináptico ( $w_i$ ) associado, os quais são responsáveis por fazer a conexão de todos os neurônios de uma camada (entrada)  $l$  com os neurônios de uma subsequente (intermediária)  $l + 1$ . Toda camada, exceto a de saída, ou cada neurônio de uma camada possui um erro associado chamado de *bias* ( $b$ ). O número de camadas intermediárias e a quantidade de neurônios precisa ser definido antes do treinamento. Pelo teorema de aproximação universal uma RNA com apenas uma camada intermediária (com uma função de ativação não linear) é suficiente para aproximar qualquer função mensurável de Borel em um espaço finito-dimensional (GOODFELLOW; BENGIO; COURVILLE, 2016). Desta forma, o problema se reduz em definir a quantidade de neurônios que compõe a camada intermediária.

Para atenuar esta complexidade, a busca pela topologia ideal foi realizada utilizando um *grid* de hiperparâmetros disponível no pacote *h2o*, este, é capaz de receber um conjunto de diferentes configurações (topologia, funções de ativação, hiperparâmetros etc.) para um sistema de RNA e fazer combinações com todas as características, retornando assim, a melhor configuração. Desta forma, foi ajustado inicialmente a quantidade de neurônios na camada intermediária. Para se ter uma ideia inicial do número de neurônios, conforme apresentado por Valença (2010), foi empregado a fórmula de Hetcht-Nielsen dada por:

$$N_{int} = 2 \cdot N_{ent} + 1 \quad (5.8)$$

Em que,  $N_{int}$  é o número de neurônios na camada intermediária e  $N_{ent}$  é o número de neurônios na camada de entrada. Sendo assim, foi adicionado ao *grid* diferentes quantidades de neurônios em torno do valor calculado.

Cada peso sináptico é associado ao próximo neurônio por meio de uma função

de ativação. Assim, foram consideradas as seguintes funções: tangente hiperbólica, *rectifier* e *maxout*. A vantagem de se utilizar estas funções é que elas permitem o uso da técnica denominada *Dropout*. Um dos problemas relacionado ao treino de uma rede é que em alguns casos os pesos sinápticos ajustados podem apresentar alta correlação entre si, o que gera sobre ajuste (*overfitting*) na rede (SRIVASTAVA et al., 2014). O uso do *Dropout* reduz a correlação (co-adaptações) entre os neurônios. Já que, para cada ciclo de treinamento ocorre o desligamento aleatório de alguns neurônios da camada oculta, ou seja, tanto a fase de propagação quanto a fase de retropropagação funcionará com uma quantidade reduzida de neurônios, para cada ciclo diferentes neurônios são desativados.

Outra forma de diminuir o sobre ajuste e evitar que apenas poucos neurônios sejam responsáveis pela estimativa da resposta é utilizando as técnicas de regularização L1 e L2. A regularização L1 permite que o modelo se torne invariante a entradas ruidosas, para isso ela torna o vetor de pesos sinápticos mais próximo de zero. Já a L2 é responsável por penalizar pesos que causam um grande impacto na rede e assim a rede pode aprender melhor a partir de todas entradas disponíveis. Estas duas técnicas geram uma mudança na função de erro da rede diferentemente do uso de *Dropout*. Sendo assim, diferentes configurações para estes três parâmetros foram testadas. Para o *Dropout* foi considerado inicialmente um valor de 0,5 (este varia de 0 até 1). Para L1 e L2 foram considerados os seguintes valores: 0; 0,001; 0,00001 e 0,000001.

O número de ciclos ou épocas que serão utilizados para treinar a rede também pode afetar o treinamento causando *overfitting* ou sendo insuficiente para uma boa generalização. Então, adicionou-se ao *grid* diferentes número de ciclos (100, 400, 500, 1000). A taxa erro de tolerância para parada do treinamento foi de 0,001. A função de erro ou de custo, utilizada para minimizar o erro de treinamento foi o erro quadrático médio, dado por:

$$EQM(\hat{\theta}) = E[\hat{\theta} - \theta]^2 \quad (5.9)$$

Esta função é minimizada via gradiente descendente estocástico sendo calculada por meio do algoritmo *backpropagation*.

### 5.2.3.2 Florestas aleatórias (FA)

As mesmas configurações da base de dados apresentada anteriormente foram consideradas no treinamento das Florestas Aleatórias (*Random Forest*).

O desenvolvimento de uma Floresta Aleatória consiste em gerar  $R$  amostras aleatórias (via *bootstrap*, ou seja, a amostragem é feita com reposição) para a base de treino. Para cada  $R_i$  uma árvore de regressão (*ntrees*, mais conhecida por árvore de decisão usadas para problemas de classificação) é criada. Em cada árvore são desenvolvidas regras de decisão (*nó*), na floresta aleatória cada *nó* pode utilizar  $v/3$  variáveis explicativas ( $v$  = número de variáveis) na construção da regra, o valor  $v/3$  (*mtries*, a seleção das variáveis é feita aleatoriamente) também pode ser otimizado. Quanto mais regras (*max\_depth*) são criadas, mais a árvore de regressão se ajusta a base de treino, no entanto se perde poder de generalização. Então para a divisão de cada *nó* são escolhidas variáveis que maximizam a redução de uma função erro, e dessa forma, é reduzido conjuntamente o viés e a variância da estimativa. Na última ramificação, chamada de folha, cada árvore de regressão gera uma estimativa da variável resposta com base nas regras construídas a partir das  $v/3$  variáveis explicativas (selecionadas aleatoriamente). A estimativa geral é dada por:

$$\hat{V}_i = \frac{1}{R} \sum_{i=1}^R f^r(V_i) \quad (5.10)$$

Os principais hiperparâmetros que podem ser ajustados no treinamento são: a quantidade de árvores de decisões (*ntrees*, foi incrementado de 1 até 1000, parando de incrementar quando se estabilizou a variância e o bias), quantidade de atributos ou variáveis que serão aleatoriamente utilizadas em cada árvore (*mtries*, neste caso foi fixado em 1) e profundidade que representa o quão profundo uma árvore pode crescer (*max\_depth*, foi utilizado um máximo de 30). Todos estes hiperparâmetros foram adicionados em um *grid* de busca.

### 5.2.3.3 Máquinas de vetores de suporte (MVS)

A construção de uma máquina de vetor de suporte para um problema de regressão não linear consiste em mapear um vetor de entrada  $\mathbf{x}$  em um espaço ( $I$ ) de características de alta dimensionalidade  $H$ , assim,  $\mathbf{x} \in \mathbb{R}^I \mapsto \phi(\mathbf{x}) \in \mathbb{R}^H$  com uma função kernel  $\phi(\mathbf{x})$  para encontrar os vetores de suporte no conjunto de treinamento, que estimam  $y$ , com a menor variação possível. Desta forma, uma MVS para regressão não linear encontra os vetores de suporte (coeficientes) minimizando a seguinte função, conforme Drucker et al. (1997):

$$L(\alpha) = -\epsilon \sum_{i=1}^N (\alpha_i^* + \alpha_i) + \sum_{i=1}^N y_i (\alpha_i^* - \alpha_i) - \frac{1}{2} \sum_{i,j=1}^N (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) \phi(x_i, x_j) \quad (5.11)$$

Com as restrições:

$$\begin{aligned}
 0 &\leq \alpha_i \leq C \quad i = 1, \dots, N \\
 0 &\leq \alpha_i^* \leq C \\
 \sum_{i=1}^N \alpha_i^* &= \sum_{i=1}^N \alpha_i
 \end{aligned} \tag{5.12}$$

E  $(\alpha_i, \alpha_i^*)$  são  $N$  pares de multiplicadores de Lagrange. No desenvolvimento dos treinos com MVS foram utilizadas as seguintes funções de kernel:

- Função linear;

$$\phi(x_i, x_j) = x_i'x_j \tag{5.13}$$

- Função polinomial;

$$\phi(x_i, x_j) = (\gamma x_i'x_j + coef0)^{gp} \tag{5.14}$$

- Função de base radial ou Gaussiana;

$$\phi(x_i, x_j) = e^{(-\gamma|x_i-x_j|^2)} \tag{5.15}$$

Desta forma, os hiperparâmetros que são ajustados para melhorar a estimativa e também evitar o superajustamento são:  $C$  que é um hiperparâmetro de custo, se  $C$  é grande, é dada maior ênfase para o erro, por outro lado, se  $C$  é pequeno maior ênfase é dada para a norma dos pesos, o que pode melhorar a generalização para novos conjuntos de dados;  $\gamma$  é responsável por controlar a suavização da decisão (definição do hiperplano), ou seja, controla a variação do modelo;  $gp$  é o grau do polinômio para a função polinomial;  $coef0$  é o coeficiente do polinômio e  $\epsilon$  define a quantidade de vetores de suporte que serão empregados. Para ajustar estes hiperparâmetros foi empregado um *grid* de busca com as atribuições:  $C$  variando de 2 até 256;  $\gamma$  variando de 0,1 até 0,9 e  $\epsilon$  indo de 0 até 0,9.  $gp$  e  $coef0$  foram setados em 3 e 0, respectivamente, para não aumentar demais a complexidade do *grid*.

### 5.3 Estatísticas para avaliação dos modelos de regressão e de aprendizagem de máquinas

A avaliação da qualidade do ajuste dos modelos de regressão e dos algoritmos de aprendizagem de máquinas foi realizada de acordo com as seguintes estatísticas: índice de ajuste de Schlaegel corrigido ( $IA_c$ ), erro padrão de estimativa em porcentagem

$(S_{xy}\%)$ , raiz do erro quadrado médio ( $REQM$ ), erro absoluto médio ( $EAM$ ) e análise gráfica dos resíduos.

O índice de ajuste ( $IA$ ) foi calculado conforme Schlaegel (1984):

$$IA = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5.16)$$

Sendo que,  $y_i$  é a variável dependente observada,  $\bar{y}_i$  é a média da variável dependente observada,  $\hat{y}_i$  é a estimativa da variável dependente e  $n$  é o número de observações. O IA corrigido para o número de parâmetros  $p$  do modelo é dado por:

$$IA_c = 1 - \left( \frac{n-1}{n-p} \right) \cdot (1 - IA) \quad (5.17)$$

O erro padrão de estimativa ( $S_{xy}\%$ ), quanto menor o ( $S_{xy}\%$ ), melhor o modelo ajustado, foi calculado pela seguinte expressão:

$$S_{xy}\% = \left( \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p}} \right) \cdot \frac{100}{\bar{y}} \quad (5.18)$$

Para o processo de validação dos algoritmos de aprendizagem de máquinas se utilizou a raiz do erro quadrático médio, expresso a seguir,

$$REQM = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (5.19)$$

e o erro absoluto médio, dado por:

$$EAM = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (5.20)$$

Sendo  $N$  o tamanho da amostra.

Os resíduos em porcentagem  $E_i\%$  foram obtido por:

$$E_i\% = \frac{(y_i - \hat{y}_i)}{y_i} \cdot 100 \quad (5.21)$$

A análise gráfica dos resíduos foi realizada com base no gráfico dos resíduos em porcentagem versus valores ajustados para diagnóstico de *outliers* e heterocedasticidade, gráfico de resíduos ( $i$ ) versus resíduos ( $i + 1$ ) para diagnóstico de autocorrelação e gráfico do histograma dos resíduos para avaliação da distribuição de probabilidade.

Com a finalidade de se atenuar o viés de amostragem, tanto para os modelos de regressão modificados quanto para os algoritmos de aprendizagem de máquinas, foi empregado reamostragens via *bootstrap*. Para cálculo da performance dos modelos e dos algoritmos foram geradas 100 amostras aleatórias de tamanho 200 a partir da base de teste, a qual possuía 241 exemplos. O resultado final da performance é dado pela média das estatísticas nas 100 amostras. Sendo assim para a comparação entre os modelos será realizada estimativas sobre as amostras dos dados de teste e calculadas as estatísticas IA, REQM e EAM e para facilitar a visualização das diferenças também será criado um ranque variando entre de 1 à 15, onde as melhores equações serão pontuadas de forma crescente. E o resultado final será representado pela somatória do ranqueamento.

Para verificar diferenças entre as equações obtidas do ajuste dos modelos de regressão e os modelos treinados com os algoritmos de aprendizagem de máquinas com o volume real dos clones de *Eucalyptus* será realizada a análise de variância onde todas as equações/modelos serão considerados o tratamento (15 Tratamentos mais 1 Tratamento da variável volume real) e a variável resposta será as estimativas e o volume real. Para cada amostra gerada será realizada uma análise de variância e na presença de diferenças estatísticas será empregado o teste de comparação de médias de Scott-Knott.

#### **5.4 Recursos computacionais para ajuste dos modelos**

Todas as análises foram realizadas com a linguagem R (R Core Team, 2016). Os procedimentos para ajustes dos modelos não lineares clássicos de produção e crescimento foram realizados por meio da função `nls` implementada conforme Bates e Watts (1988), esta baseada no método dos mínimos quadrados ordinários com o algoritmo de Gauss-Newton para as iterações dos parâmetros. Os modelos com diferentes distribuições simétricas para os resíduos foram ajustados por meio do pacote `elliptical`, conforme descrito por Cysneiros, Paula e Galea (2005). Para ajustar os modelos lineares generalizados foi utilizado a função `glm` do pacote `stats`, padrão do R. Para o ajuste dos modelos não lineares generalizados foi empregado o pacote `gnm` descrito em Turner e Firth (2015).

Os treinamentos das Redes Neurais Artificiais e das Florestas Aleatórias foram realizados com a plataforma *H2O*. Seu uso é totalmente compatível com a linguagem R por meio de um pacote de mesmo nome, `H2O` (LEDELL et al., 2019). Esta plataforma disponibiliza um cluster, permitindo assim, o desenvolvimento de análises mais com-

plexas com pouco tempo de execução. O treinamento das Máquinas de Vetores de Suporte foi realizado com o pacote `e1071`, o qual disponibiliza a função para criação do `grid` de busca com uso da validação cruzada *k-fold* (MEYER et al., 2017).

As estatísticas de qualidade de ajuste e as análises de resíduos dos modelos ajustados foram obtidas por meio do pacote `modelTools` de autoria própria.

## 5.5 Resultados e Discussão

### 5.5.1 Modelos não lineares ajustados

#### 5.5.1.1 Modelos não lineares clássicos

Os coeficientes para as equações obtidas por meio dos modelos não lineares ajustados e as estatísticas de qualidade do ajuste são apresentadas na Tabela 7.

Tabela 7 – Estatísticas de qualidade e coeficientes das equações não lineares para estimar o volume ( $\text{m}^3 \text{ árvore}^{-1}$ ) dos clones de *Eucalyptus* plantados no Polo Gesseiro do Araripe-PE.

N°	Equações	Coeficientes				IA <sub>c</sub>	S <sub>xy</sub> %	REQM
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$			
M1	Schumacher e Hall	5,96e <sup>-05</sup>	1,71400	1,06600	-	0,9835	11,1725	0,0045
M2	Spurr	6,87e <sup>-05</sup>	0,91270	-	-	0,9819	11,7171	0,0048
M3	Silva e Bailey	0,98788	-4,62119	0,93212	0,87112	0,9835	11,1953	0,0045
M4	Brody	-0,01239	-0,00651	-0,11492	-0,07977	0,9821	11,6467	0,0047

REQM em  $\text{m}^3 \text{ árvore}^{-1}$

Com base nas estatísticas calculadas para avaliar a qualidade dos ajustes, a equação obtida via modelo de Schumacher e Hall foi a que apresentou o maior IA<sub>c</sub> (0,9835), menor S<sub>xy</sub>% (11,1725) e REQM (0,0045) seguida da equação ajustada do modelo de Silva e Bailey, com estatísticas semelhantes. Quanto as medidas de qualidade da estimativa todas as equações possuem taxas de erro satisfatórias para pesquisas no setor florestal, conforme também foi constatado em Miguel et al. (2016) e Hall et al. (2019).

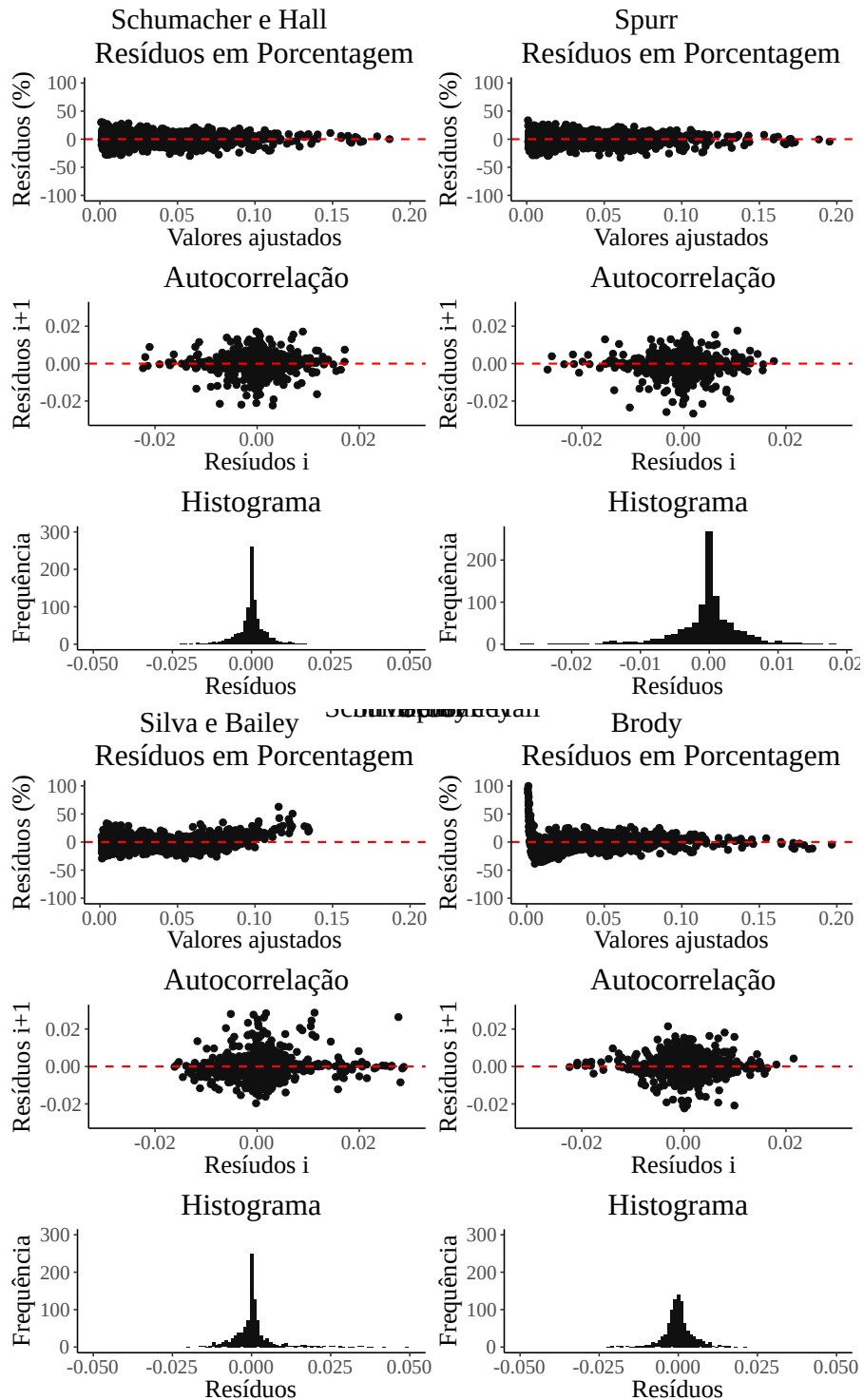
Na análise gráfica dos resíduos Figura (10) foi verificado que as equações possuem resíduos que se afastam do pressuposto de normalidade, apenas na equação resultante do modelo de Brody este afastamento é menos evidente no entanto, esta equação possui pontos de superestimativas conforme visualizado no gráfico dos

resíduos em porcentagem versus valores ajustados. As demais equações apresentam resíduos com dispersão menor e não possuem pontos que indiquem ser *outliers*. O pressuposto da normalidade em modelos de regressão em certo grau pode ser relevado, já que com o aumento do tamanho da amostra o efeito da não normalidade passa a não influenciar tão significativamente as estimativas (PINO, 2014).

Ainda de acordo com o gráfico dos resíduos em porcentagem em função dos valores ajustados foi constatado que as equações possuem um elevado grau de heterogeneidade nas variâncias do erro, este mesmo comportamento também foi verificado em Salekin et al. (2020) trabalhando com a modelagem da produção de *Eucalyptus globoidea* Blakely. Segundo Calegario et al. (2005), a presença de resíduos heterocedásticos, isto é, quando a variância não é estatisticamente semelhante em todos os pontos residuais é comum em dados biométricos. Desta forma, realizar estimativas com modelos que possuem tal característica podem levar há estimavas enviesadas, conforme apresentado extensivamente por Breusch e Pagan (1979) e White (1980). Para todas as equações geradas os resíduos não apresentaram tendência de correlação conforme exposto no gráfico de autocorrelação.



Figura 10 – Distribuição dos resíduos para as equações geradas pelos modelos clássicos de regressão.



Considerando o entendimento que equações que se desviam dos pressupostos da regressão, principalmente o de homogeneidade dos resíduos, podem gerar estimativas enviesadas, o emprego delas não é indicado ou se deve empregar com muita

ressalva para estimar o volume de madeira por árvore para o experimento com clones de *Eucalyptus* desenvolvido no Polo Gesseiro do Araripe - PE.

### 5.5.1.2 Modelos não lineares com erros simétricos

Diante do problema da falta de aderência dos resíduos ao pressuposto de normalidade foi empregado a metodologia dos modelos de regressão com distribuição simétrica dos erros. Desta forma, os modelos apresentados na Tabela 6 foram reajustados atribuindo aos resíduos as distribuições t Student, Student generalizada, Logística I, Logística II, e Exponencial potência. Para cada modelo ajustado foi selecionada a distribuição que apresentou as melhores estatísticas de ajuste e as que proporcionaram maior aderência aos resíduos.

Na Tabela 8 são apresentados os coeficientes e estatísticas dos ajustes realizados com os modelos simétricos. As distribuições que mostraram os melhores resultados foram a Exponencial Potência e a Logística I e II. Em relação a qualidade do ajuste todas as estatísticas de teste apresentaram resultados próximos ao das equações obtidas via modelos de regressão clássicos.

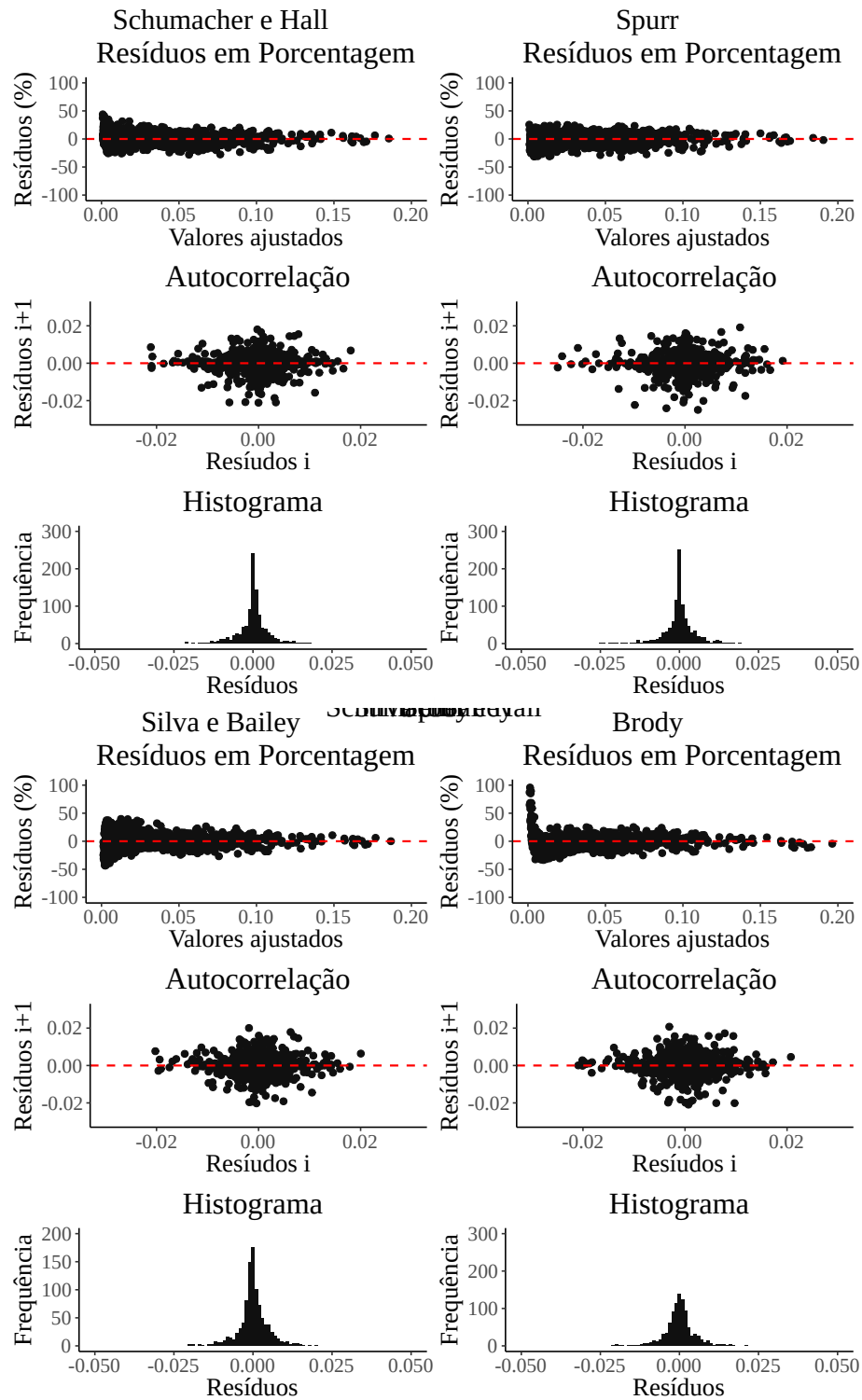
Tabela 8 – Estatísticas de qualidade e coeficientes das equações ajustadas via modelos não lineares com distribuição simétrica dos resíduos, para estimativa do volume ( $\text{m}^3 \text{árvore}^{-1}$ ) dos clones de *Eucalyptus* plantados no Polo Gesseiro.

Equações	Distribuição	Coeficientes				IA <sub>c</sub>	S <sub>xy</sub> %	REQM
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$			
Schumacher e Hall	Logística II	4,84e <sup>-5</sup>	1,67188	1,18168	-	0,9839	11,0558	0,0045
Spurr	Exponencial Potência	7,60e <sup>-5</sup>	0,89821	-	-	0,9824	11,5589	0,0047
Silva e Bailey	Logística I	1,40212	-4,62835	0,93994	0,88385	0,9835	11,1669	0,0045
Brody	Exponencial Potência	-0,01441	-0,00737	-0,11252	-0,07520	0,9824	11,5429	0,0047

REQM em  $\text{m}^3 \text{árvore}^{-1}$

Na Figura 11 é apresentada a análise dos resíduos para os modelos não lineares com erros simétricos. Por meio do gráfico dos resíduos em porcentagem versus valores ajustados foi possível constatar que a atribuição de uma distribuição com cauda mais pesada aos resíduos permitiu uma maior aderência à distribuição sugerida, no entanto, não foi capaz de estabilizar a variância, dessa forma problemas quanto a heterogeneidade dos resíduos ainda é persistente.

Figura 11 – Análise dos resíduos para as equações ajustadas por meio dos modelos não lineares com distribuição simétrica dos erros.



A metodologia dos modelos de regressão com distribuições simétricas foi avaliada por Sales et al. (2015) em um experimento com o clones de *Eucalyptus* no Agreste

de Pernambuco. Estes autores concluíram que o uso da distribuição t Student junto com o modelo de Schumacher e Hall promoveu melhores ajustes das equações em relação à distribuição Normal. Gadelha (2014) em trabalho realizado com amostras de clones do mesmo experimento, aos quatro anos de idade, ajustou modelos não lineares com as distribuições Normal, t Student e Logística II. O autor obteve resultados semelhantes entre as distribuições com uma leve superioridade obtida pelos valores dos índices de ajuste para a distribuição Normal. Sendo assim, é importante destacar que tanto para amostras pequenas quanto para amostras grandes, e povoamentos em desenvolvimento ou em ciclo de corte os modelos com distribuições simétricas apresentam estatísticas de qualidade do ajuste semelhantes à Normal.

### 5.5.1.3 Modelos não lineares generalizados

A falta de aderência aos pressupostos da regressão pelos modelos empregados também dá suporte ao uso de outras técnicas de modelagem. Ao invés de atribuir diferentes distribuições de probabilidade apenas aos resíduos foi empregado a regressão não linear generalizada, assim se avaliou a adequação da variável resposta a diferentes distribuições de probabilidade da família exponencial. Desta forma, na Tabela 9 são discutidos os resultados para os modelos não lineares generalizados ajustados.

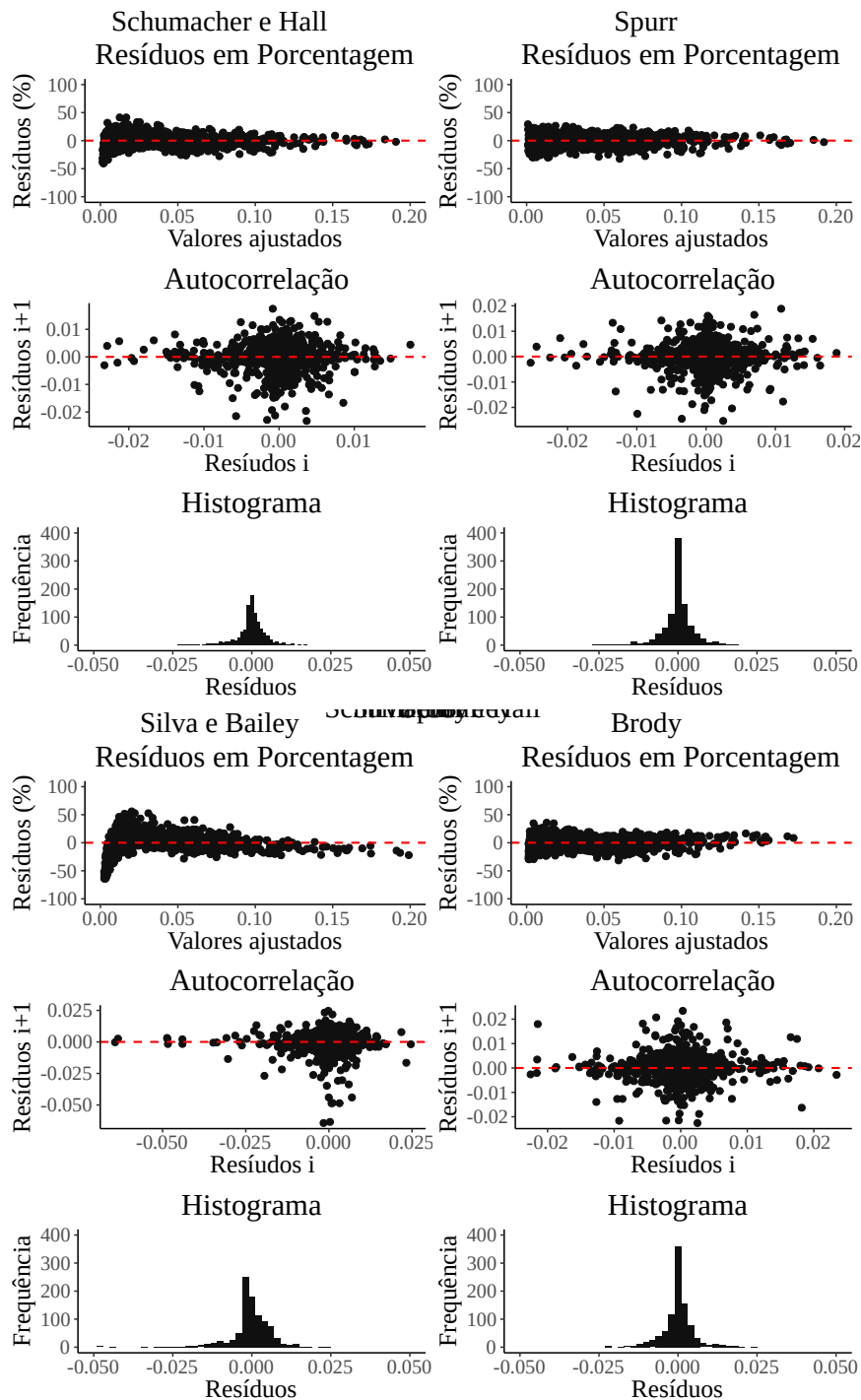
Tabela 9 – Coeficientes e estatísticas de teste das equações obtidas via modelos não lineares generalizados, para estimativa do volume por árvore ( $\text{m}^3 \text{árvore}^{-1}$ ) dos clones de *Eucalyptus* plantados na Chapada do Araripe-PE.

Modelo	Distribuição	Link	Coeficientes	Estimativas	Erro Padrão	la <sub>c</sub>	S <sub>xy</sub> %	REQM
Schumacher e Hall	Gamma	Log	$\hat{\beta}_0$	-8,37055	0,03512	0,9832	11,3071	0,0046
			$\hat{\beta}_1$	0,94743	0,00068			
			$\hat{\beta}_2$	0,96523	0,00071			
Spurr	Gamma	Canônica	$\hat{\beta}_0$	$7,26e^{-5}$	$1,59e^{-6}$	0,9823	11,5713	0,0047
			$\hat{\beta}_1$	0,90420	$3,06e^{-3}$			
			$\hat{\beta}_0^{ns}$	$2,58e^5$	$7,71e^{-5}$			
Silva e Bailey	Gamma	Canônica	$\hat{\beta}_1$	-9,76400	1,46300	0,9538	18,7185	0,0076
			$\hat{\beta}_2$	0,98270	0,00334			
			$\hat{\beta}_3$	0,97920	0,00377			
			$\hat{\beta}_0$	-0,78775	0,05234			
Brody	Gamma	Log	$\hat{\beta}_1$	8,08019	0,04080	0,9807	12,0925	0,0049
			$\hat{\beta}_2$	0,07663	0,00228			
			$\hat{\beta}_3$	0,03918	0,00094			

<sup>ns</sup> não significativo ao nível de 5% de significância. REQM em  $\text{m}^3 \text{árvore}^{-1}$ . Link representa a função de ligação entre o preditor não linear e a variável resposta.

Em relação as estatísticas de qualidade do ajuste o modelo de Schumacher e Hall com distribuição *Gamma* e função de ligação *Log* gerou a melhor equação com  $IA_c$  de 0,9832 e  $S_{xy}\%$  de 11,3071. Entretanto, com base na análise dos resíduos dada na Figura 12 e por meio do gráfico dos resíduos em porcentagem foi constatado que a variância dos erros não é homogênea.

Figura 12 – Distribuição dos resíduos para as equações geradas pelos modelos não lineares generalizados.



Mesmo com a modificação  $E(\epsilon_i^2) = \sigma^2 \cdot 1/(DAP^2 \cdot Ht)$  na matriz de variância os resíduos não tomaram a forma homocedástica indicando assim um alto grau de heterocedasticidade. Quanto ao pressuposto da independência dos erros todas as equações apresentaram resíduos independentes conforme é verificado pelo gráfico de autocorrelação, bem como também foi constatado pelo teste de Durbin - Watson (Dw) com  $Dw \approx 1,99$  com  $\alpha$  de 0,05. Considerando a aderência das equações à distribuição sugerida é evidenciado um certo grau de afastamento. Sendo assim, verifica-se que mesmo permitindo a variável resposta assumir qualquer distribuição contínua da família exponencial não se constatou aderência.

A equação obtida por meio do modelo de Silva e Bailey apresentou um desempenho inferior às demais, possivelmente isso pode ser explicado pela não significância do coeficiente ( $\hat{\beta}_0$ ) o que pode ter tornado a equação menos precisa. Bem como, a mesma apresentou o maior grau de afastamento dos pressupostos da regressão.

Os modelos não lineares generalizados foram pouco empregados na estimativa de volume em plantações florestais. No entanto, foi constatado que os modelos de Schumacher e Hall, Spurr e Brody ajustados com uma estrutura generalizada produziram equações com boa precisão.

## 5.5.2 Algoritmos de aprendizagem de máquinas

### 5.5.2.1 Redes Neurais Artificiais

Foram treinadas 300 redes neurais artificiais considerando as combinações entre número de neurônios na camada oculta, função de ativação, número de épocas, taxa de aprendizagem, L1, L2, taxa de dropout e momentum. Os hiperparâmetros de treinamento da rede selecionada estão dispostos na Tabela 10.

Tabela 10 – Hiperparâmetros de treinamento da RNA para estimar a produção dos clones de *Eucalyptus* no Polo Gesseiro.

Hiperparâmetros	RNA
Dropout	0,0
L1	1e-06
L2	0,0
Épocas	500
Taxa de Aprendizagem	0,05
Momentum	0,1

As configurações e estatísticas de qualidade para a rede neural artificial treinada para estimar o volume de clones de *Eucalyptus* no Polo Gesseiro do Araripe-PE são apresentadas na Tabela 11. Mesmo penalizando a quantidade de parâmetros presentes na rede neural as estatísticas de qualidade  $IA_c$  (0,9837) e  $S_{xy}\%$  (11,1040) ainda são considerados precisos.

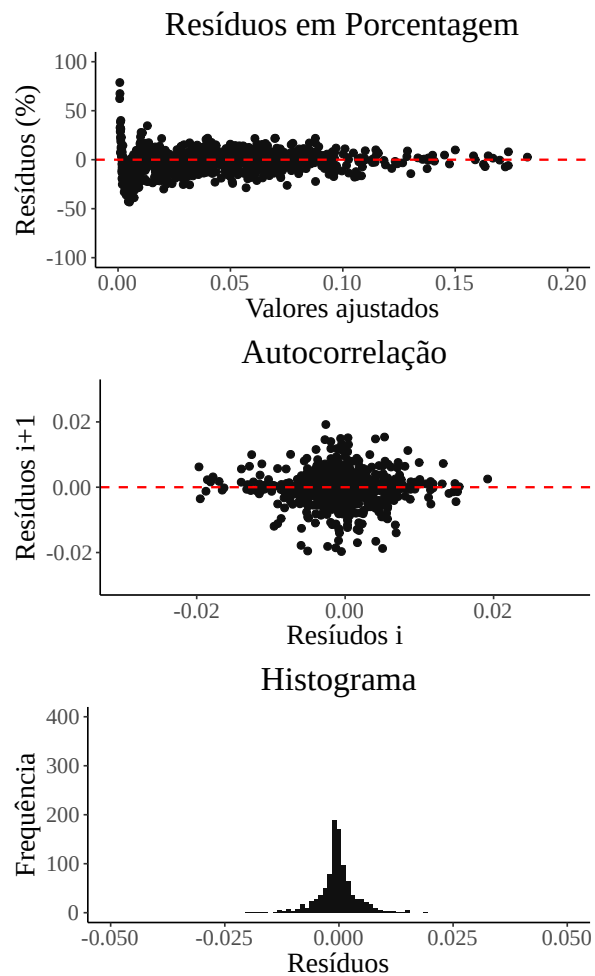
Tabela 11 – Configurações e estatística de qualidade para a RNA do tipo Perceptron de Múltiplas Camadas treinada para estimar o volume individual por  $m^3$  dos clones de *Eucalyptus* cultivados no Polo Gesseiro do Araripe.

N.	N. de neurônios - Função de ativação			Validação			Teste	
	Camada de entrada	Camada intermediária	Camada de saída	$IA_c$	$S_{xy}\%$	REQM	EAM	REQM
RNA	2: DAP, Ht	7: Maxout	1: Linear	0,9837	11,1040	0,0044	0,0030	0,0043

REQM e EAM em  $m^3$  árvore<sup>-1</sup>.

Por meio dos gráficos de resíduos analisados na Figura 13, nota-se que a RNA

Figura 13 – Análise dos resíduos das redes neurais artificiais treinadas para estimar a produção em volume dos clones de *Eucalyptus*.



apresentou alguns pontos de subestimativa para volumes pequenos, de forma geral a rede neural não apresenta *outliers* e com base nas estatísticas de qualidade de ajuste possui alto poder de generalização.

Outros estudos com as técnicas de Redes Neurais Artificiais já foram desenvolvidos nas ciências florestais, porém é difícil encontrar estudos que levam em consideração a quantidade de parâmetros na comparação com os modelos de regressão clássicos da estatística. Souza et al. (2018) realizaram o primeiro estudo com redes neurais para estimar o volume de madeira em clones de *Eucalyptus* no espaçamento 3 m x 2 m plantados no Polo Gesseiro do Araripe. Neste trabalho os autores chegaram a uma rede do tipo MLP com a topologia 2/5/1 usando a função Tangente hiperbólica na camada intermediária e Identidade na saída. Com esta configuração obtiveram um erro padrão de estimativa de 11,24014%, porém no cálculo do teste não foi incluído o número de parâmetros (*pesos sinápticos*) da rede. Mesmo assim, esse é um importante resultado para a estimativa do volume de madeira por árvore individual. Com os resultados obtidos com a RNA (erro padrão de estimativa  $S_{xy}$ % de 11,1040) é destacado que mesmo considerando a grande quantidade de parâmetros estimados, as RNA's produzem estimativas com alta precisão e poder de generalização.

Em um trabalho realizado para estimativa do volume de *Pinus taeda* na região central do Paraná, Montaño (2016) chegou a boas estimativas (erro padrão de estimativa de 12,5573%) utilizando redes neurais artificiais. Resultados semelhantes foram obtidos por Schikowski et al. (2018) trabalhando com *Acacia mearnsii* De Wild. em plantações localizadas no estado do Rio Grande do Sul. Os autores utilizaram redes neurais para estimativa do volume de madeira por árvore e obtiveram um erro de 0,006 m<sup>3</sup> (REQM). Os resultados obtidos nesta pesquisa, são superiores aos encontrados na literatura, isto pode está relacionado ao maior número de hiperparâmetros ajustados, por exemplo os parâmetros de regularização L1 e L2 que tornam a rede mais estável, ou simplesmente à variações na base de dados.

### 5.5.2.2 Florestas Aleatórias

Com o algoritmo de Florestas Aleatórias (*Random Forest*) foram treinados 12 modelos. Os hiperparâmetros, para o melhor modelo estão disponíveis na Tabela 12.



Tabela 12 – Hiperparâmetros de treinamento para Florestas Aleatórias - FA para estimar o volume dos clones de *Eucalyptus* no Polo Gesseiro.

Hiperparâmetros	FA
Número de árvores	1000
Número de atributos	1
Profundidade média	10

O hiperparâmetro número de árvores selecionado foi 1000, segundo Hastie, Tibshirani e Friedman (2009) e Genuer (2012) um maior número de árvores permite a diminuição da variância e a estabilização do bias, por isso, existe a preferência por modelos com o maior número de árvores. O modelo de floresta aleatória foi selecionado com apenas uma das variáveis preditoras em cada *nó*, pois esse, permitiu melhor generalização e ainda estima o volume utilizando as variáveis DAP e Ht quase que na mesma proporção. Dessa forma, evita-se erros de estimativas quando se tem erros de medida, ou seja, quando existe árvores com grande DAP e baixa Ht ou vice-versa a estimativa não prejudicada.

As estatísticas da qualidade de ajuste para o modelo de Floresta Aleatória são dadas na Tabela 13. Não foi possível calcular o  $la_c$  e  $S_{xy}\%$ , devido a metodologia de Florestas Aleatórias não possuir coeficientes como às demais técnicas. Nesta técnica a estimativa é realizada com base em regras, estas são influenciadas pelo número de árvores criadas e pela profundidade da árvore, sendo assim, cada estimativa pode ser obtida com base em centenas ou milhares de regras.

Tabela 13 – Configurações e estatística de qualidade para a técnica de Florestas Aleatórias - FA para estimar a produção dos clones de *Eucalyptus* no Polo Gesseiro.

N.	Variáveis preditivas	Validação		Teste	
		EAM	REQM	EAM	REQM
FA	DAP; Ht	0,0015	0,0023	0,0033	0,0051

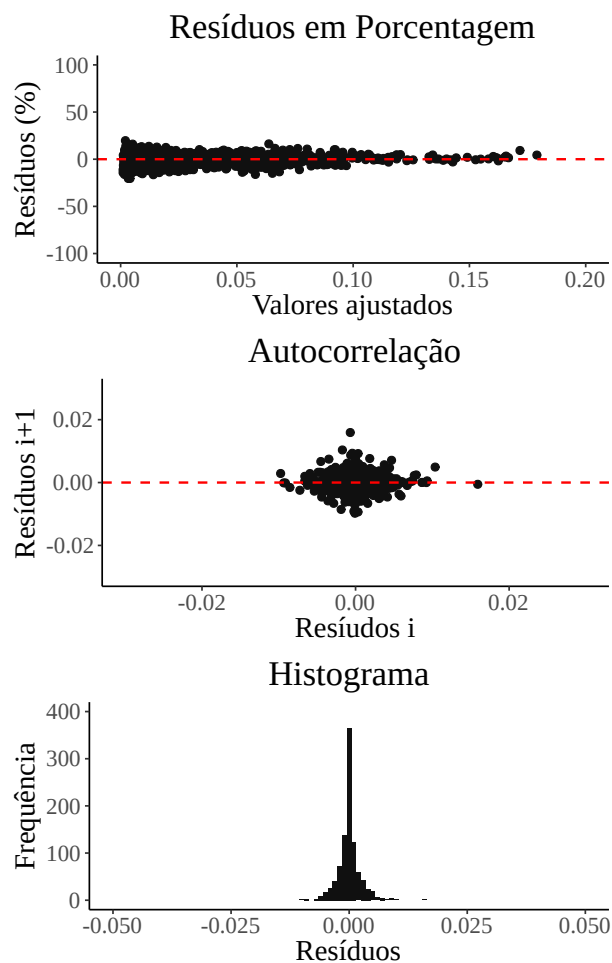
REQM e EMA em  $m^3$  árvore<sup>-1</sup>.

Trabalhos com Florestas Aleatórias são mais escassos para o problema de estimativa do volume. O melhor modelo treinado por Schikowski et al. (2018) para estimar o volume por árvore de *Acacia mearnsii* produziu um erro médio de  $0,009 m^3$  (REQM), o que equivale a 11,74% quando se considerou a REQM em porcentagem. A estimativa

do volume florestal por meio da técnica de Florestas Aleatórias foi considerada satisfatória e superior às demais técnicas (variações da técnica k-NN) em estudo realizado por Latifi, Nothdurft e Koch (2010). Os autores citados desenvolveram a pesquisa em uma área florestal localizada no norte de Karlsruhe, estado de Baden-Württemberg (Alemanha), e para estimar o volume por hectare utilizaram dados geoespaciais.

O modelo FA não apresentou pontos de superestimativas ou subestimativas, ficando os resíduos entre os valores de 30% positivo e 30% negativo (Figura 14). Pelo gráfico do histograma se nota que os resíduos possuem uma distribuição simétrica do tipo leptocúrtica com caudas longas. Um comportamento similar para a distribuição dos resíduos foi obtido por (GÖRGENS; MONTAGHI; RODRIGUEZ, 2015).

Figura 14 – Análise dos resíduos para os modelos de Florestas Aleatórias - FA (*Random Forest*) para estimar a produção em volume dos clones de *Eucalyptus*.



### 5.5.2.3 Máquinas de Vetores de Suporte

Considerando o algoritmo de Máquinas de Vetores de Suporte - MVS (*Support Vector Machine*) foram treinados 64 modelos. Os hiperparâmetros do modelo selecionado estão na Tabela 14. O valor de  $\epsilon$ , que está associado com o erro, afeta consideravelmente o número de vetores de suporte do modelo. Conseqüentemente os valores de  $l_{ac}$  e  $S_{xy}\%$  serão diretamente afetados, já que são influenciados pela quantidade de coeficientes (vetores de suporte).

Tabela 14 – Hiperparâmetros de treinamento para a técnica de Máquinas de Vetores de Suporte - MVS (*Support Vector Machine*) para estimar a produção dos clones de *Eucalyptus* no Polo Gesseiro.

Hiperparâmetros	MVS
Kernel	<i>radial</i>
Custo	90,5096
Gamma	0,2
Epsilon	0,09

Na Tabela 15 estão disponíveis as estatísticas de teste para o conjunto de validação e de teste para o modelo de Máquinas de Vetores de Suporte selecionado.

Tabela 15 – Configurações e estatística de qualidade para a técnica de Máquinas de Vetores de Suporte - MVS (*Support Vector Machine*) para estimar a produção dos clones de *Eucalyptus* no Polo Gesseiro.

N.	Variáveis preditivas	Validação			Teste	
		REQM	$l_{ac}$	$S_{xy}\%$	EAM	REQM
MVS	DAP; Ht	0,0044	0,9780	12,9196	0,0030	0,0044

REQM e EMA em  $m^3$  árvore<sup>-1</sup>.

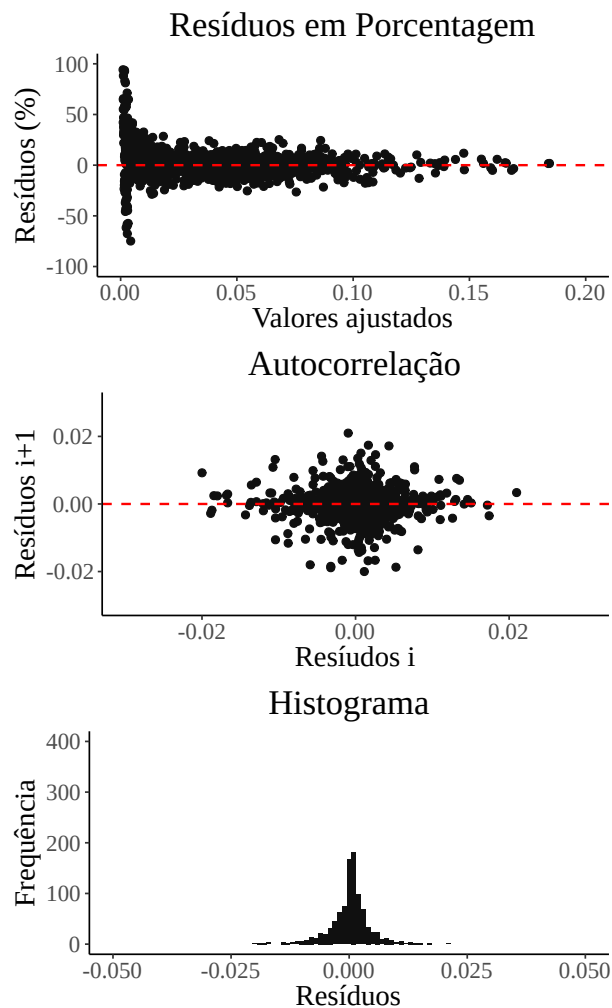
Para os testes desenvolvidos a utilização da função de *kernel* radial gerou estatísticas de qualidade semelhante a polinomial, porém houve uma menor dispersão dos resíduos quando se empregou a função de *kernel* radial. Binoti et al. (2016) também chegaram a estimativas de volume precisas quando treinaram uma MVS utilizando a função de *kernel* radial. Já Abreu (2019), obteve boas estimativas tanto para a radial quanto para a polinomial, dependendo do tipo florestal que foi estudado. Em geral o *kernel* radial tende a ser superior ao polinomial em problemas não lineares. Se

o fenômeno estudado apresenta um alto grau de não linearidade será necessário aumentar a ordem do polinômio em uma função de kernel polinomial, isto se torna um problema, pois em polinômios de alto grau a derivada de primeira ordem é constante e a segunda derivada é zero. Já para função do Kernel radial, é possível obter as infinitas derivadas, não importando assim, a complexidade do fenômeno.

O modelo MVS apresentou baixos valores de REQM e EAM quando empregado no conjunto de teste. Assim, tem-se que o modelo se ajustou bem ao problema estudado. Cordeiro et al. (2015), quando treinaram uma MVS, obtiveram boas estimativas do volume com e sem casca para espécie *Acacia mangium* com coeficiente de correlação de 0,9857 e 0,9832, respectivamente.

Com o histograma da Figura 15 se nota que os resíduos apresentam maior distribuição em torno de zero e os resíduos porcentagem não apresentam nenhum valor discrepante.

Figura 15 – Análise dos resíduos para algoritmos de Máquinas de Vetores de Suporte - MVS para estimar a produção em volume dos clones de *Eucalyptus*.



Os algoritmos de aprendizagem de máquinas se tornam uma opção para a modelagem biométrica, principalmente quando problemas com homogeneidade e falta de normalidade dos resíduos estão presentes. Segundo Özçelik et al. (2010), os modelos de RNA são capazes de aprender com a experiência, sem a necessidade de atender suposições em relação a forma da função de ajuste. Por causa destas vantagens tem crescido o interesse em algoritmos de aprendizagem de máquinas (ÖZÇELİK et al., 2017).

### 5.5.3 Comparação entre os modelos

A performance das equações obtidas por meio da regressão não linear e dos modelos obtidos via algoritmos de aprendizagem de máquinas foi avaliada no conjunto de dados de teste. Para uma maior generalização foram realizadas 100 estimativas para cada equação/modelo considerando uma reamostragem via *Bootstrap* com amostras de 200 indivíduos com reposição a partir do conjunto de dados de teste. Os valores médios das estatísticas de qualidade são apresentados na Tabela 16. Bem como, é dado as medidas do ranqueamento de cada equação/modelo.

O modelo obtido com o algoritmo de redes neurais artificiais e a equação proveniente do modelo de Schumacher com distribuição simétrica dos resíduos apresentaram resultados semelhantes e se destacaram em relação às demais. O modelo oriundo do algoritmo de máquina de vetor de suporte e a equação obtida do ajuste do modelo de Schumacher apresentam resultados próximo aos modelos citados anteriormente e também não diferem entre si.

É importante destacar que todas as estimativas realizadas com as equações/modelos para o conjunto de dados de teste apresentaram resultados próximos aos encontrados no conjunto de treino, o que evidencia a qualidade de ajuste e ausência de *overfitting*.

Tabela 16 – Resumo estatístico das melhores metodologias empregadas na estimativa da produção de clones de *Eucalyptus* no Polo Gesseiro do Araripe-PE.

<b>Modelos Estatísticos</b>	<b>IA</b>	<b>REQM</b>	<b>EAM</b>	<b>R(IA)</b>	<b>R(REQM)</b>	<b>R(EAM)</b>	$\sum_{i=1}^3 R_i$
Schumacher e Hall	0,98441	0,00443	0,00293	4	4	2	10
Spurr	0,98243	0,00470	0,00302	12	12	5	29
Silva e Bailey	0,98369	0,00453	0,00307	9	9	9	27
Brody	0,98385	0,00451	0,00311	8	8	12	28
Schumacher e Hall Simétrico	0,98499	0,00435	0,00293	2	2	1	5
Spurr Simétrico	0,98280	0,00466	0,00302	6	6	11	23
Silva e Bailey Simétrico	0,98396	0,00449	0,00309	10	10	7	27
Brody Simétrico	0,98420	0,00446	0,00308	5	5	10	20
Schumacher e Hall Generalizado	0,98392	0,00450	0,00305	7	7	8	22
Spurr Generalizado	0,98277	0,00466	0,00302	15	15	15	45
Silva e Bailey Generalizado	0,95432	0,00761	0,00465	11	11	6	28
Brody Generalizado	0,98097	0,00491	0,00317	13	13	13	39
<b>Aprendizagem de Máquinas</b>	<b>IA</b>	<b>REQM</b>	<b>EAM</b>	<b>R(IA)</b>	<b>R(REQM)</b>	<b>R(EAM)</b>	$\sum_{i=1}^3 R_i$
Redes Neurais Artificiais	0,98550	0,00427	0,00297	1	1	3	5
Florestas Aleatórias	0,98016	0,00501	0,00328	14	14	14	42
Máquinas de Vetor de Suporte	0,98498	0,00435	0,00298	3	3	4	10

R = Ranque, referente ao posicionamento dos modelos; REQM e EAM em m<sup>3</sup> árvore<sup>-1</sup>.

Como os modelos apresentaram resultados bastante semelhantes na média das estatísticas de qualidade no conjunto de dados de teste foi empregado a análise de variância para a comparação das estimativas incluindo o valor real do volume. Sendo assim, foram realizadas análises de variância para cada uma das 100 amostras geradas, em que as equações/modelos e o volume real são considerados o tratamento.

As análises de variância não mostram diferenças significativas (Valor p > 0,05) entre as estimativas realizadas por cada modelo/equação, bem como entre eles e a variável resposta volume. Desta forma, mesma na presença de resíduos heterocedásti-

cos e desvios quanto a normalidade as equações obtidas via modelos de regressão possuem ajuste precisos, bem como os algoritmos de aprendizagem de máquinas também possuem estimativas do volume estatisticamente semelhantes ao volume real.

## 5.6 Conclusões

Os modelos não lineares de regressão com distribuição simétrica dos resíduos geram equações com boa precisão para a estimativa do volume de clones de *Eucalyptus*.

Os modelos não lineares de regressão com estrutura generalizadas geram equações com boa precisão para a estimativa do volume de clones de *Eucalyptus*.

É possível obter estimativas do volume semelhantes ao volume real com os seguintes algoritmos de aprendizagem de máquinas: Redes Neurais Artificiais, Florestas Aleatórias e Máquinas de Vetor de Suporte.

Não foi constatada diferenças significativas no conjunto de dados de teste entre as estimativas do volume realizadas por meio dos modelos de regressão não linear, regressão não linear com erros simétricos, regressão não linear com estrutura generalizada, redes neurais artificiais, florestas aleatórias, máquinas de vetores de suporte e o volume de madeira real.

## Referências Bibliográficas

- ABREU, J. C. d. **Alternativas para estimar o volume de árvores individuais em formações florestais no estado de Minas Gerais**. Tese (Ciência Florestal) — Universidade Federal de Viçosa, Viçosa, MG, 2019.
- AZEVEDO, T. L. et al. Equações hipsométricas e volumétricas para um povoamento de *Eucalyptus* sp. localizado na FLONA do Ibura, Sergipe. **Revista Brasileira de Ciências Agrárias**, Universidade Federal Rural de Pernambuco (UFRPE), v. 6, n. 1, p. 105–112, 2011.
- BATES, D. M.; WATTS, D. G. **Nonlinear regression analysis and its applications**. New York, NY: Wiley Online Library, 1988. 365 p.
- BINOTI, D. H. B. et al. Support vector machine to estimate volume of eucalypt trees. **Revista Árvore**, v. 40, n. 4, p. 689–693, 2016.
- BINOTI, M. L. M. d. S. et al. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, n. 1, p. 97–105, 2015.
- BOX, G. E.; COX, D. R. An analysis of transformations. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 26, n. 2, p. 211–243, 1964.
- BREUSCH, T. S.; PAGAN, A. R. A simple test for heteroscedasticity and random coefficient variation. **Econometrica: Journal of the Econometric Society**, p. 1287–1294, 1979.
- CALEGARIO, N. et al. Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não lineares generalizados. **Scientia Forestalis, Piracicaba**, v. 69, p. 38–50, 2005.
- CORDEIRO, M. A. et al. Estimativa do volume de acácia mangium utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. **Pesquisa Florestal Brasileira**, v. 35, n. 83, p. 255–261, 2015.
- COSENZA, D. N. et al. Classificação da capacidade produtiva de sítios florestais utilizando máquina de vetor de suporte e rede neural artificial. **Scientia Forestalis**, v. 43, n. 108, p. 955–963, 2015.



CYSNEIROS, F. J. A.; PAULA, G. A.; GALEA, M. **Modelos simétricos aplicados**. São Paulo: Associação Brasileira de Estatística, 2005. v. 1. 88 p.

DRUCKER, H. et al. Support vector regression machines. In: JORDAN, M. I.; KEARNS, M. J.; SOLLA, S. A. (Ed.). **Advances in neural information processing systems**. Denver, Colorado, USA: MIT Press, 1997. p. 155–161.

FABRIKA, M.; VALENT, P.; MERGANIČOVÁ, K. Forest modelling and visualisation—state of the art and perspectives. **Central European Forestry Journal**, v. 66, n. 3-4, p. 147–165, 2019.

FILHO, L. M. A. L. et al. Modelagem do crescimento de clones de eucalyptus usando o modelo de chapman-richards com diferentes distribuições simétricas dos erros. **Ciência Florestal**, v. 22, n. 4, p. 777–785, 2012.

FURNIVAL, G. M. An index for comparing equations used in constructing volume tables. **Forest Science**, v. 7, n. 4, p. 337–341, 1961.

GADELHA, F. H. L. **Desempenho silvicultural e avaliação econômica de clones híbridos de eucaliptos plantados em diferentes regimes de manejo para fins energéticos**. 147 p. Tese (Ciências Florestais) — Universidade Federal Rural de Pernambuco, Departamento de Ciências Florestais, Recife, 2014.

GEDEON, T. D. Data mining of inputs: analysing magnitude and functional measures. **International Journal of Neural Systems**, v. 8, n. 02, p. 209–218, 1997.

GENUER, R. Variance reduction in purely random forests. **Journal of Nonparametric Statistics**, v. 24, n. 3, p. 543–562, 2012.

GENUER, R.; POGGI, J.-M.; TULEAU-MALOT, C. Vsurf: An r package for variable selection using random forests. **The R Journal**, v. 7, n. 2, p. 19–33, 2015.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. London, England: MIT press, 2016. 800 p.

GÖRGENS, E. B.; MONTAGHI, A.; RODRIGUEZ, L. C. E. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. **Computers and Electronics in Agriculture**, v. 116, p. 221–227, 2015.

GUERA, O. G. M. **Modelos matemáticos para auxílio à tomada de decisão no processo produtivo de *Pinus caribaea* Morelet var. *caribaea* Barr. & Golf**. na

**empresa florestal integral Macurije, Pinar Del Río, Cuba.** Tese (Ciências Florestais) — Universidade Federal Rural de Pernambuco, Recife, PE, 2017.

GUERA, O. G. M. et al. Alternatives to growth and yield prognosis for *Pinus caribaea* var. *caribaea* barrett & golfari. **Floresta e Ambiente**, v. 26, n. 4, 2019b.

GUISAN, A.; JR, T. C. E.; HASTIE, T. Generalized linear and generalized additive models in studies of species distributions: setting the scene. **Ecological modelling**, v. 157, n. 2-3, p. 89–100, 2002.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning**. 2. ed. New York, NY: Springer series in statistics New York, 2009. (Springer Series in Statistics).

HESS, A. F. et al. Aplicação dos modelos lineares generalizados para estimativa do crescimento em altura. **Pesquisa Florestal Brasileira**, v. 35, n. 84, p. 427–433, 2015.

LATIFI, H.; NOTHDURFT, A.; KOCH, B. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/lidar-derived predictors. **Forestry**, v. 83, n. 4, p. 395–407, 2010.

LEDELL, E. et al. **h2o: R Interface for 'H2O'**. [S.I.], 2019. R package version 3.24.0.5. Disponível em: <<https://CRAN.R-project.org/package=h2o>>.

LORENA, A. C.; CARVALHO, A. C. de. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 2007.

MELO, L. C. et al. Estimativas volumétricas em povoamento de *pinus caribaea* var. *hondurensis* no sudoeste da bahia. **Pesquisa Florestal Brasileira**, v. 33, n. 76, p. 379–386, 2013.

MENDIBURU, F. d. **agricolae: Statistical Procedures for Agricultural Research**. [S.I.], 2017. R package version 1.2-8. Disponível em: <<https://CRAN.R-project.org/package=agricolae>>.

MENDONÇA, A. R. d. et al. Growth and yield models for eucalyptus stands obtained by differential equations. **Scientia Agricola**, v. 74, n. 5, p. 364–370, 2017.

MEYER, D. et al. **e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien**. [S.I.], 2017. R package version 1.6-8. Disponível em: <<https://CRAN.R-project.org/package=e1071>>.

MEYER, H. A. The standard error of estimate of tree volume from the logarithmic volume equation. **Journal Forestry**, v. 36, p. 340–342, 1943.

MONTAÑO, R. A. N. R. **Aplicação de técnicas de aprendizado de máquina na mensuração florestal**. 102 p. Tese (Informática) — Universidade Federal do Paraná, Curitiba, PR, 2016.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**. 3. ed. New Delhi: John Wiley e Sons, 2013. 641 p.

ÖZÇELİK, R. et al. Estimating tree bole volume using artificial neural network models for four species in Turkey. **Journal of environmental management**, v. 91, n. 3, p. 742–753, 2010.

ÖZÇELİK, R. et al. Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. **Forest Science**, v. 63, n. 3, p. 291–302, 2017.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Disponível em: <<https://www.R-project.org/>>.

REIS, A. A. et al. Spatial prediction of basal area and volume in *Eucalyptus* stands using landsat tm data: an assessment of prediction methods. **New Zealand Journal of Forestry Science**, v. 48, n. 1, p. 1, 2018.

REIS, A. A. et al. Volume estimation in a *Eucalyptus* plantation using multi-source remote sensing and digital terrain data: a case study in minas gerais state, brazil. **International journal of remote sensing**, Taylor & Francis, v. 40, n. 7, p. 2683–2702, 2019.

SALES, F. d. C. V. et al. Ajustes de modelos volumétricos para o clone eucalyptus grandis x e. urophylla cultivados no agreste de pernambuco. **Floresta**, v. 45, n. 4, p. 663–670, 2015.

SANQUETTA, C. R. et al. Volume estimation of *Cryptomeria japonica* logs in southern brazil using artificial intelligence models. **Southern Forests: a Journal of Forest Science**, v. 80, n. 1, p. 29–36, 2018.

SCHIKOWSKI, A. B. et al. Modeling of stem form and volume through machine learning. **Anais da Academia Brasileira de Ciências**, v. 90, n. 4, p. 3389–3401, 2018.

SCHLAEGEL, B. E. Green ash volume and weight tables. **Res. Pap.**, SO–206, p. 1–14, 1984.

SCHNEIDER, R. et al. Growth and yield models for predicting tree and stand productivity. In: **Ecological Forest Management Handbook**. [S.l.]: ROUTLEDGE in association with GSE Research, 2016. v. 141, n. 178, p. 141–178.

SHIFLEY, S. R. et al. The past and future of modeling forest dynamics: from growth and yield curves to forest landscape models. **Landscape Ecology**, v. 32, n. 7, p. 1307–1325, 2017.

SILVA, J. A. A.; BAILEY, R. L. Considerações teóricas sobre o uso correto do índice de furnival na seleção de equações volumétricas. **Revista Árvore**, v. 15, n. 3, p. 323–327, 1991.

SILVA, J. A. A.; NETO, F. d. P. **Princípios básicos de dendrometria**. [S.l.]: Imprensa Universitária da UFRPE, 1979. 198 p.

SILVA, J. A. A. et al. Modelagem do crescimento volumétrico de clones de eucalipto (*Eucalyptus urophylla*) no Polo Gesseiro do Araripe-PE. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 13/14, p. 173–190, 2016/2017.

SILVA, J. P. M. **Prognose da produção florestal utilizando sistema Neuro-Fuzzy e Random Forest**. 69 p. Dissertação (Ciências Florestais) — Universidade Federal do Espírito Santo, Jerônimo Monteiro, ES, 2018.

SILVA, J. W. L. **Modelagem da biomassa e da quantidade de carbono de clones de *Eucalyptus* da Chapada do Araripe-PE**. 105 p. Dissertação (Biometria e Estatística Aplicada) — Universidade Federal Rural de Pernambuco, Recife, PE, 2016.

SOUZA, S. R. R. et al. Redes neurais para estimativa volumétrica de clones de *Eucalyptus* spp. no Polo Gesseiro do Araripe. **Revista Brasileira de Biometria**, v. 36, n. 3, p. 715–729, 2018.

SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. **The journal of machine learning research**, v. 15, n. 1, p. 1929–1958, 2014.

TURNER, H.; FIRTH, D. **Generalized nonlinear models in R: An overview of the gnm package**. [S.l.], 2015. R package version 1.0-8. Disponível em: <<http://CRAN.R-project.org/package=gnm>>.

VALENÇA, M. J. **Fundamentos das redes neurais: exemplos em Java**. Recife: Livro Rápido, 2010. 386 p.

VAPNIK, V. N. **Statistical learning theory**. New York: Wiley, 1998. v. 1. 736 p.

- VIEIRA, G. C. et al. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. **Science of the Total Environment**, v. 619, p. 1473–1481, 2018.
- WEBER, V. P. et al. Modelagem linear generalizada para descrever o incremento em área transversal de árvores individuais de imbuia. **FLORESTA**, v. 48, n. 1, p. 123–132, 2018.
- WHITE, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. **Econometrica: Journal of the Econometric Society**, p. 817–838, 1980.

## CAPÍTULO 6

---

### **Análise do crescimento de clones de *Eucalyptus* spp. em período climático chuvoso e de seca extrema**

#### **Resumo**

Atualmente existe uma grande preocupação no meio acadêmico sobre como as florestas são impactadas pela mudança climática. O aumento de CO<sub>2</sub> na atmosfera está diretamente relacionado com a mudança climática, e estas, podem afetar negativamente o crescimento das plantas. Sendo assim, é importante estudar o efeito das mudanças climáticas na produtividade das florestas. Desta forma, objetivou-se com esta pesquisa analisar se existe mudança de tendências nas séries climáticas da região do Araripe-PE e avaliar se existe diferenças no crescimento dos clones de *Eucalyptus* em dois cenários climáticos distintos. Para isto, foram instalados dois experimentos na Estação Experimental do Instituto Agrônomo de Pernambuco (IPA) no município de Araripina. O primeiro experimento foi desenvolvido entre os anos de 2002 e 2009 com três clones de *Eucalyptus* (C11, C39 e C41) no espaçamento 3 m x 2 m, conduzido em período com chuvas acima da normal histórica. O segundo experimento foi estabelecido entre os anos de 2010 e 2018, com os mesmos clones no mesmo arranjo espacial, no entanto, as chuvas neste período se concentraram abaixo da normal histórica. Foram coletados dados dendrométricos ao longo de todo período experimental. Os dados climáticos de precipitação e temperatura foram obtidos de uma estação meteorológica localizada próxima aos experimentos. Para análise da mudança de tendência nas séries climáticas se utilizou o teste de Mann-Kendall, regressão linear, o teste de Pettit e a metodologia *change point*. Em seguida foi realizada uma análise descritiva das variáveis climáticas nos dois períodos experimentais. Para avaliar a diferença no crescimento em volume de ambos os experimentos foram obtidas equações a partir do modelo de Chapman-Richards e posteriormente foi analisado a identidade das equações. Foi detectado que existe uma tendência de redução na precipitação e de aumento para a temperatura média do ar. Os dois períodos experimentais apresentam diferenças significativas nas variáveis climáticas. Por meio das equações obtidas foi possível identificar diferenças significativas, no crescimento em volume, entre os dois experimentos.

**Palavras-chave:** Polo Gesseiro do Araripe-PE, Mudanças Climáticas, semiárido, produtividade.

## **Abstract**

Currently, there is a great concern in academia about how forests are impacted by climate change. The increase in CO<sub>2</sub> in the atmosphere is directly related to climate change, which can negatively affect plant growth. Therefore, it is important to study the effect of climate change on forest productivity. Thus, the objective of this research was to analyze if there is a change in trends in the climatic series of the Araripe-PE region and to evaluate if there are differences in the growth of *Eucalyptus* clones in two different climatic scenarios. For this, two experiments were installed at the Experimental Station of the Agronomic Institute of Pernambuco (IPA) in the municipality of Araripina. The first experiment was carried out between the years 2002 and 2009 with three clones of *Eucalyptus* (C11, C39 and C41) in the 3 m x 2 m spacing, conducted in a period with rainfall above the historical normal. The second experiment was established in the years 2010 and 2018, with the same clones in the same spatial arrangement, however, the rains in this period were concentrated below the historical normal. Dendrometric data were collected throughout the experimental period. The climatic data of precipitation and temperature were obtained from a meteorological station located next to the experiments. The Mann-Kendall test, linear regression, the Pettit test and the *change point* methodology were used to analyze the trend change in the climate series. Then, a descriptive analysis of the climatic variables was performed in the two experimental periods. To assess the difference in volume growth of both experiments, equations were obtained from the Chapman-Richards model and later the identity of the equations was analyzed. It was detected that there is a tendency towards a reduction in precipitation and an increase in mean air temperature. The two experimental periods show significant differences in climatic variables. Through the obtained equations it was possible to identify significant differences, in volume growth, between the two experiments.

**Key-words:** Gypsum Pole of Araripe-PE, Climate Changes, semiarid, productivity.

## **6.1 Introdução**

O aumento da concentração de CO<sub>2</sub> na atmosfera, a mudança na decomposição do nitrogênio e o desenvolvimento de novas tecnologias de manejo florestal são os principais fatores que causam mudanças na produtividade das florestas em muitas

partes do mundo (REYER, 2015). A elevação de CO<sub>2</sub> na atmosfera está diretamente relacionada com a mudança climática, esta, pode afetar negativamente o crescimento das plantas, acarretando principalmente mudanças fisiológicas, tanto modificando sua taxa de respiração quanto alterando seu ciclo de desenvolvimento. É de fundamental importância para serviços de manejo florestal determinar como a produtividade das florestas será afetada em cenários futuros de mudanças no clima (REYER, 2015).

Atualmente, existe uma grande preocupação no meio acadêmico sobre como as florestas são impactadas pela mudança climática. Diversos estudos estão sendo desenvolvidos com o objetivo de entender o impacto desta mudança principalmente sobre a produtividade das florestas (KELLOMÄKI et al., 2008; WAMELINK et al., 2009; REYER et al., 2017; SANTOS et al., 2017a). Dependendo da localidade algumas espécies podem apresentar efeitos positivos com a mudança no clima, entretanto, muitas pesquisas projetam que provavelmente estes efeitos acarretem consequências negativas para as florestas (KEENAN, 2015).

A implantação de florestas de rápido crescimento no Nordeste brasileiro tem aumentado ao longo dos anos, principalmente no estado da Bahia (SANTOS et al., 2017a). Entretanto, no semiárido nordestino a concentração desses plantios ainda é escassa, sendo Pernambuco, desde os anos 2002, pioneiro em estudos com clones resistentes ao estresse hídrico (SILVA et al., 2013). No Polo Gesseiro do Araripe o principal objetivo destes plantios é produzir madeira para atender a demanda energética das indústrias do gesso localizadas na região (SILVA et al., 2016/2017; SILVA et al., 2018). O Polo Gesseiro possui uma demanda energética anual estimada em 652.680 m<sup>3</sup> de lenha, sendo esta proveniente majoritariamente da vegetação Caatinga (GADELHA; SILVA; FERREIRA, 2018).

Com a intensificação da mudança climática a região semiárida do Nordeste brasileiro sofrerá com a ocorrência de eventos extremos do clima, principalmente com longos períodos de seca (LACERDA et al., 2014/2015). Considerando o semiárido nordestino, até o ano de 2040, é provável que exista uma redução de até 20% na precipitação e um aumento de até 1°C na temperatura do ar, segundo o "Relatório do Painel Brasileiro de Mudanças Climáticas"(PBMC, 2013). A distribuição de recursos hídricos e nutrientes influenciam fortemente a produção de madeira em florestas plantadas (RYAN et al., 2010). Desta forma, torna-se importante o estudo da tendência do crescimento de clones de *Eucalyptus* sob diferentes períodos climáticos, principalmente para se estabelecer o conhecimento necessário sobre a adaptação e práticas de manejo destas florestas em condições extremas do clima (SANTOS et al., 2017a).

Diante do provável aumento nos casos de eventos extremos em regiões semiá-



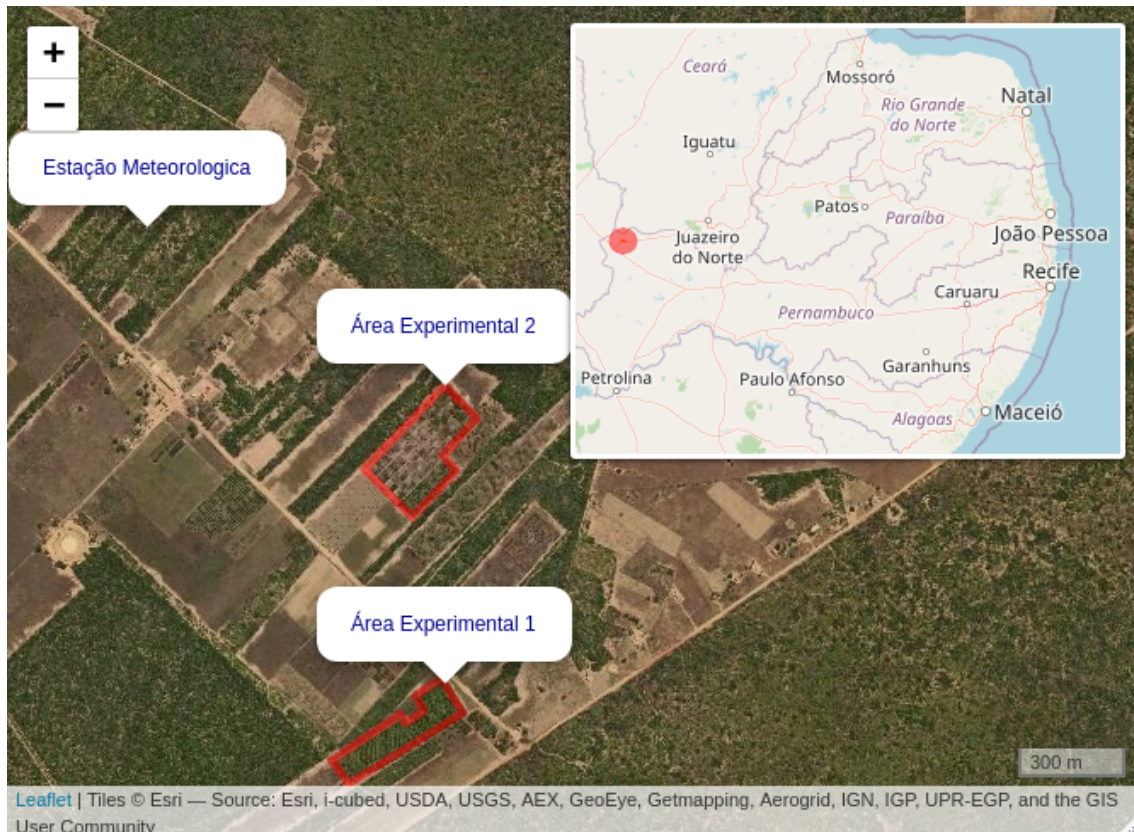
ridas, objetivou-se com esta pesquisa analisar se existe mudança na tendência das séries climáticas da região do Araripe-PE e avaliar se existe diferenças no crescimento dos clones de *Eucalyptus* em um cenário chuvoso e em outro de seca extrema.

## 6.2 Material e Métodos

### 6.2.1 Área experimental e planejamento do experimento

Foram instalados dois experimentos na Estação Experimental do Instituto Agrônomo de Pernambuco (IPA) no município de Araripina (coordenadas 07° 29' 00" S e 40° 36' 00" W, altitude de 816 metros), região semiárida pernambucana (Figura 16). Segundo a classificação de Köppen, o clima da região é do tipo BSh, semiárido quente, caracterizado por escassez de chuvas com bastante irregularidade em sua distribuição, a precipitação média anual é de 740 mm, apresenta índices de elevada evaporação e temperatura média do ar, alta (25 °C). A vegetação característica desse tipo de clima é xerófila (Caatinga). O solo é do tipo LVAd1, Latossolos Vermelho - Amarelo Distróficos (SANTOS et al., 2011).

Figura 16 – Localização geográfica das áreas experimentais e estação meteorológica do estudo conduzido no Instituto Agrônomo de Pernambuco, Araripina-PE.



### 6.2.1.1 Experimento 1

O experimento 1 foi instalado no mês de março de 2002 em uma área de 2,5 hectares. A colheita foi realizada em setembro de 2009, configurando assim um povoamento de 7,5 anos de idade. Nesse experimento foram utilizados 15 clones de *Eucalyptus* (tratamentos) e quatro repetições, com espaçamento 3 m x 2 m. As parcelas possuem tamanho de 14 m x 21 m, com área total de 294 m<sup>2</sup>, composta por 49 árvores, ao eliminar o efeito de bordadura a parcela era composta por 25 árvores (SILVA et al., 2013). O objetivo desse estudo foi avaliar quais os clones mais produtivos para a região semiárida.

### 6.2.1.2 Experimento 2

O experimento 2 foi instalado em março de 2010 em uma área de 7,5 hectares. Os dados foram coletados no mês de março de 2018, configurando um povoamento de 8 anos de idade. O experimento era composto por três clones (dois clones de *Eucalyptus brassiana* e um de *Eucalyptus urophylla*), cinco diferentes espaçamentos (2 m x 1 m, 2 m x 2 m, 3 m x 2 m, 3 m x 3 m e 4 m x 2 m) e quatro repetições, disposto em um delineamento inteiramente casualizado com arranjo fatorial (3 x 5).

Para a comparação do incremento volumétrico (m<sup>3</sup> ha<sup>-1</sup>) nos dois períodos climáticos foram considerados apenas os clones *E. urophylla* (C39 e C41) e o clone *E. brassiana* (C11). As comparações foram realizadas apenas com árvores do espaçamento 3 m x 2 m comum aos dois experimentos.

## 6.2.2 Dendrometria e volume

Para cada experimento, após os 18 meses de idade, foram tomadas medidas de diâmetro à altura do peito (DAP) e altura total (Ht) em intervalos de 6 meses até o fim do experimento. Bem como, para o experimento 1, aos 42 meses de idade, foi realizada uma amostragem de 75 árvores conforme Alves et al. (2007). Para o experimento 2 foram realizadas amostragens aos 30, 36 e 42 meses com amostras contendo 12, 33 e 36 indivíduos, respectivamente (GADELHA et al., 2015). Com os dados de volume real nas idades das amostragens foi calculado o volume real e determinado o fator de forma de 0,50 para o experimento 1 e 0,48 para o experimento 2, que foi empregado para cálculo do volume, com base nos dados de DAP e Ht coletados a cada 6 meses, nas duas situações experimentais.

Ao final de cada experimento foi realizado o procedimento de cubagem das árvores pelo método de Smalian (SILVA; NETO, 1979). Com o auxílio de uma suta foram tomadas medidas de diâmetro ao longo do fuste nos seguintes intervalos: 0,30 m; 0,50 m; 0,70 m; 0,90 m; 1,10 m; 1,30 m; 1,50 m; 1,70 m; 1,90 m; 2,30 m, a partir destes intervalos foram obtidas medidas a cada um metro. A altura real foi medida com uso de uma fita métrica. O volume de madeira em cada seção foi obtido por;

$$v_i = \frac{g_i + g_{i+1}}{2} \cdot l_i \quad (6.1)$$

em que,

$v_i$  = é o volume da seção  $i$  ( $m^3$ );

$g_i$  = área basimétrica no início da seção  $i$  ( $m^2$ );

$g_{i+1}$  = área basimétrica no fim da seção  $i$  ( $m^2$ );

$l_i$  = comprimento da seção  $i$ .

$g_i$  é calculado por meio da seguinte fórmula;

$$g_i = \frac{\pi \cdot d_i^2}{4} \quad (6.2)$$

e  $d_i$  é o diâmetro da seção na posição que se pretende calcular.

O volume de madeira por árvore individual  $V_j$ , em ( $m^3$ ), foi calculado da seguinte forma;

$$V_j = \sum_{i=1}^n v_i \quad (6.3)$$

### 6.2.3 Dados climáticos

As variáveis climáticas foram obtidas de uma estação meteorológica localizada na Estação Experimental do Instituto Agrônomo de Pesquisa no município de Araripina-PE. Os dados de precipitação foram obtidos na sede da Estação Experimental. Os dados de temperatura foram obtidos pelo sistema web Agritempo desenvolvido pelo Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura (CEPA-GRI/UNICAMP) e Empresa Brasileira de Pesquisa Agropecuária (Embrapa Informática Agropecuária) (CEPAGRI/UNICAMP; EMBRAPA, 2019).

## 6.2.4 Comparação da produtividade em dois períodos climáticos

Na análise da precipitação foram utilizados dados na escala mensal e anual referente ao período de 1980 à 2018. A temperatura média do ar foi analisada na escala mensal e no período de 2000 à 2018. Para as análises de tendências em todas as variáveis foram utilizados os testes de Mann-Kendall e análise de regressão, para a detecção de pontos de mudanças foi utilizado o teste de Pettitt e a metodologia *Change point*. Antes da aplicação dos testes de tendências, foi realizada a decomposição da série nas componentes aleatório, tendência e sazonalidade para posterior recomposição da série sem a componente sazonal.

### 6.2.4.1 Análise de tendência e pontos de mudanças em séries climáticas

Inicialmente, foi utilizado o teste de Mann-Kendall na detecção de tendências para as variáveis climáticas. Segundo Gonçalves e Back (2018) o teste considera as seguintes hipóteses:

$H_0$ : não há tendência negativa nem positiva na série, ou seja, os termos da série formam uma amostra aleatória de  $T$  valores independentes e identicamente distribuídos *i.i.d.*

$H_1$ : existe uma tendência negativa ou positiva na série, ou seja, a distribuição das variáveis aleatórias  $X_i$  e  $X_k$  não é *i.i.d.* para alguns termos  $i$ 's da série,  $k \leq T$ ,  $T$  é a dimensão da série.

A estatística do teste foi calculada como segue:

$$S = \sum_{k=1}^{T-1} \sum_{i=k+1}^T \text{sgn}(X_i - X_k) \quad (6.4)$$

Em que,

$$\text{sgn}(X_i - X_k) = \begin{cases} 1 & \text{se } X_i - X_k > 0 \\ 0 & \text{se } X_i - X_k = 0 \\ -1 & \text{se } X_i - X_k < 0 \end{cases} \quad (6.5)$$

Considerando que  $H_0$  não seja rejeitada verifica-se que a estatística  $S$  é normalmente distribuída por meio da variância de  $S$ , tal como:

$$Var(S) = \frac{1}{18} \left[ T(T-1)(2T+5) - \sum_{k=1}^p t_p(t_p-1)(2t_p+5) \right] \quad (6.6)$$

Sendo,  $p$  o número de empates e  $t_p$  o número de pontos no grupo  $k$ . Desta forma, a estatística  $Z$  do teste de Mann-Kendall foi determinada por:

$$Z(MK) = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & se \ S > 0 \\ 0 & se \ S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & se \ S < 0 \end{cases} \quad (6.7)$$

Por fim, se o valor de  $Z$  for positivo a tendência da série é crescente, se negativo a tendência é decrescente. O teste foi realizado com o pacote `trend` da linguagem R, ao nível de 5% de significância (POHLERT, 2018).

As tendências das séries foram estimadas com base em um modelo linear do tipo:

$$Y = \alpha + \beta X + \epsilon, \quad (6.8)$$

Assim, é possível avaliar se existe uma tendência por meio do teste de significância do coeficiente angular. A análise se resume ao cálculo do intervalo de confiança para o coeficiente  $\hat{\beta}$ , a tendência é significativa se os limites do intervalo de confiança não englobarem o valor 0.

Para identificação dos pontos de mudanças de tendência nas séries estudadas foi empregado o teste de Pettitt (1979) e a metodologia *change point* proposta por Chen e Gupta (2000). O teste de Pettitt foi utilizado com nível de 5% de significância. Este é um teste não paramétrico e utiliza parte da estatística proposta por Mann-Whitney ( $U_{t,T}$ ). As hipóteses do teste são:  $H_0$ : a série não apresenta um ponto de mudança e  $H_1$ : existe um ponto de mudança na série. Segundo Pettitt (1979) o valor da estatística do teste é calculado por:

$$K_T = \max_{1 \leq t \leq T} |U_{t,T}| \quad (6.9)$$

Em que,

$$\begin{aligned}
 U_{t,T} &= U_{t-1,T} + V_{t,T} \\
 V_{t,T} &= \sum_{j=1}^T \text{sgn}(X_t - X_j), \quad e \\
 \text{sgn}(x) &= \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x = 0 \\ -1 & \text{se } x < 0 \end{cases}
 \end{aligned} \tag{6.10}$$

Desta forma, o ponto de mudança é obtido pela estatística  $K_T$  que é dado pelo valor máximo absoluto de  $|U_{t,T}|$ , em função de um nível de significância (P) calculado em 6.11, conforme Uliana et al. (2015):

$$P \approx 2e^{\left(\frac{-6(K_T^2)}{T^3+T^2}\right)}. \tag{6.11}$$

Sendo que,  $K_T^2$  é o valor crítico do teste e  $T$  é o tamanho da série histórica. O teste foi realizado por meio do pacote `trend` disponível no R.

A metodologia *change point* também empregada na detecção de mudanças possui a vantagem de ser capaz de identificar múltiplos pontos em séries temporais. Desta forma, as hipóteses do teste são:

$$H_0: F_1 = F_2 = \dots = F_t;$$

$$H_1: F_1 = \dots = F_{k_1} \neq F_{k_1+1} = \dots = F_{k_2} \neq F_{k_2+1} = \dots = F_{k_{\lambda+1}} = \dots = F_t.$$

Em que,  $1 < k_1 < k_2 < \dots < k_{\lambda} < t$  são os múltiplos pontos de mudanças  $\lambda$ , nas posições  $k_1, k_2, \dots, k_{\lambda}$ . Se  $F_1, F_2, \dots, F_t$  seguem uma distribuição de probabilidade comum  $F(\theta)$ , então a hipótese nula é testada sobre os parâmetros populacionais  $\theta_i$ , com  $i = 1, \dots, t$  e  $\theta \in \mathbb{R}^t$  (CHEN; GUPTA, 2012). A estatística do teste pode ser computada conforme Chen e Gupta (2000):

$$\min \left[ \sum_{i=1}^{\lambda+1} [C(x_{k_1}, \dots, x_{\lambda})] + \beta f(k) \right] \tag{6.12}$$

Sendo que,  $[C(x_{k_1}, \dots, x_{\lambda})]$  é uma função custo aplicada para cada segmento da série, usualmente a log-verossimilhança negativa e  $f(k)$  é uma penalidade para

diminuir o sobre-ajustamento, o número de pontos de mudanças e a autocorrelação (TCHORBADJIEFF; ANGELOV, 2017). Para proceder esta análise foi utilizado o pacote `changeoint` com a função `cpt.meanvar`, ou seja, a análise foi realizada em torno da média e da variância, para isso, a análise de normalidade das séries foi realizada pelo teste de Shapiro-Wilk. A função da penalidade foi definida em “MBIC” e o método utilizado foi o “PELT” (KILLICK; ECKLEY, 2014).

#### 6.2.4.2 Análise descritiva dos períodos climáticos e comparação da produtividade

Para uma melhor compreensão do comportamento das variáveis temporais foram construídos gráficos de tendências. A distribuição dos dados das séries foi analisada por meio das seguintes medidas: mínimo, média, mediana, máximo, desvio padrão, curtose e assimetria. Bem como, foram criados histogramas de frequência. Todas estas análises foram desenvolvidas para cada série nos períodos climáticos.

A comparação do crescimento em volume nos dois períodos climáticos foi realizada por meio da identidade das equações obtidas com o modelo de Chapman-Richards, dado por:

$$V_i = \beta_0 \cdot (1 - e^{-\beta_1 \cdot I_i})^{\beta_2} + \epsilon_i \quad (6.13)$$

Em que:

$V_i$  = volume de madeira ( $\text{m}^3 \text{ ha}^{-1}$ );

$I_i$  = idade (meses);

$\beta_0$ ,  $\beta_1$ , e  $\beta_2$  = parâmetros que serão ajustados;

$\epsilon_i$  = erro aleatório  $\sim \text{NID}(0, \sigma^2)$ .

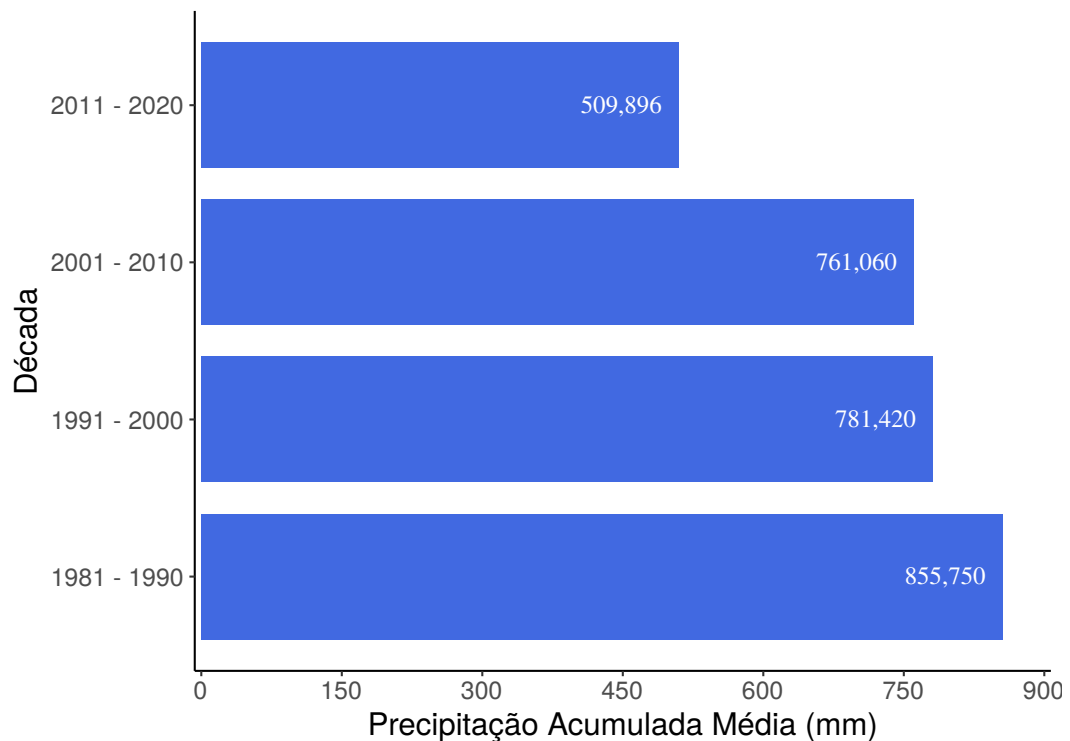
Após obtenção das equações para cada experimento foram realizadas estimativas do volume em um novo conjunto de dados, com ambas as equações, e em seguida se procedeu a metodologia proposta por Leite e Oliveira (2002) para avaliar se as estimativas são estatisticamente diferentes. Se as estimativas forem semelhantes, pode-se concluir que as equações possuem a mesma estrutura e o crescimento em volume não difere de um experimento para outro.

## 6.3 Resultados e Discussão

### 6.3.1 Análise histórica das séries climáticas

A precipitação histórica na região da Chapada do Araripe-PE nas últimas 4 décadas tem sofrido baixas recorrentes, com queda considerável nos últimos 10 anos (Figura 17).

Figura 17 – Precipitação histórica por década na Estação Experimental do IPA no município de Araripina-PE.



A região semiárida nordestina é conhecida por apresentar um clima com bastante irregularidades. Entretanto, é importante investigar se essa mudança é apenas um evento atípico ou pode ser uma nova tendência da série. Pois, anos seguidos de diminuição na precipitação e aumento da temperatura contribuem para o aparecimento de uma atmosfera mais quente, provocando índices mais elevados de seca no solo e maiores perdas de umidade pela evapotranspiração (LACERDA; LOPES, 2016/2017), o que gera uma condição extrema para a vegetação da região.

Pelo teste de Mann - Kendall foi constatado tendência negativa para as variáveis de precipitação e tendência positiva para a temperatura média do ar (Tabela 17).



Tabela 17 – Resultados dos testes de Pettitt, Mann-Kendall e da análise de regressão linear para variáveis do clima oriundas da estação experimental do Instituto Agrônômico de Pernambuco - IPA.

Variáveis	Teste de Pettit	Teste de Mann-Kendall	Análise de Regressão	
			b	Intervalo de Confiança (95%)
Precipitação Anual (mm)	190*	-0,2600*	-7,8180	-14,5848 a -1,0502*
Precipitação Mensal (mm)	11404*	-0,0958*	-0,6675	-1,1281 a -0,2069*
Temperatura Média Mensal (°C)	11024*	0,4560*	0,1769	0,1517 a 0,2021*

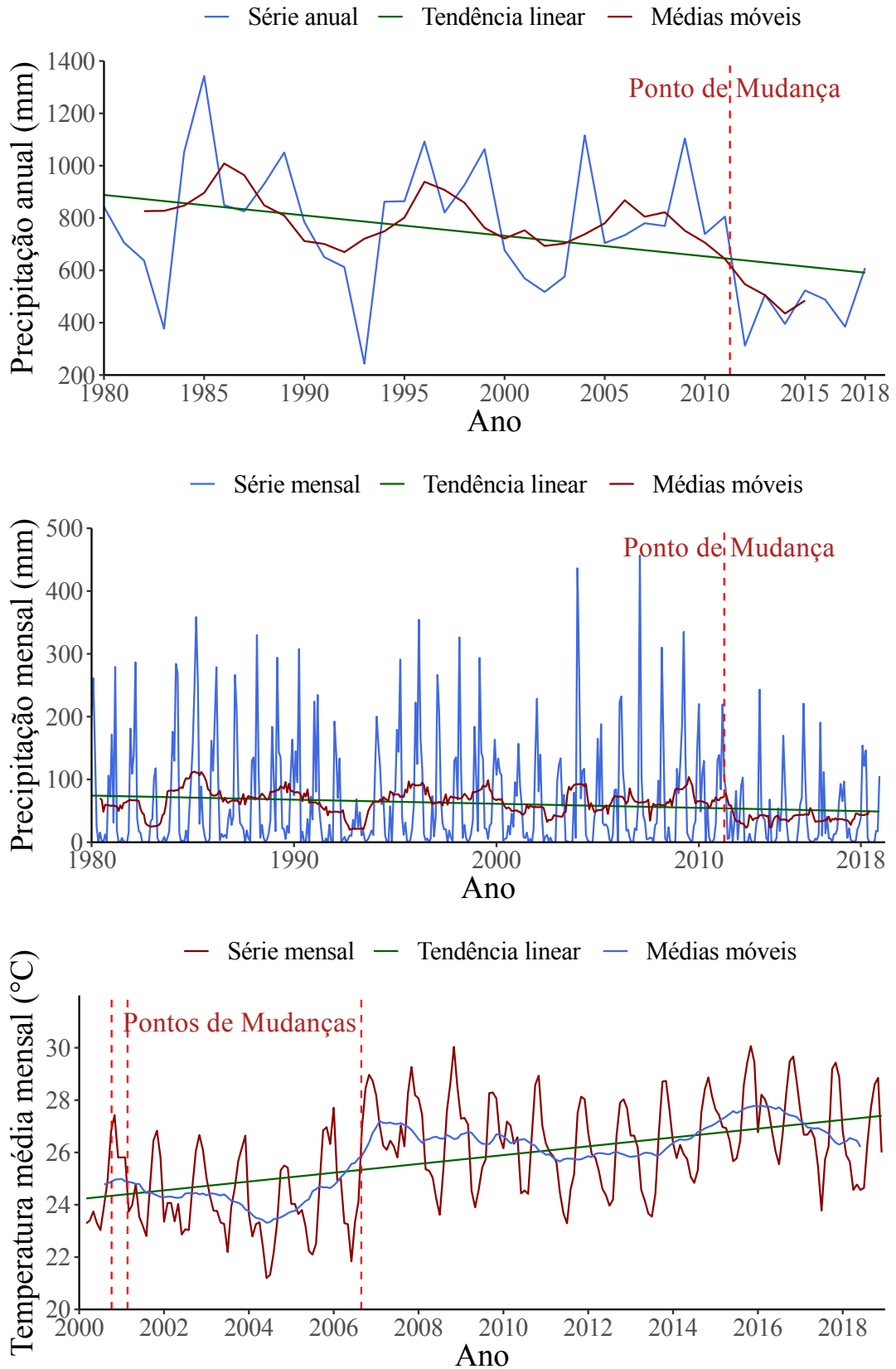
**b**, coeficiente angular da regressão, indica a taxa de inclinação da reta; \*Significativo a 5% de significância.

Pela análise de regressão se observa tendência significativa na diminuição da precipitação acumulada da ordem de 7,8180 mm por ano. Considerando a precipitação mensal houve tendência de queda de 0,6675 mm por mês. Para a temperatura média do ar ocorreu tendência de aumento, entretanto, cabe ressaltar que apesar da tendência ser significativa, para uma afirmação categórica da mesma, seria necessário ao menos um conjunto de dados com 30 anos de duração. Mesmo assim, a elevação na temperatura está presente, pois estudos realizados com dados meteorológicos de 1961 a 2009 do município de Araripina, indicam que houve, neste período, um aumento de 4° C na temperatura máxima (LACERDA et al., 2014/2015).

Para a precipitação acumulada Lacerda et al. (2014/2015) indicam que houve uma diminuição média de 275 mm, aproximadamente 57%, com base em dados oriundos de oito estações pluviométricas (instaladas no Vale do rio Pajeú, Pernambuco), no mesmo período anteriormente citado. Estes estudos corroboram as análises desenvolvidas nesta pesquisa, pois indicam tendências semelhantes.

Em relação aos pontos de mudanças, para as séries de precipitação anual e mensal foi encontrado apenas um ponto que convergiu no ano de 2012 em ambas as séries, já para a série de temperatura média do ar os pontos de mudanças foram detectados no mês de novembro de 2001, fevereiro de 2002 e junho de 2006, bem mais cedo que para os encontrados na série de precipitação (Figura 18). A identificação destes pontos realizada com a metodologia de Chen e Gupta (2000) foram corroboradas pelo teste de Pettit, ambos ao nível de 5% de probabilidade de erro. A linha de tendência das séries realizadas pela análise de regressão linear e a suavização por médias móveis também é destacado na Figura 18.

Figura 18 – Análise de tendência para a precipitação histórica acumulada anual, mensal e temperatura média histórica mensal no município de Araripina-PE.

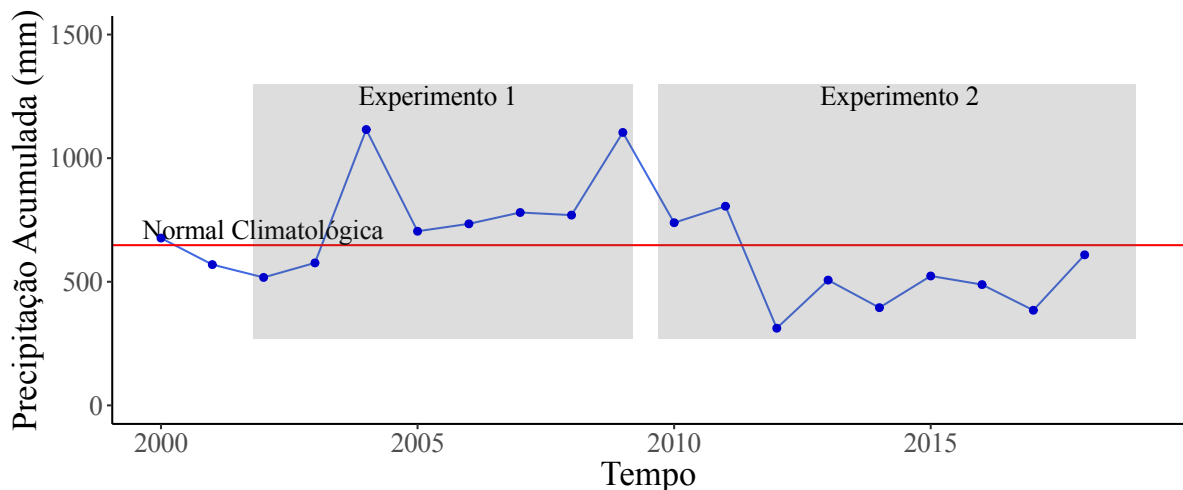


Com base nos testes realizados ficou evidente que existiu mudança na tendência das variáveis climáticas da região, o que corrobora os resultados obtidos no relatório do Painel Brasileiro de Mudanças Climáticas - PBMC (2013), o qual indica a possibilidade de que a região semiárida do Nordeste apresente uma redução de 20% na precipitação até o ano de 2040 e um aumento de até 1 °C na temperatura do ar. Segundo Lacerda e Lopes (2016/2017) o principal fenômeno formador de chuvas no semiárido Nordeste é a chamada Zona de Convergência Intertropical (ZCIT) e afirmam que os eventos de seca extrema dos últimos anos possivelmente foram causados pela mudança na dinâmica de atuação da ZCIT. Os mesmos ainda indicam que estes eventos podem se tornar mais persistentes, pois ultimamente existe uma tendência de reposicionamento da ZCIT que tem se estabelecido na faixa mais ao norte de sua posição média. Esta mudança é uma das fontes dos baixos níveis de formação de chuvas no semiárido nordestino.

### 6.3.2 Comparação das variáveis climáticas dos períodos experimentais

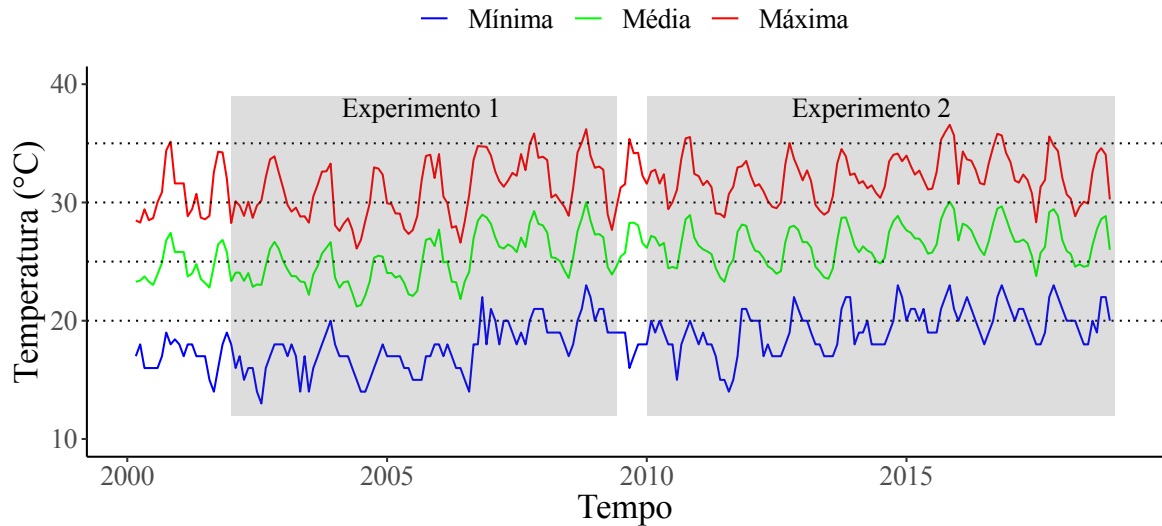
O gráfico da precipitação acumulada anual, com destaque para os dois períodos da realização dos experimentos, é apresentado na Figura 19.

Figura 19 – Precipitação acumulada anual, com destaque para os períodos experimentais, registrada na Estação Experimental do IPA no município de Araripina-PE.



O gráfico de dispersão das temperaturas mínima, média e máxima, também com destaque para os dois períodos experimentais, é apresentado na Figura 20.

Figura 20 – Temperatura histórica, com destaque para os períodos experimentais, registrada na Estação Experimental do IPA no município de Araripina-PE.



Para o experimento 1 a média da precipitação foi de 754,90 mm (variando no intervalo de 517,60 mm à 1115,70 mm), já no segundo experimento a precipitação média foi de 529,40 mm, variando no intervalo de 312,20 mm à 805,60 mm (Tabela 18).

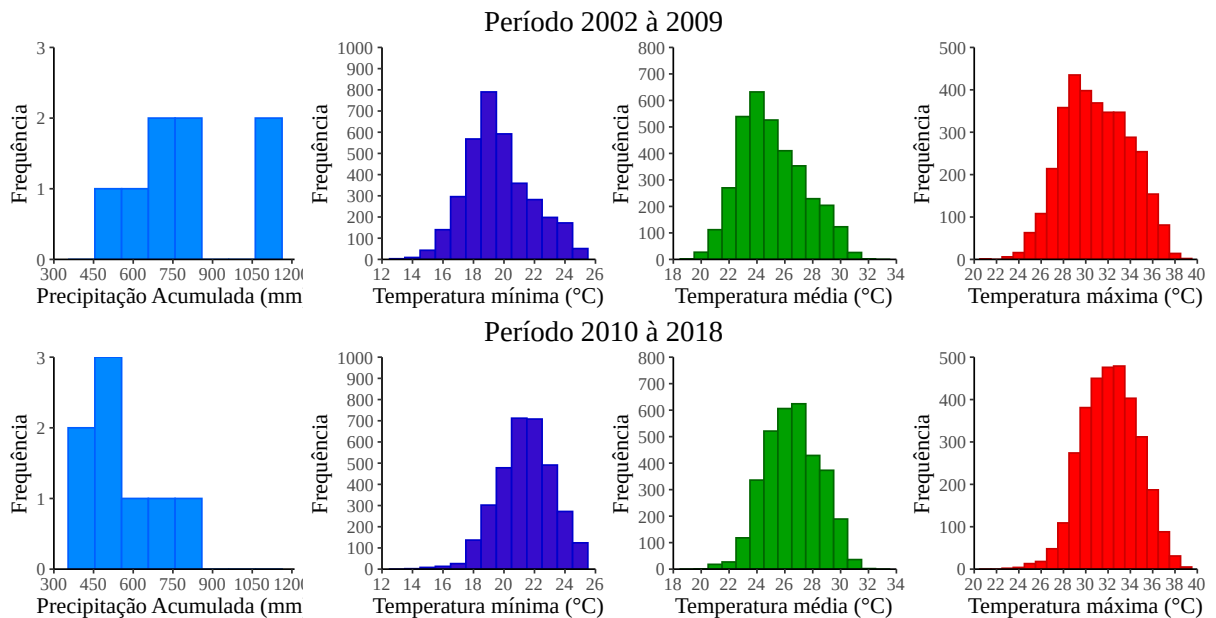
Tabela 18 – Análise estatística descritiva para as variáveis climáticas nos dois períodos de realização dos experimentos implantados no Polo Gesseiro do Araripe-PE.

Estatísticas	Experimento 1 (2000 à 2009)			Experimento 2 (2010 à 2018)				
	PP (mm)	Temperatura °C			PP (mm)	Temperatura °C		
		Mín.	Méd.	Máx.		Mín.	Méd.	Máx.
Mínimo	517,60	13,00	19,00	21,00	312,20	14,00	20,00	23,00
1° Quartil	601,40	18,00	23,00	29,00	395,30	20,00	25,00	30,00
Mediana	719,4	19,00	25,00	31,00	506,80	21,00	27,00	32,00
Média	754,90	19,66	25,08	31,00	529,40	21,40	26,55	32,20
3° Quartil	777,50	21,00	27,00	33,00	609,00	23,00	28,00	34,00
Máximo	1115,70	27,00	32,00	39,00	805,60	26,00	32,00	39,00
Assimetria	0,84	0,35	0,36	0,06	0,45	-0,24	-0,02	-0,09
Curtose	2,51	2,87	2,54	2,34	2,10	3,09	2,59	2,81
Desvio Padrão	206,70	2,17	2,34	2,99	163,85	1,81	1,95	2,49

PP: Precipitação Pluviométrica; Mín: Mínima; Méd: Média; Max: Máxima.

Com base no gráfico de histograma é possível identificar que a distribuição da precipitação foi menor cerca de 200 mm no segundo experimento (Figura 21).

Figura 21 – Histogramas para os dados de precipitação e temperatura nos experimentos.



Para os dados de temperatura (mínima, média e máxima) foi constatado que houve um aumento de mais de 1 °C para as estatísticas de (Mínimo, 1° Quartil, Mediana, Média e 3° Quartil). Pelo histograma da Figura 21 houve indícios de deslocamento positivo na distribuição dos dados, comparando o período de 2002 à 2009 com o período de 2010 à 2018. Outro fator importante para se destacar é que o valor de assimetria mudou de um período para o outro, passando de uma assimetria positiva para negativa, isto indica que anteriormente os valores mais altos de temperatura ocorriam com menor frequência, posteriormente esse fenômeno mudou, os valores mais baixos passaram a ocorrer em menor frequência. Para os períodos estudados os valores de curtose indicam que a distribuição é leptocúrtica, ou seja, a maior parte dos dados se distribuem em torno do centro da distribuição, neste caso próximo da média. A mudança de apenas 1 °C na temperatura pode parecer insignificante, porém mesmo um pequeno aumento na temperatura média do globo, na ordem de grandeza de apenas 1 °C, pode ser o suficiente para ocasionar mudanças no crescimento e na capacidade de regeneração de diversas espécies florestais (KIRSCHBAUM; FISCHIN, 1996).

É de conhecimento científico que os valores de temperatura estão aumentando em todo o globo, segundo Marengo (2010) nos últimos 41 anos a temperatura máxima no Nordeste aumentou de 1,5 °C a 2 °C. Ainda segundo o autor com as perspectivas das mudanças climáticas, o semiárido da região Nordeste que possui uma curta

estação chuvosa poderia, no futuro com clima mais quente e seco, torna-se uma região árida. Isto mostra a real necessidade de se fazer estudos comparativos de produção condicionados às condições climáticas para esta região. Já que, para Krug (2010), as florestas existentes irão quase certamente enfrentar condições climáticas que diferem das atuais. Ainda para a autora a situação se agrava para as florestas com baixa intensidade de manejo ou nenhum manejo, como é o caso da Caatinga, pois para as mesmas existem menos planejamentos para a sua adaptação, o que gera aumento na vulnerabilidade dessa vegetação às mudanças climáticas. Sendo assim, os estudos com clones de *Eucalyptus* são imprescindíveis para mitigar os efeitos da mudança climática local sobre a vegetação Caatinga.

### 6.3.3 Crescimento em volume de clones *Eucalyptus* em condições climáticas distintas

Para se determinar a viabilidade do plantio de clones *Eucalyptus* em condições semiáridas se faz necessária a comparação da produção nos regimes climáticos presentes na região.

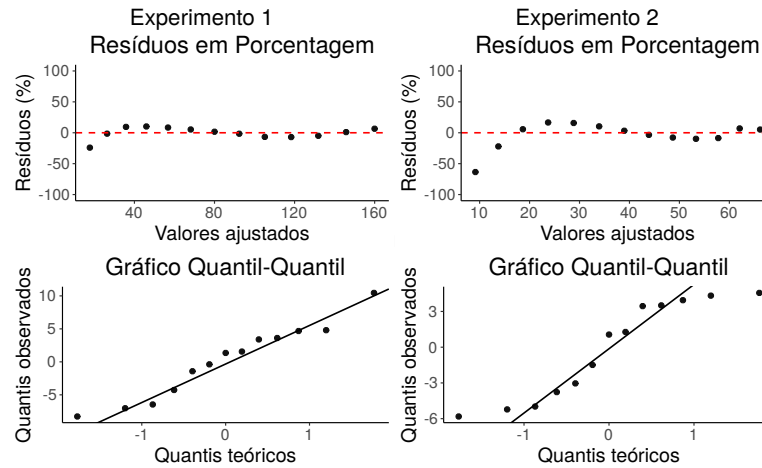
Para avaliar diferenças no crescimento dos períodos experimentais se comparou a similaridade das equações de crescimento do modelo de Chapman-Richards, ajustadas para cada situação. Com base nas estatísticas empregadas para avaliar a qualidade dos ajustes, as duas equações se mostraram apropriadas para a estimativa do crescimento em volume (Tabela 19).

Tabela 19 – Estatísticas para avaliar a qualidade das equações empregadas para comparação do crescimento nos dois períodos experimentais.

Experimentos	Equações	IA <sub>aj</sub>	S <sub>xy</sub> (%)	REQM (m <sup>3</sup> ha <sup>-1</sup> )
1	$\hat{V}_i = 9,527e + 03 \cdot (1 - e^{-5,970e-04 \cdot I_i})^{1,385}$	0,9833	7,2134	5,5014
2	$\hat{V}_i = 132,3314 \cdot (1 - e^{-0,0116 \cdot I_i})^{1,6028}$	0,9495	11,5160	4,0166

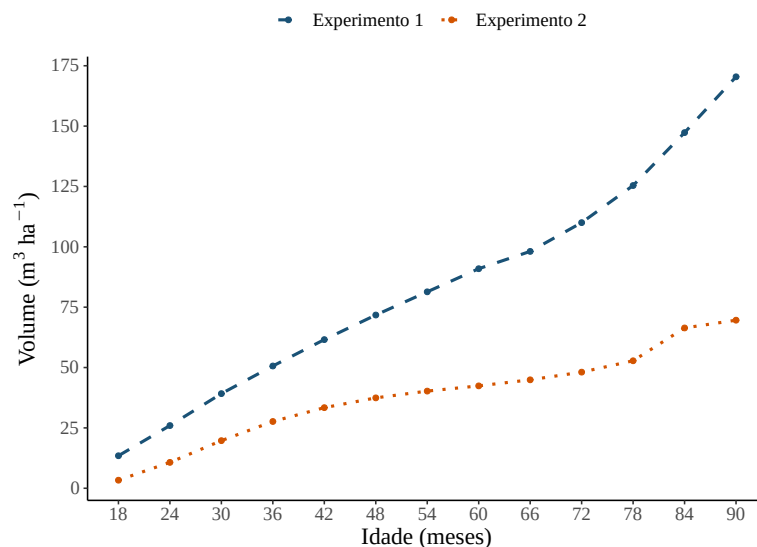
Por meio do teste de Shapiro-Wilk ambas as equações possuem resíduos que seguem a distribuição normal padrão, equação do experimento 1 (p-valor = 0,7122) e equação do experimento 2 (p-valor = 0,0585). Com base na análise gráfica dos resíduos as estimativas geradas pelas equações não apresentam *outliers*, indicando a qualidade dos ajustes (Figura 22).

Figura 22 – Análise dos resíduos para as equações de crescimento nos dois períodos experimentais. Experimento 1 (esquerda) e Experimento 2 (direita).



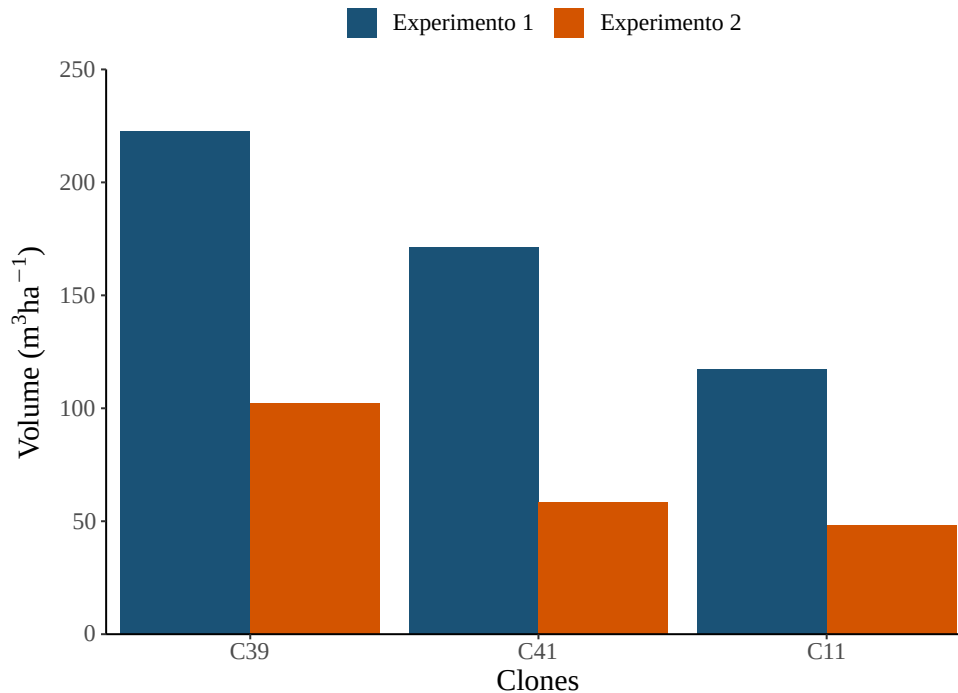
De acordo com a metodologia proposta por Leite e Oliveira (2002) foi estimado o crescimento para um novo conjunto de dados com ambas as equações. Pelo teste  $F(H_0)$  de Graybill's ( $F$  calculado de 3,9823) e teste  $t$  pareado (5,8117) as estimativas são diferentes ao nível de 5% de significância. Existe grande correlação linear entre as estimativas (0,9956) e o erro médio absoluto foi de (0,4796). Desta forma, é possível afirmar que o crescimento nos períodos experimentais ocorre de forma distinta, não sendo possível empregar uma única equação para realizar as estimativas nos dois períodos. As curvas de crescimento reais para os dois casos de estudo evidenciam as diferenças no crescimento ao longo do tempo (Figura 23).

Figura 23 – Curvas de crescimento em volume para os dois experimentos avaliados.



Ao final dos experimentos as diferenças na produção final são bem acentuadas para todos os três clones estudados (Figura 24).

Figura 24 – Produtividade de clones de *Eucalyptus* em dois períodos climáticos no Polo Gesseiro do Araripe-PE.



É grande a diferença de produtividade entre os dois regimes, porém é importante destacar que o período do Experimento 2 apresentou uma condição climática extrema e atípica para a região, conforme análises realizadas nas subseções 6.3.1 e 6.3.2. Com base nos dados da Figura 19 é destacado que houve 7 anos seguidos com níveis de precipitação abaixo da normal climatológica, evento que nunca foi registrado desde a época em que se tem registros dos valores da precipitação para a localidade.

Mesmo com esse evento de seca extrema a produção dos clones de *Eucalyptus* é economicamente viável segundo análises desenvolvidas por Gadelha, Silva e Ferreira (2018). O volume de madeira dos clones nesta condição ainda é bastante superior aos dos PMF's. Segundo Pimentel et al. (2016) em estudo sobre a viabilidade econômica de um PMF no povoado Catolé município de Serra Talhada-PE a estimativa do volume foi  $34,75 \text{ m}^3 \text{ ha}^{-1}$ , considerando um ciclo de corte de 15 anos, bastante inferior ao volume registrado no Experimento 2 desta pesquisa. Este valor é semelhante ao encontrado por Santos et al. (2017b) em área de Caatinga no município de Desterro-PB, que estimou um volume de  $30,037 \text{ m}^3 \text{ ha}^{-1}$ . No geral, é viável o plantio de clones de *Eucalyptus* em ambiente semiárido, porém é importante levar em consideração os clones que



apresentam melhor resistência ao estresse hídrico, principal limitador do incremento em volume nessas regiões.

#### **6.4 Conclusão**

A variável precipitação apresenta tendência de queda no período compreendido entre 1980 e 2018. Por outro lado, a temperatura média mensal apresenta tendência positiva. Foram identificados os anos onde ocorreram mudanças na tendência das variáveis precipitação e temperatura média mensal, indicando que o efeito da mudança climática já pode está atuando no clima da região. Os dois períodos experimentais apresentam diferenças significativas nas variáveis climáticas, sendo o período do experimento 1 que representa o melhor cenário climático. Por meio das equações obtidas foi possível identificar diferenças significativas, no crescimento em volume, entre os dois experimentos.

## Referências Bibliográficas

CEPAGRI/UNICAMP; EMBRAPA, I. A. **Agriempo**. 2019. Disponível em: <www.agriempo.gov.br>.

CHEN, J.; GUPTA, A. K. **Parametric Statistical Change Point Analysis**. Kansas City, MO: Springer Science Business Media LLC, 2000. 190 p.

CHEN, J.; GUPTA, A. K. **Parametric statistical change point analysis: with applications to genetics, medicine, and finance**. 2. ed. New York: Springer Science & Business Media, 2012. 281 p.

GADELHA, F. H. L.; SILVA, J. A. A.; FERREIRA, R. L. C. Viabilidade econômica do cultivo de clones de *Eucalyptus* em função do espaçamento e do sistema silvicultural, em Araripina-PE. **Anais da Academia Pernambucana de Ciência Agrônoma**, v. 15, n. 2, p. 105–123, 2018.

GONÇALVES, F. N.; BACK, Á. J. Análise da variação espacial e sazonal e de tendências na precipitação da região sul do Brasil. **Revista de Ciências Agrárias**, v. 41, n. 3, p. 11–20, 2018.

KEENAN, R. J. Climate change impacts and adaptation in forest management: a review. **Annals of Forest Science**, v. 72, n. 2, p. 145–167, 2015.

KELLOMÄKI, S. et al. Sensitivity of managed boreal forests in Finland to climate change, with implications for adaptive management. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 363, n. 1501, p. 2339–2349, 2008.

KILLICK, R.; ECKLEY, I. A. changepoint: An R package for changepoint analysis. **Journal of Statistical Software**, v. 58, n. 3, p. 1–19, 2014.

KIRSCHBAUM, M.; FISCHIN, A. Climate change impacts on forests. In: WATSON, R.; ZINYOWERA, M. C.; MOSS, R. H. (Ed.). **Climate change 1995: impacts, adaptation and mitigation of climate change: scientific-technical analysis**. England: Cambridge: Cambridge University Press, 1996. p. 95–129.

KRUG, T. Impacto, vulnerabilidade e adaptação das florestas à mudança do clima. **Parcerias estratégicas**, v. 13, n. 27, p. 43–72, 2010.

LACERDA, F. F.; LOPES, G. M. B. Extremos de chuva e seca em Pernambuco: Uma realidade que veio para ficar. **Anais da Academia Pernambucana de Ciência Agronômica**, v. 13/14, p. 42–46, 2016/2017.

LACERDA, F. F. et al. Alterações climáticas globais; uma realidade em pernambuco. **Anais da Academia Pernambucana de Ciência Agronômica**, v. 11/12, p. 121–154, 2014/2015.

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

MARENGO, J. A. Vulnerabilidade, impactos e adaptação à mudança do clima no semi-árido do brasil. **Parcerias estratégicas**, v. 13, n. 27, p. 149–176, 2010.

PBMC. **Contribuição do Grupo de Trabalho 1 ao Primeiro Relatório de Avaliação Nacional do Painel Brasileiro de Mudanças Climáticas**. Rio de Janeiro, Brasil: Sumário Executivo GT1, 2013. 24 p.

PETTITT, A. N. A non-parametric approach to the change-point problem. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, v. 28, n. 2, p. 126–135, 1979.

PIMENTEL, D. J. O. et al. Análise de rentabilidade de manejo florestal sustentável no assentamento Catolé, em Serra Talhada-PE. **South American Journal of Basic Education, Technical and Technological**, v. 3, n. 1, p. 42–49, 2016.

POHLERT, T. **trend: Non-Parametric Trend Tests and Change-Point Detection**. [S.l.], 2018. R package version 1.1.1. Disponível em: <<https://CRAN.R-project.org/package=trend>>.

REYER, C. Forest productivity under environmental change—a review of stand-scale modeling studies. **Current Forestry Reports**, v. 1, n. 2, p. 53–68, 2015.

REYER, C. P. et al. Are forest disturbances amplifying or canceling out climate change-induced productivity changes in european forests? **Environmental Research Letters**, v. 12, n. 3, p. 034027, 2017.

ROCHA, K. D. **Produtividade volumétrica de clones de *Eucalyptus* spp. na região do Pólo gesseiro do Araripe**. 111 p. Tese (Ciências Florestais) — Universidade Federal Rural de Pernambuco, Departamento de Ciências Florestais, Recife, 2012.

RYAN, M. G. et al. Factors controlling eucalyptus productivity: how water availability and stand structure alter production and carbon allocation. **Forest Ecology and Management**, v. 259, n. 9, p. 1695–1703, 2010.

SANTOS, A. C. A. et al. Influência da variabilidade edafoclimática no crescimento de clones de eucalipto no nordeste baiano. **Pesquisa Florestal Brasileira**, v. 37, n. 91, p. 259–268, 2017a.

SANTOS, H. et al. O novo mapa de solos do Brasil: legenda atualizada. **Embrapa Solos-Documentos**, Rio de Janeiro: Embrapa Solos, p. 67, 2011.

SANTOS, W. S. et al. Análise florística-fitossociológica e potencial madeireiro em área de caatinga submetida a manejo florestal. **Agropecuária Científica no Semiárido**, v. 13, n. 3, p. 203–211, 2017.

SILVA, J. A. A. et al. Produtividade volumétrica de clones de *Eucalyptus* spp. no Polo Gesseiro do Araripe, Pernambuco. **Anais da Academia Pernambucana de Ciência Agronômica**, v. 10, p. 240–260, 2013.

TCHORBADJIEFF, A.; ANGELOV, I. Change point analysis as a tool to detect abrupt cosmic ray muons variations. In: GEORGIEV, K.; TODOROV, M.; GEORGIEV, I. (Ed.). **Advanced Computing in Industrial Mathematics**. Sofia, Bulgaria: Springer, 2017. p. 395–406.

ULIANA, E. M. et al. Análise de tendência em séries históricas de vazão e precipitação: uso de teste estatístico não paramétrico. **Ambiente & Água-An Interdisciplinary Journal of Applied Science**, v. 10, n. 1, p. 82–88, 2015.

WAMELINK, G. et al. Modelling impacts of changes in carbon dioxide concentration, climate and nitrogen deposition on carbon sequestration by european forests and forest soils. **Forest Ecology and Management**, v. 258, n. 8, p. 1794–1805, 2009.

## CAPÍTULO 7

---

### **Variáveis climáticas aplicadas em modelos não lineares e algoritmos de aprendizagem de máquinas para estimativa do crescimento volumétrico de clones de *Eucalyptus* spp.**

#### **Resumo**

Objetivou-se com esta pesquisa avaliar se a adição de variáveis climáticas em modelos de crescimento clássicos e nos algoritmos de aprendizagem de máquinas geram estimativas do crescimento em volume com maior eficiência. Foi conduzido um experimento, na Estação Experimental do Instituto Agrônomo de Pernambuco, com três clones de *Eucalyptus* distribuídos em cinco níveis de espaçamentos, em delineamento inteiramente aleatório e arranjo fatorial (3 x 5) com quatro repetições. As variáveis empregadas foram: volume ( $\text{m}^3 \text{ha}^{-1}$ ), idade (meses), precipitação (mm), radiação solar acumulada ( $\text{Mj m}^{-2}$ ), umidade relativa (%), temperatura média do ar ( $^{\circ}\text{C}$ ), temperatura mínima média ( $^{\circ}\text{C}$ ), temperatura mínima absoluta ( $^{\circ}\text{C}$ ), temperatura máxima média ( $^{\circ}\text{C}$ ) e temperatura máxima absoluta ( $^{\circ}\text{C}$ ). As variáveis do clima foram selecionadas com base no procedimento de *Stepwise*, teste não paramétrico de importância das variáveis obtido via rede neural artificial e valor de inflação da variância. Para o crescimento volumétrico foram ajustados os modelos de Schumacher, Clutter-Jones, Prodan, Chapman-Richards e Silva-Bailey, com e sem adição de variáveis climáticas. Bem como, foram treinados os algoritmos de redes neurais artificiais, máquinas de vetores de suporte e floresta aleatória para estimativa do crescimento em volume, com e sem adição das variáveis climáticas. Os modelos foram avaliados por meio do índice de ajuste de Schlaegel corrigido ( $IA_c$ ), o erro padrão de estimativa ( $S_{xy}\%$ ), a raiz do erro quadrático médio (REQM), o erro absoluto médio (EAM) e a análise gráfica dos resíduos. As equações obtidas com o modelo de Schumacher e de Chapman-Richards apresentaram os melhores ajustes. A inclusão das variáveis climáticas nas equações de crescimento melhoram as estimativas. No entanto, estas equações possuem problemas quanto a aderência aos pressupostos da regressão. Os algoritmos de inteligência artificial foram capazes de estimar o crescimento florestal com eficiência semelhante aos modelos clássicos de crescimento. Estes se tornam uma importante ferramenta

para o setor florestal por serem robustos a falta de normalidade e homogeneidade da variância dos resíduos. A adição das variáveis climáticas aos algoritmos inteligentes não proporcionaram melhora significativa nas estimativas.

**Palavras-chave:** Polo Gesseiro do Araripe-PE, Schumacher, Chapman-Richards, redes neurais artificiais, máquinas de vetores de suporte, florestas aleatórias.

## Abstract

The objective of this research was to evaluate whether the addition of climatic variables in classic growth models and in Artificial Intelligence techniques generates estimates of growth in volume with greater efficiency. An experiment was conducted at the Experimental Station of the Instituto Agrônomo de Pernambuco, with three *Eucalyptus* clones distributed in five spacing levels, in a completely randomized design and factorial arrangement (3 x 5) with for replications. The variables used were: volume ( $\text{m}^3 \text{ ha}^{-1}$ ), age (months), precipitation (mm), accumulated solar radiation ( $\text{Mj m}^{-2}$ ), relative humidity (%), average air temperature ( $^{\circ}\text{C}$ ), minimum temperature average ( $^{\circ}\text{C}$ ), minimum absolute temperature ( $^{\circ}\text{C}$ ), temperature average maximum ( $^{\circ}\text{C}$ ) and absolute maximum temperature ( $^{\circ}\text{C}$ ). The climate variables were selected based on the Stepwise test, a non-parametric test of the importance of the variables obtained via the artificial neural network and the variance inflation value. For volumetric growth, the models of Schumacher, Clutter-Jones, Prodan, Chapman-Richards and Silva-Bailey were adjusted, with and without climatic variables. As well, artificial neural network algorithms, support vector machines and random forest were trained to estimate growth in volume, with and without addition of climatic variables. The models were evaluated using the corrected Schlaegel adjustment index ( $\text{AI}_c$ ), the standard error of estimate ( $\text{S}_{xy}\%$ ), the root of mean square error (RMSE), the error mean absolute value (EMA) and the graphical analysis of residues. The equations obtained from the Schumacher and Chapman-Richards model presented the best adjustments. The inclusion of climatic variables in the growth equations improves the estimates. However, these equations have problems with adherence to the regression assumptions. Artificial intelligence algorithms were capable of estimating forest growth with efficiency similar to classic growth models. These become an important tool for the forestry sector because they are robust due to the lack of normality and homogeneity of the variance of the residues. The addition of the climatic variables to the intelligent algorithms did not provide significant improvement in the estimates.

**Key-words:** Gypsum Pole of Araripe-PE, Schumacher, Chapman-Richards, artificial

neural network, support vector machines, random forest.

## 7.1 Introdução

Nos últimos anos, os efeitos das Mudanças Climáticas têm se intensificado (HUTCHINS et al., 2019; LOTZE et al., 2019; OLANIYI; OLUTIMEHIN; FUNMILAYO, 2019). Portanto, tem crescido o interesse no desenvolvimento de modelos biológicos de crescimento para povoamentos florestais que incluam variáveis climáticas em sua estrutura (MACHADO et al., 2014).

O crescimento de uma floresta é definido como o aumento das dimensões de uma ou de várias árvores em determinado tempo (BARRAZA, 2014). Uma forma de representar o crescimento de um povoamento é por meio de modelos biológicos. No desenvolvimento científico atual se pode definir duas linhas filosóficas sobre o desenvolvimento de modelos (MAESTRI et al., 2013). A primeira linha se concentra no estudo de modelos descritivos, empíricos ou estatísticos que são empregados frequentemente por especialistas do manejo florestal. A segunda está relacionada com o desenvolvimento de modelos baseados em processos biológicos e esses são mais complexos para uso direto no manejo (SCOLFORO et al., 2013).

O principal problema relacionado ao emprego direto dos modelos baseados em processos está ligado com a quantidade de parâmetros necessários para representar a fisiologia da planta, o que gera um aumento na complexidade de seu ajuste e interpretação (FERRAZ-FILHO et al., 2011). Desta forma, tem crescido o interesse no uso de modelos híbridos, os quais permitem a inclusão de variáveis climáticas de fácil obtenção.

Os algoritmos de aprendizagem de máquinas também constituem um conjunto de ferramentas que ganharam destaque na área de ciências florestais. Os mais difundidos no meio florestal são as redes neurais artificiais, as máquinas de vetores de suporte e as florestas aleatórias. Diversos estudos são reportados com o uso destas técnicas na área florestal (BHERING et al., 2015; GARCÍA-GUTIÉRREZ et al., 2015; BINOTI et al., 2016; ÖZÇELIK et al., 2017; ABREU, 2019). Entretanto, em sua maioria as aplicações dos algoritmos de aprendizagem de máquinas fazem uso apenas das variáveis dendrométricas como entrada para estimativa da produção volumétrica, do crescimento em volume ou altura, do perfil do fuste, etc (BINOTI et al., 2015; BONETE et al., 2020; CASTRO et al., 2020). Dessa forma, se faz necessário a investigação da adição das variáveis do clima nos algoritmos de aprendizagem de máquinas.

A inclusão de variáveis climáticas em modelos biológicos e nos algoritmos

de aprendizagem de máquinas pode gerar ganhos na estimativa do crescimento. Principalmente quando o povoamento está inserido em regiões geográficas sujeitas a variações extremas do clima (ELLI et al., 2017). Tendo em vista que, com o aumento dos efeitos da mudança climática no semiárido brasileiro a ocorrência de eventos extremos do clima será mais frequente, principalmente com longos períodos de seca (LACERDA et al., 2014/2015), o desenvolvimento de modelos biológicos e de aprendizagem de máquinas para o crescimento volumétrico com variáveis climáticas em sua estrutura é de extrema para agentes responsáveis por planos de manejo florestal.

Diante do exposto, objetivou-se, com esta pesquisa, modelar o crescimento volumétrico em povoamentos de *Eucalyptus* spp. e avaliar a contribuição de variáveis climáticas ao desempenho dos modelos de crescimento clássicos nos algoritmos de aprendizagem de máquinas.

## 7.2 Material e Métodos

### 7.2.1 Descrição do experimento e variáveis dendrométricas

O experimento foi montado no mês de março de 2010 com clones de eucaliptos, em uma área de 7,5 hectares, localizada na Estação Experimental do Instituto Agrônomo de Pernambuco (IPA), coordenadas 07° 28' 04,2" S e 40° 24' 37,7" W, altitude de 816 metros, no município de Araripina-PE (Figura 25). A precipitação média anual da região é de 740 mm e possui elevados índices de evaporação, a temperatura média do ar alta é de 25 °C. Segundo classificação de Köppen, o clima da região é do tipo BSh, semiárido quente, caracterizado por longos períodos de seca e chuvas irregulares.

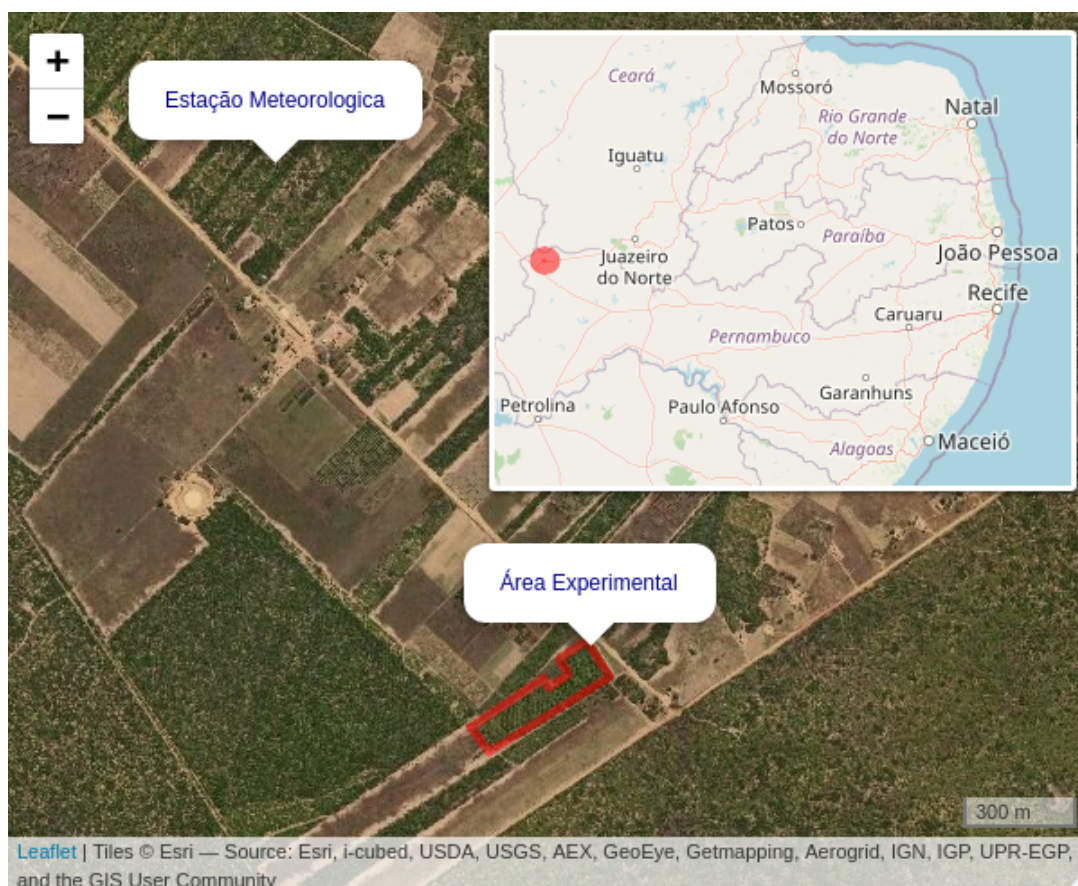
Foram empregados no experimento três clones de eucaliptos (C39 e C41 provenientes de *E. urophylla* e C11 obtido do *E. brassiana*) distribuídos em cinco espaçamentos (2 m x 1 m, 2 m x 2 m, 3 m x 2 m, 3 m x 3 m e 4 m x 2 m). Estes foram dispostos em um delineamento experimental inteiramente aleatório com arranjo fatorial (3 x 5), com quatro repetições. Em cada parcela foram plantadas 49 árvores, desprezando-se o efeito de bordadura, só as 25 árvores da área útil foram consideradas.

A partir do mês de setembro de 2011 foram medidas as variáveis dendrométricas altura - Ht (m) e diâmetro a altura do peito - DAP (cm) em intervalos de seis meses até o mês de março de 2018. Nos anos de 2012, 2013 e 2014 foram realizadas amostragens para realização da cubagem rigorosa e posteriormente determinação do fator de forma (GADELHA, 2014). Em março de 2018 foi determinado o volume (m<sup>3</sup>) por meio da cubagem rigorosa de todas as árvores do experimento, empregando a metodologia de



Smalian, conforme SILVA e Neto (1979). Após a determinação do fator de forma para os quatro anos, estes foram empregados para cálculo do volume em todos os períodos semestrais nos quais foram tomadas as medidas de Ht e DAP. Esses dados permitiram determinar o volume de madeira ( $m^3 ha^{-1}$ ) por clone nas diferentes idades (meses).

Figura 25 – Localização da área experimental com clones de *Eucalyptus* inserida no município de Araripina-PE.



### 7.2.2 Variáveis climáticas

As variáveis climáticas foram disponibilizados pela Agência Pernambucana de Águas e Clima (APAC). Os dados foram obtidos em escala diária, posteriormente foram transformados para escala mensal e em seguida foram calculados para o período semestral de forma a sintetizar as informações de cada período de coleta. As variáveis disponibilizadas foram: precipitação (mm), radiação solar acumulada ( $Mj m^{-2}$ ), umidade relativa (%), temperatura média do ar ( $^{\circ}C$ ), temperatura mínima média ( $^{\circ}C$ ), temperatura mínima absoluta ( $^{\circ}C$ ), temperatura máxima média ( $^{\circ}C$ ) e temperatura máxima absoluta ( $^{\circ}C$ ). A estação meteorológica (coordenadas,  $7^{\circ}27'29,9''S$  e  $40^{\circ} 25'01,2'' W$ ) estava localizada à 1,27 Km do local do experimento.

### 7.2.3 Modelos de crescimento clássicos

Diversos modelos de crescimento são discutidos na literatura especializada, no entanto as estimativas mais precisas são obtidas com um contingente específico de modelos. Sendo assim, neste trabalho foram empregados os modelos que apresentaram melhores estimativas para o crescimento florestal conforme apresentado em Elli et al. (2017) e Guera et al. (2019). As expressões dos modelos são dadas na Tabela 20.

Tabela 20 – Modelos clássicos de crescimento para estimativa do volume em nível de povoamento.

Nº	Modelos	Autores
MC1	$V_i = \beta_0 \cdot e^{(\beta_1 \cdot \frac{1}{I_i})} \cdot \epsilon_i$	Schumacher
MC2	$V_i = \beta_0 \cdot (1 + \beta_1 \cdot I_i^{\beta_2})^{\beta_3} + \epsilon_i$	Clutter e Jones
MC3	$V_i = \frac{I_i^2}{(\beta_0 + \beta_1 \cdot I_i + \beta_2 \cdot I_i^2)} + \epsilon_i$	Prodan
MC4	$V_i = \beta_0 \cdot (1 - e^{-\beta_1 \cdot I_i})^{\beta_2} + \epsilon_i$	Chapman e Richards
MC5	$V_i = \beta_0 \cdot e^{\beta_1 \cdot \beta_2^{I_i}} + \epsilon_i$	Silva e Bailey

Em que:  $V_i$  = volume ( $m^3 \text{ ha}^{-1}$ );  $I_i$  = Idade (meses);  $\beta_0, \beta_1$  e  $\beta_2$  = parâmetros do modelo;  $\epsilon_i$  = erro aleatório.

A inclusão de variáveis climáticas em modelos de crescimento pode ser realizada de diferentes formas, no entanto uma das mais empregadas na literatura é associar estas variáveis com o parâmetro assintótico do modelo. Desta forma, se for possível fazer uso de um fator modificador no parâmetro assintótico de qualquer modelo de crescimento, é possível determinar o efeito climático no crescimento do povoamento florestal (SCOLFORO et al., 2013). Para determinar o melhor fator modificador (Mod.), foram testadas três maneiras diferentes de inclusão:

- Primeira forma:

$$Mod. = \kappa_1 \cdot VC_1 + \kappa_2 \cdot VC_2 + \dots + \kappa_n \cdot VC_n \quad (7.1)$$

- Segunda forma:

$$Mod. = VC_1^{\kappa_1} \cdot VC_2^{\kappa_2}, \dots, VC_n^{\kappa_n} \quad (7.2)$$

- Terceira forma:

$$Mod. = \frac{\beta_0}{1 - (VC_1^{\kappa_1} \cdot VC_2^{\kappa_2}, \dots, VC_n^{\kappa_n})} \quad (7.3)$$

Considerando a assíntota ( $\beta_0$ ) do modelo de Schumacher, o fator modificador (Mod.) foi adicionado conforme:

$$V_i = Mod. \cdot e^{(\beta_1 \cdot \frac{1}{T_i})} + \epsilon_i \quad (7.4)$$

Em que:

*Mod.* = fator modificador;  $\kappa_1, \kappa_2, \kappa_n$  = parâmetros da regressão;  $VC_1, VC_2, VC_n$  = variáveis climáticas definidas na subseção 7.2.2. Para os demais modelos a inclusão ocorreu de forma semelhante.

Para ajuste dos modelos foi utilizado o método dos mínimos quadrados ordinários com o algoritmo iterativo de Gauss-Newton e em casos de não convergência deste foi empregado o algoritmo de Levenberg-Marquardt. Todos os procedimentos estatísticos foram realizados com a linguagem de programação R (R Core Team, 2016).

#### 7.2.4 Estimativa do crescimento via algoritmos da aprendizagem de máquinas

Para treinar os algoritmos de inteligência artificial, o conjunto de dados com um total de 168 amostras, foi aleatoriamente dividido nos subconjuntos de treino (80%) e teste (20%).

##### 7.2.4.1 Rede neural artificial

No treino da rede neural artificial (RNA) os dados foram normalizados com a finalidade de diminuir a amplitude de valores extremos e aumentar a velocidade na convergência do algoritmo. Neste estudo foi empregada a rede neural do tipo perceptron de múltiplas camadas (Multilayer perceptron MLP) com a arquitetura formada por uma camada de entrada (idades e variáveis do clima), uma intermediária (neurônios ocultos) e uma de saída (volume  $m^3 ha^{-1}$ ). As camadas são interligadas por pesos sinápticos ( $w$ ), os quais serão ajustados em cada ciclo de treinamento e corrigidos em função do erro quadrático médio. Esta função é minimizada via gradiente descendente estocástico sendo calculada por meio do algoritmo backpropagation. Na camada intermediária foram consideradas as funções de ativação tangente hiperbólica, rectifier e maxout e na camada de saída, a função linear. Para se ter uma ideia inicial do número de neurônios na camada intermediária, conforme apresentado por Valença (2010), foi empregado a fórmula de Hetch-Nielsen, em que  $N_{intermediária} = 2N_{entrada} + 1$ , sendo assim, foram consideradas diferentes quantidades de neurônios em torno do valor calculado.

O número de ciclos ou épocas que serão utilizados para treinar a rede também pode afetar o ajuste causando overfitting ou sendo insuficiente para uma boa generalização. Então, foram considerados diferentes número de ciclos (200, 500 e 1000). A taxa erro de tolerância para parada do treinamento foi de 0,001. Aliado a estes critérios de parada foi utilizado a validação cruzada com o método k-fold. Sendo assim, o conjunto de treinamento foi dividido em 5 partes iguais. Com esse método é realizado o treinamento de cada rede k vezes, com (5-1) partes, ficando uma parte para validação, dessa forma, é gerado uma medida mais eficiente sobre a capacidade do algoritmo na predição do volume. Considerando a quantidade de parâmetros e hiperparâmetros (número de neurônio, número de ciclo, função de ativação, etc) a serem testados/ajustados foi empregado um grid em que todas as arquiteturas foram criadas e testadas aleatoriamente, bem como foram calibrados todos os hiperparâmetros do algoritmo.

#### 7.2.4.2 Floresta aleatória

Para o treino da floresta aleatória (FA) a normalização não se faz necessária, já a validação cruzada foi empregada da mesma forma como apresentada para a RNA. No treinamento do algoritmo FA,  $K$  amostras aleatórias são geradas via amostragem bootstrap para a base de treino. Para cada  $K_i$  uma árvore de regressão é gerada (*ntrees*, mais difundida em fenômenos de classificação como árvore de decisão). Para cada árvore de regressão são criadas regras de decisão (nó), cada regra gerada pode empregar por padrão  $v/3$  variáveis explicativas ( $v = \text{variáveis}$ ), o parâmetro  $v/3$  (*mtries*, a seleção das variáveis é realizada aleatoriamente) será otimizado. Quanto maior a profundidade da árvore maior o ajuste sobre a base de treino, o parâmetro (*max\_depth*, quantidade de regras) é otimizado de forma a garantir o maior poder de generalização do algoritmo. Dessa forma, para a divisão de cada nó são escolhidas variáveis que maximizam a redução de uma função erro, e dessa forma, é reduzido conjuntamente o viés e a variância da estimativa. Na última ramificação, chamada de folha, cada árvore de regressão gera uma estimativa da variável resposta com base nas regras construídas a partir das  $v/3$  variáveis explicativas. A estimativa geral é dada por:

$$V_i = \frac{1}{k_i} \sum_{i=1}^k f^k(V_i) \quad (7.5)$$

Os hiperparâmetros que foram ajustados durante o treinamento são: o número de árvores de regressão (*ntrees*, este variou de 1 até 1000, parando de incrementar quando se estabilizou a variância e o bias), o número de variáveis que foram aleato-

riamente empregadas em cada árvore (*mtries*, variou de 2 a 4) e a profundidade da árvore (*max\_depth*, fixado em um máximo de 30). Da mesma forma como foi desenvolvido para a RNA, os hiperparâmetros foram adicionados em um grid de busca e suas combinações foram testadas.

### 7.2.4.3 Máquina de vetor de suporte

O algoritmo máquina de vetor de suporte (MVS) para regressão de fenômenos não lineares consiste em mapear um vetor de entrada  $\mathbf{x}$  em um espaço ( $I$ ) de características de alta dimensionalidade  $H$ , desta forma,  $\mathbf{x} \in \mathbb{R}^I \mapsto \phi(\mathbf{x}) \in \mathbb{R}^H$ ,  $\phi(\mathbf{x})$  representa uma função kernel empregada para buscar os vetores de suporte, presentes no conjunto de treino, que estimam a variável resposta ( $y$ , nesse caso volume de madeira), com a menor variação possível. Logo, uma MVS para problemas de regressão não linear busca os vetores de suporte (parâmetros ou coeficientes) minimizando a função, conforme Drucker et al. (1997):

$$L(\alpha) = -\epsilon \sum_{i=1}^N (\alpha_i^* + \alpha_i) + \sum_{i=1}^N y_i (\alpha_i^* - \alpha_i) - \frac{1}{2} \sum_{i,j=1}^N (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) \phi(x_i, x_j) \quad (7.6)$$

Com as restrições:

$$\begin{aligned} 0 &\leq \alpha_i \leq C \quad i = 1, \dots, N \\ 0 &\leq \alpha_i^* \leq C \\ \sum_{i=1}^N \alpha_i^* &= \sum_{i=1}^N \alpha_i \end{aligned} \quad (7.7)$$

Sendo  $(\alpha_i, \alpha_i^*)$ ,  $N$  pares de multiplicadores de Lagrange. No treino da MVS foram consideradas as funções kernel: linear, polinomial e de base radial (gaussiana).

Da mesma forma foi construído um *grid* de busca de maneira a otimizar todos os hiperparâmetros necessários para melhorar a estimativa e evitar o super ajustamento. Os hiperparâmetros otimizados foram:  $C$  (variando de 2 até 256) este é um hiperparâmetro de custo, quando  $C$  é grande, maior ênfase é dada ao erro, já, quando  $C$  é pequeno é dada maior ênfase para a norma dos pesos, melhorando assim a generalização em novos conjuntos de dados;  $\gamma$  (variando de 0,1 até 0,9) este controla a suavização da decisão, ou seja, define o hiperplano ótimo, controlando assim a variação do modelo;  $\epsilon$  (indo de 0 até 0,9) está relacionado com o número de vetores de suporte que serão utilizados, os demais hiperparâmetros são específicos para as funções de kernel utilizadas e foram considerados com o valor padrão.

Os algoritmos de aprendizagem de máquinas possuem a natureza aleatória para a seleção das variáveis e para a entrada de cada elemento da amostra, assim a estimativa da variável resposta realizada por eles depende de como os dados das variáveis explicativas serão apresentados ao algoritmo. Para se avaliar os algoritmos de forma eficiente foi empregada a amostragem via *bootstrap* do conjunto de teste. Onde no cálculo da performance dos algoritmos foram geradas 100 amostras aleatórias de tamanho 24 a partir da base de teste, a qual possuía 34 exemplos. Logo, a performance final de cada algoritmo é dada pela média das estatísticas obtidas nas 100 amostras.

Os treinamentos das RNA's e das FA's foram realizados com a plataforma H<sub>2</sub>O. A mesma é compatível com a linguagem R por meio do pacote H<sub>2</sub>O (LEDELL et al., 2019). O treinamento das MVS's foi realizado com o pacote e1071 escrito na linguagem R (MEYER et al., 2017).

### 7.2.5 Seleção das variáveis climáticas

A adição de um conjunto de variáveis climáticas em um modelo de crescimento pode acrescentar problemas relacionados à multicolinearidade devido a presença de correlação entre as variáveis. Para contornar esta dificuldade foi empregado a metodologia de *Stepwise* para seleção de variáveis. Foi utilizado o procedimento *forward* no qual foi ajustado o modelo  $Mod. = \kappa_1 \cdot VC_1 + \kappa_2 \cdot VC_2 + \dots + \kappa_n \cdot VC_n + \epsilon$  e posteriormente realizada a exclusão das variáveis com menor importância conforme teste de Akaike (AIC). O teste AIC utiliza a distância de Kulbach-Leibler que é calculada entre o modelo de aproximado e o modelo completo. Desta forma, o AIC é uma medida superior ao teste F-parcial e ao R<sup>2</sup> (coeficiente de determinação) na seleção de variáveis.

Em seguida também foi calculado o valor de inflação da variância (VIF) para medir a multicolinearidade. O VIF exprime o aumento na variância de um coeficiente da equação se os demais preditores apresentam elevada correlação. Um valor de VIF superior a 10 indicam a existência de multicolinearidade. As variáveis que apresentaram altos valores de VIF foram excluídas dos modelos.

Todos esses métodos de seleção de variáveis são lineares, sendo assim variáveis que apresentem relações não lineares podem não ser incluídas. Desta forma, as relações não lineares foram medidas por meio do grau de importância das variáveis no treino da RNA. Para este procedimento foi empregado a metodologia desenvolvida por Gedeon (1997), esta consiste em remover pares de pesos da matriz de neurônios que ligam as variáveis de entrada com os neurônios da camada oculta, assim, pares de

variáveis são removidos para se avaliar o efeito na precisão da rede.

### 7.2.6 Avaliação das estimativas

Para avaliação da qualidade dos ajustes, foram empregados os seguintes testes estatísticos: Índice de ajuste de Schlaegel corrigido ( $IA_c$ ), erro padrão de estimativa em porcentagem ( $S_{xy}$  %), raiz do erro quadrático médio (REQM  $m^3 ha^{-1}$ ), erro absoluto médio (EAM  $m^3 ha^{-1}$ ). Bem como, a normalidade dos resíduos foi analisada pelo teste de Shapiro-Wilk, a homogeneidade das variâncias foi observada pelo teste de Breush-Pagan e autocorrelação foi medida pelo teste de Durbin-Watson. Também foi realizada a análise gráfica dos resíduos.

## 7.3 Resultados e Discussão

### 7.3.1 Análise da importância das variáveis climáticas

As variáveis climáticas que apresentam maior importância na estimativa do crescimento em volume ( $m^3 ha^{-1}$ ) selecionadas pelo procedimento *Stepwise* são apresentadas na Tabela 21. Essas variáveis possuem uma relação linear com o volume, apresentando coeficiente de determinação ( $R^2$ ) de 0,72.

Tabela 21 – Variáveis climáticas selecionadas pelo método *setwise*.

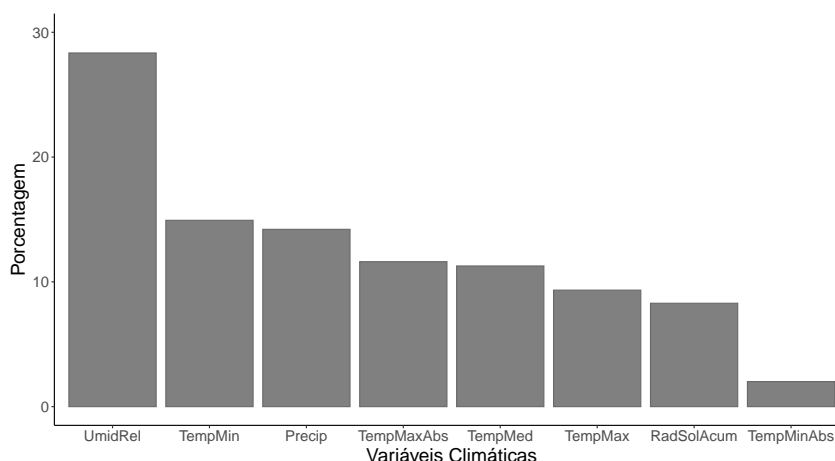
Variáveis	Estimativa	Erro Sd.	t valor	Pr(> t )
(Intercepto)	-411,7906	37,1593	-11,08	<0,01
RadSolAcum	0,0125	0,0022	5,74	<0,01
TempMaxAbs	0,8347	0,1491	5,60	<0,01
TempMin	13,9607	1,7964	7,77	<0,01
TempMinAbs	-5,7682	1,3983	-4,13	<0,01
UmidRel	2,3771	0,3030	7,85	<0,01

RadSolAcum = radiação solar acumulada ( $Mj m^{-2}$ ), TempMaxAbs = temperatura máxima absoluta ( $^{\circ}C$ ), TempMin = temperatura mínima média ( $^{\circ}C$ ), TempMinAbs = temperatura mínima absoluta ( $^{\circ}C$ ), UmidRel = umidade relativa (%).

Na Figura 26 são dadas as variáveis com maior valor de importância para estimativa do volume por meio do uso da técnica de redes neurais artificiais. Com as RNA's é possível captar relações entre as variáveis com propriedades diferentes da

linear. Como a maioria dos testes estatísticos para cálculo da importância de variáveis possuem o pressuposto da linearidade entre os preditores e o predito, os algoritmos de aprendizagem de máquinas possuem a vantagem de quantificar correlações de diferentes naturezas.

Figura 26 – Importância de variáveis climáticas na estimativa do volume obtidas via redes neurais artificiais.



RadSolAcum = radiação solar acumulada ( $\text{Mj m}^{-2}$ ), TempMed = temperatura média ( $^{\circ}\text{C}$ ), TempMax = temperatura máxima média ( $^{\circ}\text{C}$ ), TempMaxAbs = temperatura máxima absoluta ( $^{\circ}\text{C}$ ), TempMin = temperatura mínima média ( $^{\circ}\text{C}$ ), TempMinAbs = temperatura mínima absoluta ( $^{\circ}\text{C}$ ), UmidRel = umidade relativa (%), Precip = Precipitação (mm).

Com base nas duas análises apresentadas é possível observar que diferentes variáveis foram indicadas. Analisando o resultado obtido com as redes neurais artificiais se identifica que algumas das variáveis com maior grau de importância possuem elevada correlação entre si. Desta forma, para evitar problemas de multicolinearidade nos modelos, o conjunto de variáveis final foi selecionado com base no valor de inflação da variância (VIF) dado na Tabela 22.

Com base nessas análises as variáveis candidatas para entrarem nos modelos, de acordo com o seu grau de importância foram: umidade relativa (%), Precipitação (mm), temperatura máxima média ( $^{\circ}\text{C}$ ), temperatura máxima absoluta ( $^{\circ}\text{C}$ ), radiação solar acumulada ( $\text{Mj m}^{-2}$ ) e temperatura mínima absoluta ( $^{\circ}\text{C}$ ).

### 7.3.2 Modelagem do crescimento clássica

As estatísticas de qualidade e as equações obtidas para predição do crescimento em volume dos clones de *Eucalyptus* são mostradas na Tabela 23. De forma geral todas as equações apresentaram bons ajustes, o erro padrão de estimativa ficou em



Tabela 22 – Variáveis selecionadas pelo valor de inflação da variância (VIF).

Variáveis	VIF
RadSolAcum	2,47
TempMax	4,56
TempMaxAbs	1,86
TempMinAbs	6,14
UmidRel	3,77
Precip	1,35

RadSolAcum = radiação solar acumulada (Mj m<sup>-2</sup>), TempMax = temperatura máxima média (°C), TempMaxAbs = temperatura máxima absoluta (°C), TempMinAbs = temperatura mínima absoluta (°C), UmidRel = umidade relativa (%), Precip = Precipitação (mm).

torno de 20%, com exceção da equação obtida por meio do modelo de Silva e Bailey com **S<sub>xy</sub>** acima de 24%.

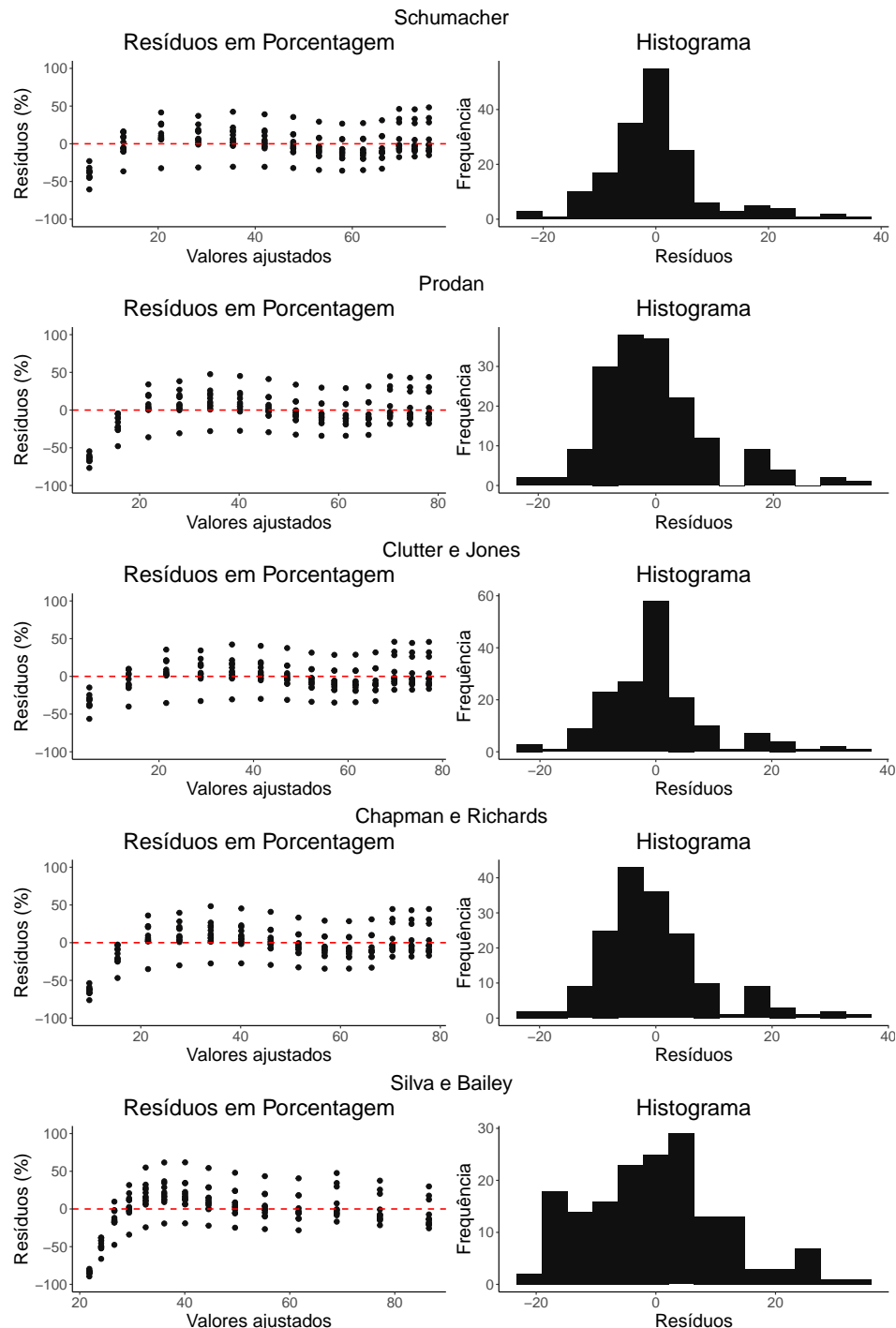
Tabela 23 – Estatísticas e coeficientes das equações para estimativa do crescimento em volume (m<sup>3</sup> ha<sup>-1</sup>) para os clones de *Eucalyptus* plantados no Polo Gesseiro do Araripe-PE.

Nº	Equações	IA <sub>c</sub>	S <sub>xy</sub> %	REQM	EAM	SW	BP	DW
MC1	$\hat{V}_i = 136,9680 \cdot e^{(-56,8060 \cdot \frac{1}{I_i})}$	0,8452	20,3739	9,4389	6,5529	< 0,05	< 0,05	0,751
MC2	$\hat{V}_i = 3.506^{ns} \cdot (1 - 1,1111^{ns} \cdot I_i^{-0,0041^{ns}})^{1,5010^{ns}}$	0,8485	20,1583	9,2826	6,3986	< 0,05	< 0,05	0,715
MC3	$\hat{V}_i = \frac{I_i^2}{(22,8182 + 0,4480^{ns} \cdot I_i + 0,0056 \cdot I_i^2)}$	0,8434	20,4976	9,4676	6,9437	< 0,05	< 0,05	0,943
MC4	$\hat{V}_i = 109,6000 \cdot (1 - e^{-0,0019 \cdot I_i})^{1,9590}$	0,8423	20,5651	9,4988	6,9770	< 0,05	< 0,05	0,681
MC5	$\hat{V}_i = 0,0011^{ns} \cdot e^{-9,5409^{ns} \cdot 1,0016^{I_i}}$	0,7727	24,6907	11,4043	8,9901	< 0,05	< 0,05	0,258

REQM e EMA em m<sup>3</sup> ha<sup>-1</sup>, SW = Teste de Shapiro-Wilk, BW = Teste de Breush-Pagan, DW = Teste de Durbin-Watson (para todos os testes dos resíduos são apresentados os p-Valores), *ns* = não significativo pelo teste t ao nível de 5% de significância.

As equações provenientes dos modelos de Clutter-Jones e Prodan possuem coeficientes que não foram estatisticamente significantes, o que para predições em outras bases de dados pode acarretar em estimativas errôneas. Todas as equações apresentaram desvios quanto a normalidade dos resíduos e falta de homogeneidade da variância dos erros. A análise gráfica dos resíduos é dada na Figura 27.

Figura 27 – Análise dos resíduos para as equações clássicas de crescimento em volume de clones de *Eucalyptus*.



A normalidade, pressuposto da análise de regressão, em certo grau pode ser tolerável. Esta não introduz viés nas estimativas dos coeficientes, mas sim nas dos desvios padrões, o que pode afetar a validade dos resultados de testes de hipóteses (BERNIER; FENG; ASAKAWA, 2011; PINO, 2014). Em contraste com a normalidade,

problemas com a falta de homogeneidade podem ser mais críticos na modelagem, pois em geral quando o tamanho da amostra aumenta, o efeito da ausência de normalidade não afeta tão severamente as estimativas, já problemas com a homogeneidade das variâncias produzem estimativas viesadas mesmo com o aumento da amostra (YANG; TU; CHEN, 2019).

Todas as equações apresentaram a tendência de subestimar o crescimento volumétrico nos meses iniciais (Figura 27). Este mesmo comportamento foi observado por Guera et al. (2019b) quando ajustaram o modelo de Schumacher para estimar o crescimento de *Pinus caribaea* var. *caribaea* na província de Pinar del Río, Cuba. Já para o modelo de Chapman-Richards os autores encontram um comportamento de superestimativa.

A inclusão das variáveis climáticas foi realizada apenas nos modelos de Schumacher e Chapman-Richards que resultaram nas melhores equações. Desta forma, na Tabela 24 é mostrado as estatísticas de qualidade de ajuste e as equações com adição das variáveis climáticas.

Tabela 24 – Estatísticas e coeficientes das equações com adição de variáveis climáticas para estimativa do crescimento em volume ( $m^3 ha^{-1}$ ) para os clones de *Eucalyptus* plantados no Polo Gesseiro do Araripe-PE.

Nº	Equações	IA <sub>e</sub>	S <sub>xy</sub> %	REQM	EAM	SW	BP	DW
MC1	$\hat{V}_i = (-0,0569 \cdot PP + 0,5439 \cdot Tmax + 2,5588 \cdot TminAbs + 1,1295 \cdot UR) \cdot e^{(-52,5031 \cdot \frac{1}{T_i})}$	0,8534	19,8295	9,1033	6,4589	< 0,05	< 0,05	0,9466
MC4	$\hat{V}_i = (-0,0381 \cdot PP + 0,4566 \cdot Tmax + 1,3686 \cdot TminAbs + 0,7085 \cdot UR) \cdot (1 - e^{0,0369 \cdot T_i})^{3,2547}$	0,8488	20,1353	9,1868	6,5704	< 0,05	< 0,05	0,8973

REQM e EMA em  $m^3 ha^{-1}$ , SW = Teste de Shapiro-Wilk, BW = Teste de Breush-Pagan, DW = Teste de Durbin-Watson (para todos os testes dos resíduos são apresentados os p-Valores), PP = Precipitação (mm), TMax = temperatura máxima média (°C), TMinAbs = temperatura mínima absoluta (°C), UR = umidade relativa (%).

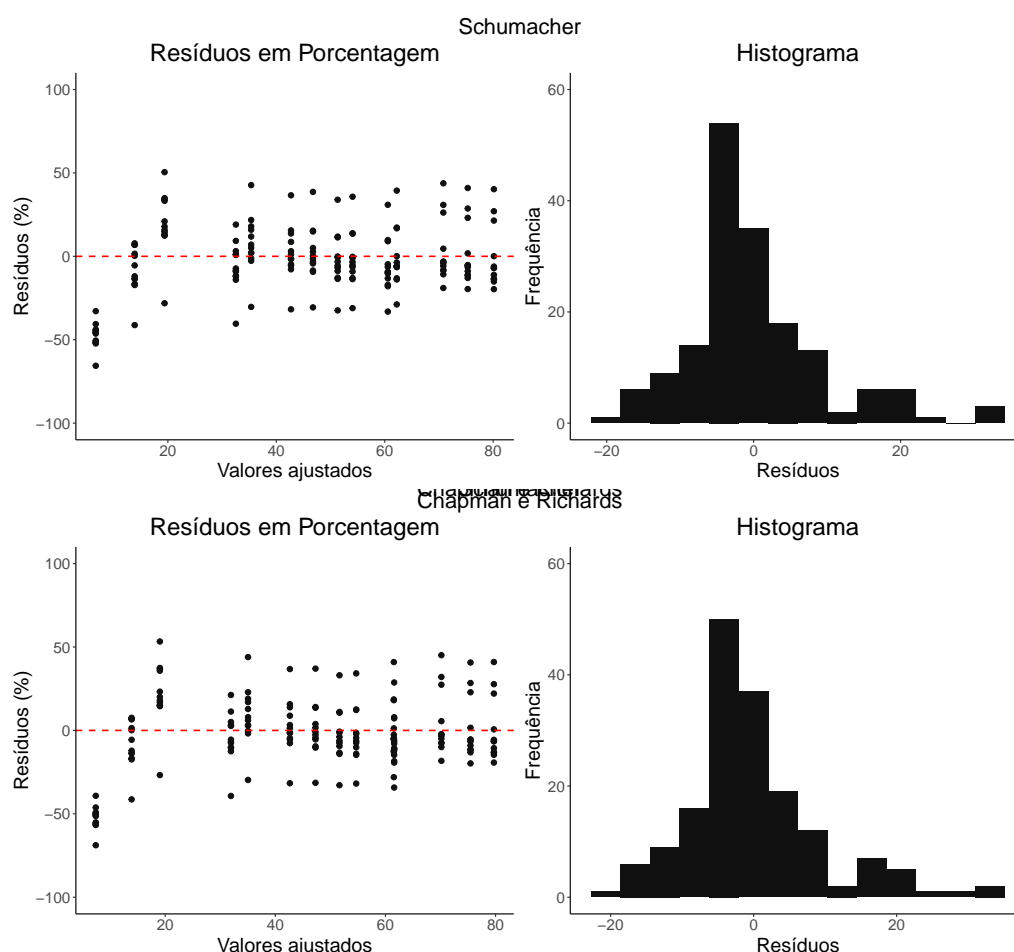
O fator modificador linear  $Mod. = \kappa_1 \cdot VC_1 + \kappa_1 \cdot VC_2 + \dots + \kappa_n \cdot VC_n$  foi o que permitiu os melhores ajustes. A terceira forma de inclusão do fator modificador (*Mod.*) não foi eficiente e se mostrou de difícil convergência.

A inclusão das variáveis climáticas da região permitiu uma melhora na acurácia das equações de crescimento, principalmente no modelo de Schumacher. A adição destas variáveis relacionadas com as condições do clima acontece principalmente pelo fato da região apresentar valores extremos opostos nos semestres de medição. Para o primeiro semestre (outubro a março), a Temperatura máxima (alta) e a umidade relativa (baixa) do ar chegam aos seus valores mais extremos. Já no segundo semestre (abril à setembro), a precipitação mensal chegou aos seus maiores níveis, bem como é

registrado os menores valores para a temperatura mínima. Desta forma, as variáveis utilizadas nas equações conseguem explicar as condições climáticas comuns aos dois semestres de medição. No entanto, se as variáveis fossem coletadas na escala mensal os efeitos climáticos poderiam ser melhor expressados e os ganhos na acurácia poderiam ser maiores.

As equações obtidas ainda apresentam problemas quanto a normalidade e homogeneidade da variância dos resíduos (Figura 28).

Figura 28 – Análise dos resíduos para as equações com adição de variáveis climáticas para estimativa do crescimento em volume de clones de *Eucalyptus*.



Elli et al. (2017) também conseguiram resultados satisfatórios para equação de crescimento com adição de variáveis climáticas quando trabalharam com as seguintes espécies: *Eucalyptus urophylla* S.T. Blake x *Eucalyptus grandis* Hill ex Maiden, *Mimosa scabrella* Benth., *Parapiptadenia rigida*, *Peltophorum dubium* (Spr.) Taubert e *Schizolobium parahybae* (Vell.) Blake. Para estas equações, os autores incluíram as variáveis precipitação (mm), temperatura mínima (°C) e temperatura mínima absoluta (°C). A

entrada das variáveis de temperatura mínima nas equações possivelmente se deu pelo fato dos autores terem conduzido o experimento na região Sul onde o frio pode ser um fator condicionante para desenvolvimento das espécies.

Por outro lado, Zamin et al. (2013) não obtiveram melhoras na precisão quando adicionaram variáveis climáticas ao modelo de Chapman-Richards para estimativa do crescimento de *Araucaria angustifolia* e *Pinus taeda* em fase juvenil. O fato dos autores não obterem sucesso na estimativa pode estar ligado a forma de como o fator modificador foi modelado, já que eles empregaram o fator que apresentou os piores resultados neste estudo.

### 7.3.3 Modelagem do crescimento com algoritmos de aprendizagem de máquinas

Foram treinadas 295 RNA's para estimativa do crescimento em volume sem a inclusão de variáveis climáticas. A melhor RNA foi obtida com os seguintes hiperparâmetros: neurônios (1:3:1), com a função de ativação *Rectifier* na camada oculta e a função linear na camada de saída; 750 épocas, taxas de regularização L1 e L2 de 0,001; taxa de aprendizagem de 0,1 e *momentum* de 0,1. Para o algoritmo de MVS as melhores configurações foram: função de *kernel radial*, parâmetro custo de 11,3137, *gamma* de 0,4 e *epsilon* de 0,29. O melhor ajuste obtido com a floresta aleatória possui os seguintes hiperparâmetros: número de árvores = 1000, número de folhas = 14 e profundidade = 5,5. Todas as estatísticas de qualidade de ajuste estão contidas na Tabela 25.

Tabela 25 – Resumo estatístico dos algoritmos de aprendizagem de máquinas para estimativa do crescimento em volume de clones de *Eucalyptus* spp.

Algoritmos de aprendizagem de máquinas	Validação				Teste	
	IA <sub>c</sub>	S <sub>xy</sub> %	EAM	REQM	EAM	REQM
Redes Neurais Artificiais	0,8575	19,8071	5,9962	8,9882	6,6563	9,7635
Floresta Aleatória	-	-	6,4522	9,7834	7,8030	10,1050
Máquinas de Vetores de Suporte	0,8033	23,2730	6,1543	9,2409	6,5174	9,7713

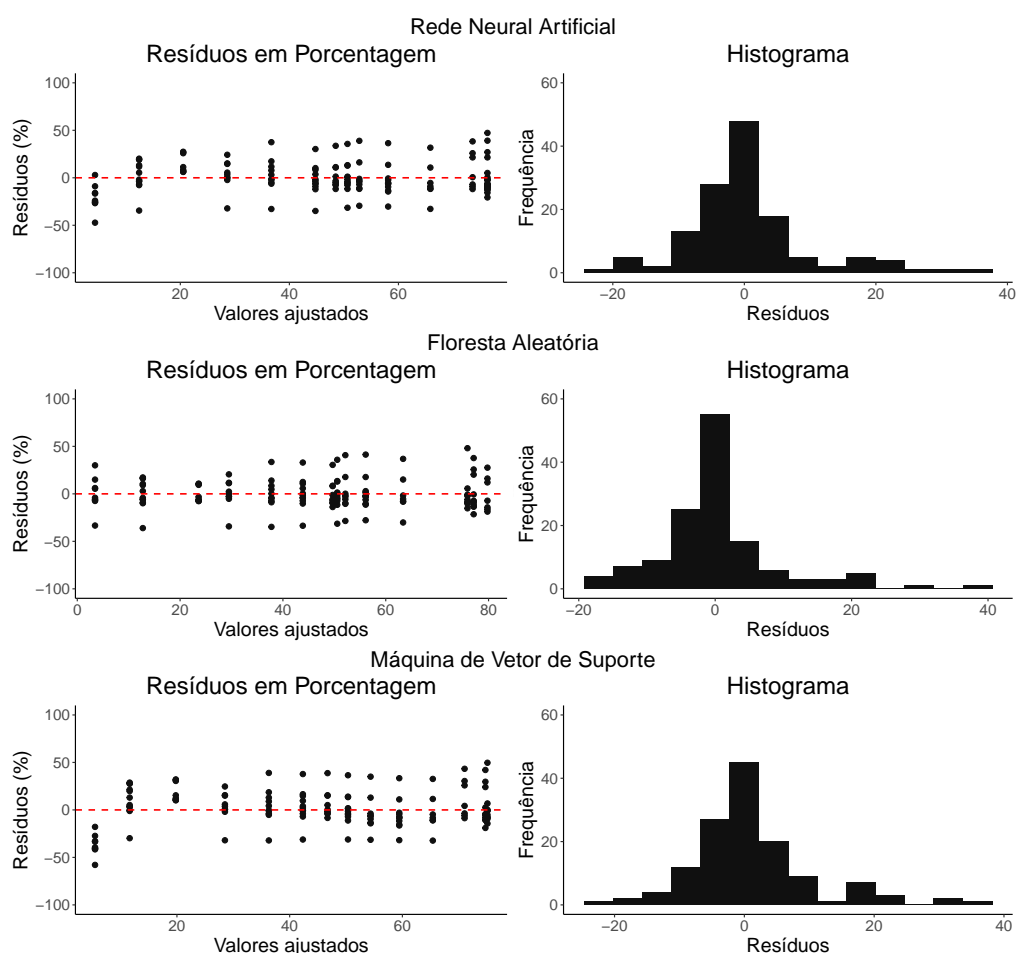
REQM e EMA em m<sup>3</sup> ha<sup>-1</sup>

Os algoritmos de aprendizagem de máquinas apresentaram resultados seme-

lhantes aos encontrados por meio dos modelos biológicos clássicos. No entanto, as vantagens desses são a robustez quanto aos pressupostos da análise de regressão impostos aos modelos de crescimento (ÖZÇELİK et al., 2017). Entretanto, a principal desvantagem dos algoritmos de aprendizagem de máquinas é que estes não possuem uma fácil interpretação biológica não sendo possível derivar outras informações sobre o crescimento como é factível com os modelos biológicos.

O mesmo comportamento de subestimativa apresentado nas equações de crescimento foi observado nos algoritmos RNA e MVS, diferente do que foi encontrado com o algoritmo FA que mostrou comportamento dos resíduos aleatório na linha de base zero. Os gráficos dos resíduos são apresentados na Figura 29.

Figura 29 – Análise dos resíduos para os algoritmos de aprendizagem de máquinas treinados para estimar o crescimento em volume de clones de *Eucalyptus*.



Para avaliação da inclusão das variáveis climáticas foram treinadas 300 RNA's. A RNA que apresentou as melhores estatísticas de qualidade possui os seguintes hiperparâmetros: neurônios (5:4:1), com a função de ativação *Rectifier* na camada

oculta e a função linear na camada de saída; 1000 épocas, taxas de regularização L1 de  $10^{-6}$ ; taxa de aprendizagem de 0,1 e *momentum* de 0,1. Quando foi adicionado as variáveis climáticas no algoritmo de MVS, as melhores configurações foram: função de *kernel radial*, parâmetro custo de 1,4142, *gamma* de 0,1 e *epsilon* de 0,25. O algoritmo de floresta aleatória selecionado possui os seguintes hiperparâmetros: número de árvores = 1000, número de folhas = 14 e profundidade = 6. O testes realizados para comparação dos algoritmos estão presentes na Tabela 26.

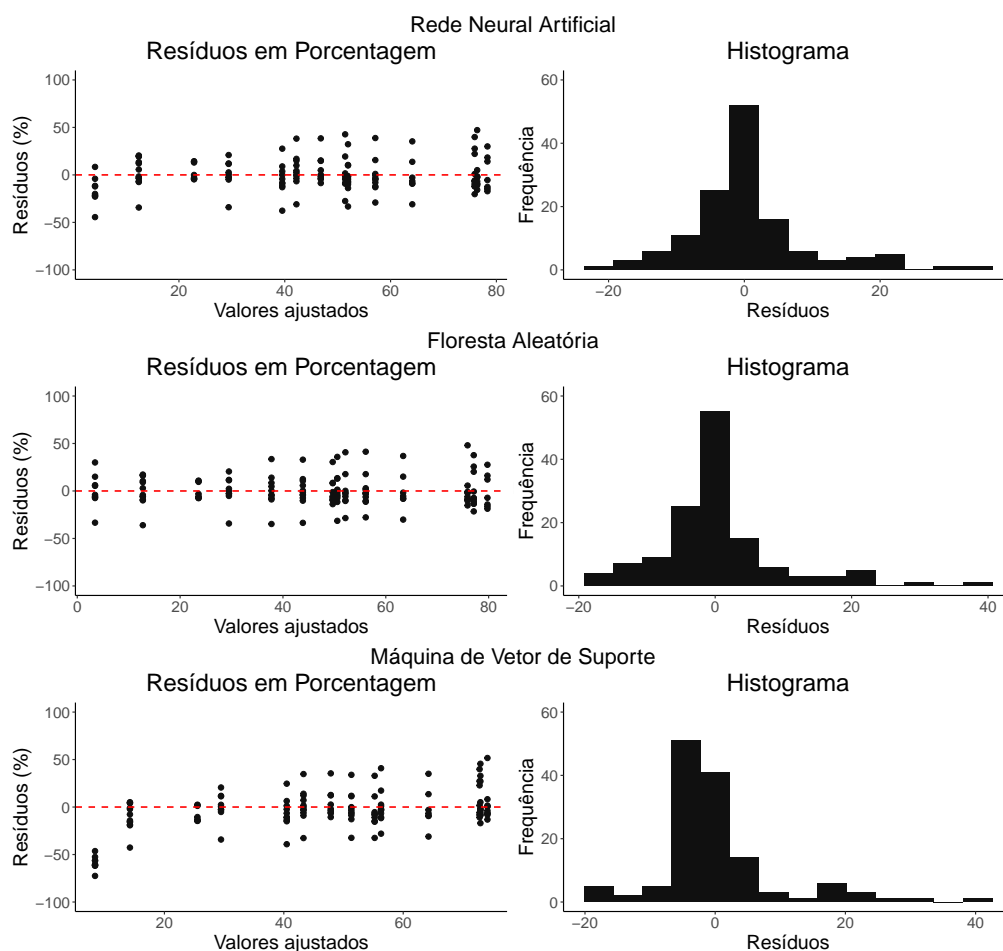
Tabela 26 – Resumo estatístico dos algoritmos de aprendizagem de máquinas com adição de variáveis climáticas para estimativa do crescimento em volume de clones de *Eucalyptus* spp.

Algoritmos de aprendizagem de máquinas	Validação				Teste	
	IA <sub>c</sub>	S <sub>xy</sub> %	EAM	REQM	EAM	REQM
Redes Neurais Artificiais	0,8392	21,0458	5,7588	8,8534	6,8726	10,2321
Floresta Aleatória	-	-	6,5147	9,7706	7,7790	10,1010
Máquinas de Vetores de Suporte	0,78	24,2482	6,0526	9,1237	6,5174	9,7713

REQM e EMA em m<sup>3</sup> ha<sup>-1</sup>

Os três algoritmos de aprendizagem de máquinas apresentam resultados semelhantes. As MVS's apresentaram valores de estatísticas de qualidade melhores no conjunto de teste em relação aos demais algoritmos. A inclusão de variáveis nos algoritmos de aprendizagem de máquinas não proporcionou ganho significativo no desempenho. Isso pode estar relacionado com o tempo decorrido entre coletas, possivelmente a realização de coletas mensais poderia ajudar no entendimento do efeito das variáveis climáticas sobre o crescimento em volume de madeira. Os resíduos para os algoritmos de RNA, MVS e FA com adição das variáveis climáticas possuem distribuição semelhante (Figura Figura 30).

Figura 30 – Análise dos resíduos para os algoritmos de aprendizagem de máquinas com adição de variáveis climáticas treinados para estimar o crescimento em volume de clones de *Eucalyptus*.



Pela análise gráfica dos resíduos não foram detectados outliers para as estimativas realizadas por meio dos três algoritmos inteligentes.

## 7.4 Conclusões

A inclusão de variáveis climáticas nas equações obtidas dos modelos de Schumacher e Chapman-Richards permitiu melhores estimativas do crescimento em volume de clones de *Eucalyptus* spp. plantados na região semiárida pernambucana.

Na estimativa do crescimento volumétrico de clones *Eucalyptus*, os algoritmos de aprendizagem de máquinas RNA, MVS e FA apresentaram resultados semelhantes aos modelos de crescimento clássicos. Os algoritmos inteligentes são robustos a falta de normalidade, heterogeneidade e autocorrelação dos resíduos se mostrando apropriados para problemas florestais onde estas situações são comuns.





## Referências Bibliográficas

- ABREU, J. C. d. **Alternativas para estimar o volume de árvores individuais em formações florestais no estado de Minas Gerais**. Tese (Ciência Florestal) — Universidade Federal de Viçosa, Viçosa, MG, 2019.
- BARRAZA, G. Q. N. **Sistema de crecimiento y rendimiento para las principales especies comerciales de *Pinus* en Durango, México**. 194 p. Tese (Doctor en Ciencias) — Institución de Enseñanza e Investigación en Ciencias Agrícolas, Montecillo, Texcoco, Edo. de Mexico, 2014.
- BERNIER, J.; FENG, Y.; ASAKAWA, K. Strategies for handling normality assumptions in multi-level modeling: a case study estimating trajectories of health utilities index mark 3 scores. **Health Rep**, v. 22, p. 45–51, 2011.
- BHERING, L. L. et al. Application of neural networks to predict volume in eucalyptus. **Crop Breeding and Applied Biotechnology**, v. 15, n. 3, p. 125–131, 2015.
- BINOTI, D. H. B. et al. Support vector machine to estimate volume of eucalypt trees. **Revista Árvore**, v. 40, n. 4, p. 689–693, 2016.
- ELLI, E. F. et al. Climatic factors defining the height growth curve of forest species. **iForest-Biogeosciences and Forestry**, v. 10, n. 3, p. 547, 2017.
- FERRAZ-FILHO, A. C. et al. Dominant height projection model with the addition of environmental variables. **Cerne**, v. 17, n. 3, p. 427–433, 2011.
- GADELHA, F. H. L. **Desempenho silvicultural e avaliação econômica de clones híbridos de eucaliptos plantados em diferentes regimes de manejo para fins energéticos**. 147 p. Tese (Ciências Florestais) — Universidade Federal Rural de Pernambuco, Departamento de Ciências Florestais, Recife, 2014.
- GARCÍA-GUTIÉRREZ, J. et al. A comparison of machine learning regression techniques for lidar-derived estimation of forest variables. **Neurocomputing**, v. 167, p. 24–31, 2015.
- GEDEON, T. D. Data mining of inputs: analysing magnitude and functional measures. **International Journal of Neural Systems**, v. 8, n. 02, p. 209–218, 1997.

GUERA, O. G. M. et al. Alternatives to growth and yield prognosis for *Pinus caribaea* var. *caribaea* barrett & golfari. **Floresta e Ambiente**, v. 26, n. 4, 2019b.

HUTCHINS, D. A. et al. Climate change microbiology—problems and perspectives. **Nature Reviews Microbiology**, v. 17, n. 6, p. 391–396, 2019.

LACERDA, F. F. et al. Alterações climáticas globais; uma realidade em pernambuco. **Anais da Academia Pernambucana de Ciência Agronômica**, v. 11/12, p. 121–154, 2014/2015.

LOTZE, H. K. et al. Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. **Proceedings of the National Academy of Sciences**, v. 116, n. 26, p. 12907–12912, 2019.

MACHADO, S. d. A. et al. Efeito de variáveis climáticas no crescimento mensal de *Pinus taeda* e *Araucaria angustifolia* em fase juvenil. **Floresta e ambiente**, v. 21, n. 2, p. 170–181, 2014.

MAESTRI, R. et al. Modelagem do crescimento florestal considerando variáveis do ambiente: Revisão. **Scientia Agraria**, v. 14, n. 3, p. 103–110, 2013.

OLANIYI, O.; OLUTIMEHIN, I.; FUNMILAYO, O. Review of climate change and its effect on nigerian ecosystem. **International journal of Rural Development, Environment and Health Research**, v. 3, n. 3, 2019.

ÖZÇELİK, R. et al. Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. **Forest Science**, v. 63, n. 3, p. 291–302, 2017.

PINO, F. A. A questão da não normalidade: Uma revisão. **Revista de economia agrícola**, v. 61, n. 2, p. 17–33, 2014.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Disponível em: <<https://www.R-project.org/>>.

SCOLFORO, J. R. S. et al. Dominant height model for site classification of eucalyptus grandis incorporating climatic variables. **International Journal of Forestry Research**, v. 2013, 2013.

SILVA, J. A. A.; NETO, F. d. P. **Princípios básicos de dendrometria**. [S.l.]: Imprensa Universitária da UFRPE, 1979. 198 p.

YANG, K.; TU, J.; CHEN, T. Homoscedasticity: an overlooked critical assumption for linear regression. **General psychiatry**, v. 32, n. 5, 2019.

ZAMIN, N. T. et al. Effect of climate variables on monthly growth in modeling biological yield of araucaria angustifolia and pinus taeda in the juvenile phase. **International Journal of Forestry Research**, v. 2013, 2013.

## CAPÍTULO 8

---

### **Modelos não lineares e algoritmos de aprendizagem de máquinas para prognose do volume de clones de *Eucalyptus* spp. com adição de variáveis climáticas**

#### **Resumo**

Objetivou-se com esta pesquisa ajustar modelos biológicos não lineares e avaliar algoritmos de aprendizagem de máquinas para prognose do volume de clones de *Eucalyptus* spp. com adição de variáveis climáticas. Os dados foram coletados em um experimento na Estação Experimental do Instituto Agronômico de Pernambuco (IPA), no Polo Gesseiro do Araripe-PE. Foram realizadas coletas em períodos de 6 meses até a idade de 96 meses, as variáveis obtidas foram altura e diâmetro à altura do peito, bem foram realizadas amostragens ao longo do tempo para cubagem de árvores e cálculo do fator de forma, sendo obtido o volume por árvore em cada período de medição. As variáveis climáticas foram obtidas de uma estação meteorológica localizada há 1,27 km de distância do experimento. Foram ajustados modelos biológicos clássicos para prognose e treinado os seguintes algoritmos de aprendizagem de máquinas: redes neurais artificiais (RNA), máquinas de vetores de suporte (MVS) e florestas aleatórias (FA). A análise da qualidade dos ajustes foi realizada por meio dos testes: Índice de ajuste de Schlaegel corrigido ( $IA_c$ ), erro padrão de estimativa em porcentagem ( $S_{xy}$  %), raiz do erro quadrático médio (REQM  $m^3 ha^{-1}$ ) e erro absoluto médio (EAM  $m^3 ha^{-1}$ ). Bem como, análise dos resíduos. As melhores equações para prognose do volume de clones de *Eucalyptus* foram obtidas dos modelos de Mitscherlich, Brito-Silva e Clutter-Jones. O algoritmo de FA apresentou as melhores estatísticas de qualidade dentre os demais algoritmos de aprendizagem de máquinas e estatísticas semelhantes aos modelos clássicos. A adição de variáveis climáticas só proporcionou os melhores ajustes nos modelos biológicos clássicos.

**Palavras-chave:** Polo Gesseiro do Araripe-PE, Brito-Silva, Mitscherlich, redes neurais artificiais, máquinas de vetores de suporte, florestas aleatórias.

## Abstract

The objective of this research was to adjust non-linear biological models and to evaluate Artificial Intelligence techniques for prognosis of the volume of clones of *Eucalyptus* spp. with the addition of climatic variables. The data were collected in an experiment at the Experimental Station of the Agronomic Institute of Pernambuco (IPA), at the Gypsum Pole of Araripe-PE. Collections were carried out in periods of 6 months up to the age of 96 months, the variables obtained were height and diameter at breast height. Sampling was performed over time for cubing trees and calculating the form factor, obtaining the volume per tree in each measurement period. The climatic variables were obtained from a meteorological station located 1.27 km away from the experiment. Classic biological models for prognosis were adjusted and the following intelligent algorithms were trained: artificial neural networks (ANN), support vector machines (SVM) and random forests (RF). The analysis of the quality of the adjustments was performed using the tests: Corrected Schlaegel adjustment index ( $AI_c$ ), standard error of percentage estimate ( $S_{xy}$  %), root of the mean square error (RMSE  $m^3 ha^{-1}$ ) and mean absolute error (MAE  $m^3 ha^{-1}$ ). As well as residues analysis. The best equations for prognosis of the volume of *Eucalyptus* clones were obtained from the Mitscherlich, Brito-Silva and Clutter-Jones models. The RF algorithm presented the best quality statistics among the other intelligent algorithms and similar to the classic models. The addition of climatic variables only provided the best adjustments in classic biological models.

**Key-words:** Gypsum Pole of Araripe-PE, Brito-Silva, Mitscherlich, artificial neural network, support vector machines, random forest.

## 8.1 Introdução

A prognose da produção está inserida em um conjunto de três elementos fundamentais do manejo florestal (CAMPOS; LEITE, 2017). Diferente dos modelos de crescimento em que a variável de interesse (volume, altura, diâmetro, etc) é modelada em função apenas do tempo, nos modelos de prognose a variável de interesse é modelada em função do seu valor inicial e do tempo inicial e final (BINOTI et al., 2015). Quanto a terminologia da palavra em modelos de crescimento se tem a predição e nos modelos de prognose se tem a previsão ou projeção.

Modelos de prognose são bastantes difundidos na literatura de ciências florestais (SILVA, 2014/2015; SILVA et al., 2016/2017; CASTRO et al., 2017). Este modelos

geralmente utilizam as variáveis índice de sítio, área basal e idade para representar o efeito do ambiente e dos tratos silviculturais no desenvolvimento das árvores em um povoamento florestal (SILVA, 2014/2015). Dentre os modelos mais empregados na prognose de povoamentos florestais estão os Chapman-Richards e Clutter-Jones (CAMPOS; LEITE, 2017), no entanto, outros modelos com grande potencial foram desenvolvidos como é o caso dos modelos de Silva e Bailey (SILVA, 1986) e Brito e Silva (BRITO et al., 2007). A área da modelagem na biometria florestal está em contante desenvolvimento.

Nos últimos anos tem crescido o uso de algoritmos de aprendizagem de máquinas em problemas florestais causado principalmente pela busca de estimativas com maior acurácia (DIAMANTOPOULOU et al., 2015; BINOTI et al., 2016; GUERA et al., 2019b). Os principais algoritmos de aprendizagem de máquinas empregadas são redes neurais artificiais (RNA), máquinas de vetores de suporte (MVS) e florestas aleatórias (FA), estas já foram reportadas em diversos estudos (GÖRGENS; MONTAGHI; RODRIGUEZ, 2015; ÖZÇELIK et al., 2017; SOUZA et al., 2018; REIS et al., 2019; ABREU, 2019).

A adição de variáveis climáticas em modelos de prognose foi reportada em algumas pesquisas (MAESTRI, 2003; FERRAZ-FILHO et al., 2011). Já em algoritmos de aprendizagem de máquinas o estudo destas variáveis foi pouco explorado. A inclusão de variáveis climáticas em modelos biológicos pode proporcionar ganhos na estimativa principalmente por adicionarem as variações relacionadas ao clima local, o que não acontece em modelos clássicos (ELLI et al., 2017).

Desta forma, objetivou-se com esta pesquisa ajustar modelos biológicos não lineares e avaliar algoritmos de aprendizagem de máquinas para prognose do volume de clones de *Eucalyptus* spp. com adição de variáveis climáticas.

## **8.2 Material e Métodos**

### **8.2.1 Obtenção dos dados dendrométricos e climáticos**

Os dados dendrométricos são provenientes de um povoamento experimental com 96 meses de idade realizado na Estação Experimental do Instituto Agrônomo de Pernambuco (IPA) (coordenadas 07° 28' 04,2" S e 40° 24' 37,7" W, altitude de 816 metros), no Polo Gesseiro do Araripe-PE. Todos os detalhes da realização do experimento e da área experimental são detalhados no Capítulo 7.

Todas as informações referentes as variáveis meteorológicas estão contidas

no Capítulo 7. Bem como, a metodologia de seleção das variáveis para comporem os modelos modificados.

### 8.2.2 Modelos estatísticos para prognose

Para prognose do volume de madeira do povoamento foram empregados os modelos que estão dispostos na Tabela 27.

Tabela 27 – Modelos de prognose ajustados para os clones de *Eucalyptus* plantados no Polo Gesseiro do Araripe-PE.

Nº	Modelos	Autores
MP1	$V_f = V_i - \beta_0 \cdot (\kappa^{t_f} - \kappa^{t_i}) + \epsilon_i$	Mitscherlich
MP2	$V_f = V_i - \frac{e^{(\beta_0 \cdot e^{\kappa \cdot t_f} + \beta_1 \cdot e^{\lambda \cdot t_f})}}{e^{(\beta_0 \cdot e^{\kappa \cdot t_i} + \beta_1 \cdot e^{\lambda \cdot t_i})}} + \epsilon_i$	Brito e Silva
MP3	$V_f = V_i \cdot e^{\beta_0(\kappa t_f^\kappa - \kappa t_i^\kappa)} + \epsilon_i$	Silva e Bailey
MP4	$V_f = V_i \cdot \left(\frac{1 - e^{\kappa t_f}}{1 - e^{\kappa t_i}}\right)^{\beta_0} + \epsilon_i$	Chapman e Richards
MP5	$V_f = V_i \cdot \left(\frac{1 - e^{\kappa t_f}}{1 - e^{\kappa t_i}}\right)^3 + \epsilon_i$	Bertalanffy
MP6	$V_f = V_i - \beta_0 \cdot \left(e^{-\kappa t_f^{\beta_1}} - e^{-\kappa t_i^{\beta_1}}\right) + \epsilon_i$	Weibull
MP7	$V_f = \left[V_i^{\beta_0} + \beta_1 \cdot (t_f^\kappa - t_i^\kappa)\right]^{\frac{1}{\beta_0}} + \epsilon_i$	Clutter e Jones

Em que:  $V_f$  = volume final ( $\text{m}^3 \text{ha}^{-1}$ );  $V_i$  = volume inicial ( $\text{m}^3 \text{ha}^{-1}$ );  $t_f$  = tempo final (meses);  $t_i$  = tempo inicial (meses);  $\kappa$ ,  $\lambda$ ,  $\beta_0$  e  $\beta_1$  = parâmetros que serão ajustados;  $\epsilon_i$  = erro aleatório.

A primeira forma de para adição das variáveis climáticas foi realizada conforme proposto por Maestri (2003) e Ferraz-Filho et al. (2011). Considerando que o incremento do volume nos modelos de prognose está relacionado com o parâmetro de inclinação do modelo, foi atribuído um fator modificador linear para este parâmetro. O fator modificador (*Mod.*) foi definido por:

$$Mod. = \alpha_1 \cdot VC_1 + \alpha_2 \cdot VC_2 + \dots + \alpha_n \cdot VC_n \quad (8.1)$$

Considerando o modelo de Mitscherlich, temos:

$$V_f = V_i - \beta_0 \cdot \left[(\kappa \cdot Mod.)^{t_f} - (\kappa \cdot Mod.)^{t_i}\right] + \epsilon_i \quad (8.2)$$



Em que:

$Mod.$  = fator modificador;  $\alpha_1, \alpha_2, \alpha_n$  = parâmetros da regressão;  $VC_1, VC_2, VC_n$  = variáveis climáticas definidas na subseção 7.2.2. Para os demais modelos a inclusão ocorreu de forma semelhante.

Considerando que a mudança no parâmetro de inclinação pode alterar o comportamento biológico do modelo o fator modificador foi relacionado com o parâmetro assintótico do modelo, conforme:

$$V_f = V_i - Mod. \cdot (\kappa^{t_f} - \kappa^{t_i}) + \epsilon_i \quad (8.3)$$

### 8.2.3 Prognose do volume com algoritmos de aprendizagem de máquinas

Para prognose do volume foram treinados os seguintes algoritmos: Redes Neurais Artificiais, Florestas Aleatórias e Máquinas de Vetores de Suporte, tal como foram definidos no Capítulo 5. Todos procedimentos de treinamento foram realizados conforme discutido nos Capítulos 5 e 7. As variáveis de entrada empregadas nos algoritmos foram: volume inicial ( $m^3 \text{ ha}^{-1}$ , referente ao tempo anterior ao que se quer prever), idade inicial (meses), idade final (meses) e as variáveis climáticas discutidas anteriormente. O volume final ( $m^3 \text{ ha}^{-1}$ ) era o que se pretendia prever.

### 8.2.4 Avaliação das estimativas

As avaliações das equações e dos algoritmos de aprendizagem de máquinas foram realizadas por meio dos testes: Índice de ajuste de Schlaegel corrigido ( $IA_c$ ), erro padrão de estimativa em porcentagem ( $S_{xy} \%$ ), raiz do erro quadrático médio (REQM  $m^3 \text{ ha}^{-1}$ ) e erro absoluto médio (EAM  $m^3 \text{ ha}^{-1}$ ). Para os modelos estatísticos os seguintes pressupostos foram avaliados: normalidade dos resíduos (teste de Shapiro-Wilk), homogeneidade das variâncias (teste de Breush-Pagan) e autocorrelação dos resíduos (teste de Durbin-Watson). Também foi realizada a análise gráfica dos resíduos para as duas propostas de modelagem. Os demais detalhes dos testes e das metodologias de avaliações foram apresentados no Capítulo 5. Todas as análises foram realizadas na linguagem de programação R (R Core Team, 2016).

### 8.3 Resultados e Discussão

#### 8.3.1 Modelos estatísticos para prognose do volume

Os coeficientes das equações e as estatísticas para avaliar a qualidade de ajuste são expostos na Tabela 28. De forma geral, todas as equações geradas possuem elevados valores de índice de ajuste ( $IA_c$ ), baixo erro padrão de estimativa ( $S_{xy}\%$ ) e a raiz do erro quadrático médio (REQM) também com baixos valores. Por estes critérios, todas as equações podem ser utilizadas na estimativa do crescimento volumétrico.

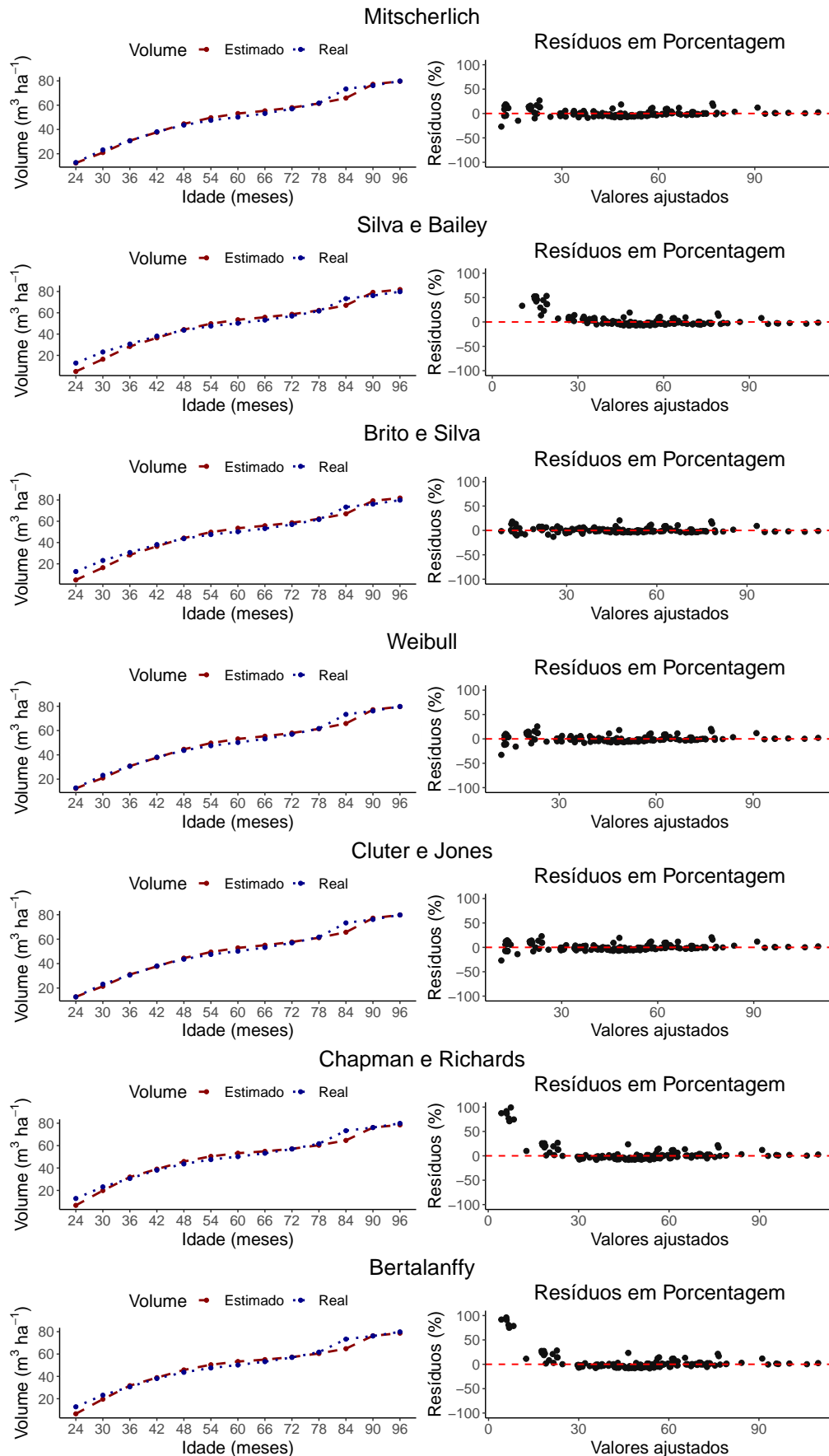
Tabela 28 – Estatísticas e coeficientes das equações para prognose do volume ( $m^3 ha^{-1}$ ) dos clones de *Eucalyptus* implantados no Polo Gesseiro do Araripe-PE.

Equações	Coeficientes				$IA_c$	$S_{xy}\%$	REQM	EAM	SW	BP	DW
	$\hat{r}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\lambda}$							
Mitscherlich	0,9874	151,7	-	-	0,9807	6,0472	3,0	2,0653	<0,05	<0,05	<0,05
Brito e Silva	-0,1423	-26,7292	-0,0045 <sup>ns</sup>	-3,8207 <sup>ns</sup>	0,9855	5,2924	2,6085	1,7067	<0,05	<0,05	<0,05
Silva e Bailey	-1,3802	125,6431	-	-	0,9778	6,5106	3,2300	2,0824	<0,05	0,235	<0,05
Clutter e Jones	0,0362 <sup>ns</sup>	0,7796	398,4 <sup>ns</sup>	-	0,9821	5,8556	2,8956	1,9203	<0,05	<0,05	<0,05
Weibull	-1,44E <sup>-5ns</sup>	-1,86E <sup>6ns</sup>	0,3983	-	0,9819	5,8925	2,9138	1,9830	<0,05	0,05	<0,05
Chapman e Richards	-0,0329	3,1790	-	-	0,9706	7,4952	3,7184	2,6149	<0,05	<0,05	<0,05
Bertalanffy	-0,0311	-	-	-	0,9705	7,4736	3,7197	2,6138	<0,05	<0,05	<0,05

REQM e EMA em  $m^3 ha^{-1}$ , SW = Teste de Shapiro-Wilk, BW = Teste de Breush-Pagan, DW = Teste de Durbin-Watson (para todos os testes dos resíduos são apresentados os p-Valores), *ns* = não significativo pelo teste t ao nível de 5% de significância.

Apesar de todas as equações apresentarem boas estimativas, algumas se mostraram superiores para estimativa do volume. As equações de Chapman e Richards e a de Bertalanffy possuem o desempenho mais pobre dentre as demais e são bem semelhantes entre si. Isto se explica pelo fato do modelo de Chapman e Richards ser uma generalização do de Bertalanffy (SILVA, 2014/2015). O modelo de Chapman e Richards é mais flexível apenas por permitir ajustar o parâmetro  $\beta_0$ , enquanto que no de Bertalanffy este parâmetro é a constante 3. E segundo Silva (1986) esta pequena modificação permite o ajuste do modelo para diferentes fenômenos da natureza. A equação de Bertalanffy ainda apresentou um erro maior ao estimar o volume durante a fase jovem, conforme é visto no gráfico dos resíduos na Figura 31.

Figura 31 – Curvas de crescimento real e estimada e distribuição dos resíduos para as equações de prognose do volume de clones de *Eucalyptus*.



A equação de Silva e Bailey possui baixos valores de  $S_{xy}\%$ , RMSE e EAM, porém para o volume inicial apresentou um erro elevado. No entanto, é importante destacar que em geral modelos de crescimento tendem a superestimar ou subestimar a variável resposta (volume, altura ou DAP) nas idades iniciais, como exemplo, pode-se citar os trabalhos de Brito et al. (2007) e Martins, Soares e Silva (2014). O que justifica esta perda de eficiência na fase inicial é alta taxa de crescimento das plantas quando são jovens. Com isto, tem-se um aumento da variabilidade do volume nesta fase.

Os modelos que proporcionaram as melhores equações foram os de Mitscherlich, Brito-Silva, Clutter-Jones e Weibull. Estas equações são bastante semelhantes, com os valores das estatísticas de qualidade bem próximos. Silva et al. (2016/2017), trabalharam com clones de *Eucalyptus urophylla* no Polo Gesseiro do Araripe, também encontraram bons ajustes para esses modelos. Cabe ressaltar que os autores citados acima alertaram para o detalhe de que o coeficiente  $\hat{\beta}_0$  de algumas dessas equações possuíam um valor de difícil interpretação biológica, o que também aconteceu nesta pesquisa. Mesmo diante desses bons ajustes alguns coeficientes não são estatisticamente significativos ao nível de 5% de probabilidade de erro pelo teste t. Quando a estimativa dos parâmetros apresenta valores de difícil interpretação, pode ser um indício de falha na convergência do algoritmo iterativo.

Ao se observar o gráfico dos resíduos em porcentagem é possível destacar que todos os modelos apresentaram uma tendência em subestimar o volume em um determinado momento do crescimento. Analisando os gráficos com as curvas de crescimento se pode notar que houve um aumento na taxa de crescimento entre as idades de 78 e 90 meses. Este aumento na taxa de crescimento provavelmente é o que causa os valores de subestimativas em torno dos meses destacados. A principal razão para o que aconteceu está relacionado com os níveis de precipitação. Com base no gráfico de precipitação anual apresentado na subseção 6.3.1, figura 18, tem-se que ocorreu um pequeno aumento nos níveis de precipitação nos anos de 2015 e 2016 o que pode ter acarretado em um aumento da taxa de crescimento.

Apesar dos ótimos valores de estatística da qualidade para as equações desenvolvidas, nenhuma delas apresentou normalidade para os resíduos, conforme foi analisado pelo teste de Shapiro-Wilk com 5% de probabilidade de erro. Bem como, todas as equações apresentaram autocorrelação dos resíduos e apenas as equações obtidas por meio dos modelos de Silva-Bailey e Weibull possuem homogeneidade na variância dos resíduos. Mesmo esse pressuposto não sendo muito levado em consideração, é extremamente importante ser atendido. Pois, equações com falha no atendimento aos pressupostos da regressão podem levar a estimativas errôneas.

Sendo assim, é importante desenvolver novos modelos que sejam robustos a estes problemas.

Após avaliação das equações para prognose do volume, as três melhores foram selecionadas para serem reajustadas com a inclusão das variáveis climáticas. Desta forma, os modelos selecionados foram Mitscherlich, Brito-Silva e Weibull. Foi dado preferência ao modelo de Weibull no lugar de Clutter-Jones, pois este possui uma estrutura menos complexa e apresentou ajustes semelhantes. As estatísticas da qualidade dos ajustes para as equações modificadas são mostradas na Tabela 29. Cabe ressaltar que a modificação do parâmetro de inclinação do modelo não proporcionou ajustes superiores e ainda causou problemas com a convergência dos algoritmos iterativos e os coeficientes encontrados não apresentaram significância estatística.

Tabela 29 – Estatísticas e coeficientes das equações com adição de variáveis climáticas para prognose do volume ( $\text{m}^3 \text{ha}^{-1}$ ) para os clones de *Eucalyptus* plantados no Polo Gesseiro do Araripe-PE.

Nº	Equações	IA <sub>c</sub>	S <sub>xy</sub> %	REQM	EAM	SW	BP	DW
MP1	$\hat{V}_f = V_i - (-0,1646 \cdot PP + 8,2647 \cdot Tmax) \cdot (0,9914^{t_f} - 0,9914^{t_i})$	0,9828	5,7271	2,8320	1,9220	< 0,05	< 0,05	0,3731
MP2	$\hat{V}_f = V_i - \frac{e^{[(0,0192 \cdot PP - 1,0570 \cdot Tmax) \cdot e^{-0,1452 \cdot t_f} - 3,0944 \cdot e^{-0,0067 \cdot t_f}]}{e^{[(0,0192 \cdot PP - 1,0570 \cdot Tmax) \cdot e^{-0,1452 \cdot t_i} - 3,0944 \cdot e^{-0,0067 \cdot t_i}]}}$	0,9853	5,3493	2,6278	1,7338	< 0,05	0,07	<0,05
MP6	$\hat{V}_f = V_i - (-383,8 \cdot PP + 20830 \cdot Tmax) \cdot (e^{-0,00001 t_f^{0,5752}} - e^{-0,00001 t_i^{0,5752}})$	0,9830	5,6704	2,7948	1,8870	< 0,05	0,129	<0,05

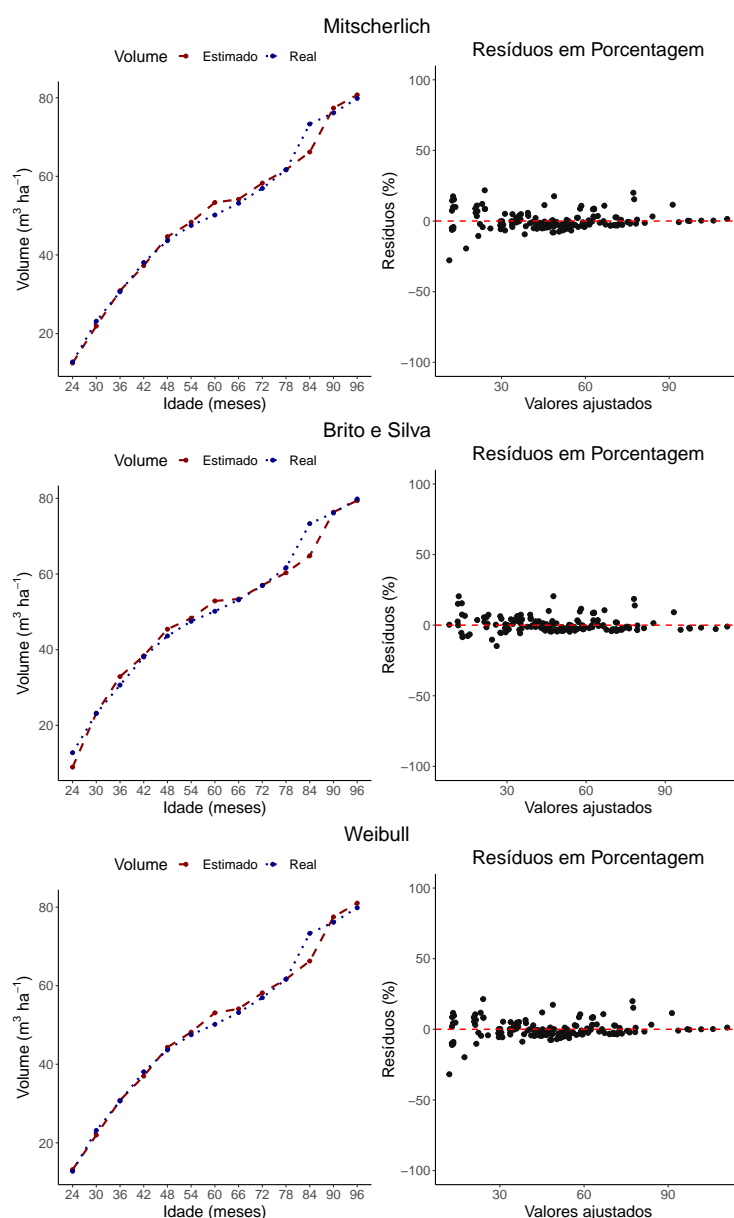
REQM e EMA em  $\text{m}^3 \text{ha}^{-1}$ , SW = Teste de Shapiro-Wilk, BW = Teste de Breush-Pagan, DW = Teste de Durbin-Watson (para todos os testes dos resíduos são apresentados os p-Valores), PP = Precipitação (mm), Tmax = temperatura máxima média ( $^{\circ}\text{C}$ ).

Com exceção da equação obtida pelo modelo de Brito-Silva a adição de variáveis climáticas proporcionou ganho na qualidade das estatísticas de avaliação, aumentado os valores de IA<sub>c</sub>, REQM e EAM e diminuindo o erro padrão (S<sub>xy</sub>%). Para as equações provenientes de Brito-Silva e Weibull a inclusão das variáveis do clima corrigiu o problema da falta homogeneidade na variância dos resíduos e para a equação obtida de Mitscherlich resolveu o problema da autocorrelação dos resíduos. Sendo assim, uma transformação logarítmica na variável resposta pode ajudar a estabilizar a variância do modelo e solucionar os problemas com os demais pressupostos. Já que o problema da falta de normalidade pode ser tolerável em certos níveis (PINO, 2014), a autocorrelação pode ser tratada empregando o método dos mínimos quadrados generalizados, esse é empregado para estimar um parâmetro desconhecido na presença de erros heterocedásticos ou quando as observações apresentam correlações (SILVA, 2016).

Os gráficos dos resíduos e as curvas de crescimento são apresentadas na Figura 32.

A inclusão de variáveis climáticas em modelos de prognose também foi avaliada por Ferraz-Filho et al. (2011), esses autores desenvolveram equações para prognose da altura dominante de *Eucalyptus* aos 5,5 anos de idade e afirmaram que a inclusão das variáveis climáticas permitiram projeções mais precisas que as equações tradicionais. Quando se adiciona as variáveis do clima em modelos biológicos se permite que a informação de algumas mudanças do ambiente sejam também modeladas, aumentando assim a precisão das projeções.

Figura 32 – Curvas de crescimento real e estimada e distribuição dos resíduos para as equações de prognose do volume de clones de *Eucalyptus* com adição de variáveis climáticas.



### 8.3.2 Prognose do volume por meio de algoritmos de aprendizagem de máquinas

Para prognose do volume no povoamento de clones de *Eucalyptus* spp. foram treinadas 300 RNA's. A melhor topografia da rede foi selecionada com os seguintes hiperparâmetros: neurônios (3:6:1), como função de ativação foram selecionadas a *Maxout* e a linear para camada oculta e de saída; foram necessárias 750 épocas para o treinamento; foram selecionadas as taxas de regularização L1 e L2 de  $10^{-6}$ ; taxa de aprendizagem de 0,006 e *momentum* de 0,1. Para o algoritmo de MVS as melhores configurações foram: função de *kernel radial*, parâmetro custo de 22,6274, *gamma* de 0,1 e *epsilon* de 0,08. Após o treino de 24 algoritmos de florestas aleatórias o que apresentou os melhores ajustes possuía os hiperparâmetros: número de árvores criadas igual a 1000, número de folhas médias geradas de 58,85 e profundidade média de regra de decisão de 9,11. Os resultados das análises estatísticas para avaliação da qualidade de ajuste são apresentadas na Tabela 30.

Tabela 30 – Resumo estatístico dos algoritmos de aprendizagem de máquinas para prognose do volume de povoamentos de clones de *Eucalyptus* spp.

Algoritmos de aprendizagem de máquinas	Validação				Teste	
	IA <sub>c</sub>	S <sub>xy</sub> %	EAM	REQM	EAM	REQM
Redes Neurais Artificiais	0,9897	4,5424	1,3648	1,9859	3,5900	5,3698
Floresta Aleatória	-	-	1,8749	2,9369	2,5330	3,6990
Máquinas de Vetores de Suporte	0,9891	4,6692	1,3989	2,0050	3,6384	5,4381

REQM e EMA em m<sup>3</sup> ha<sup>-1</sup>

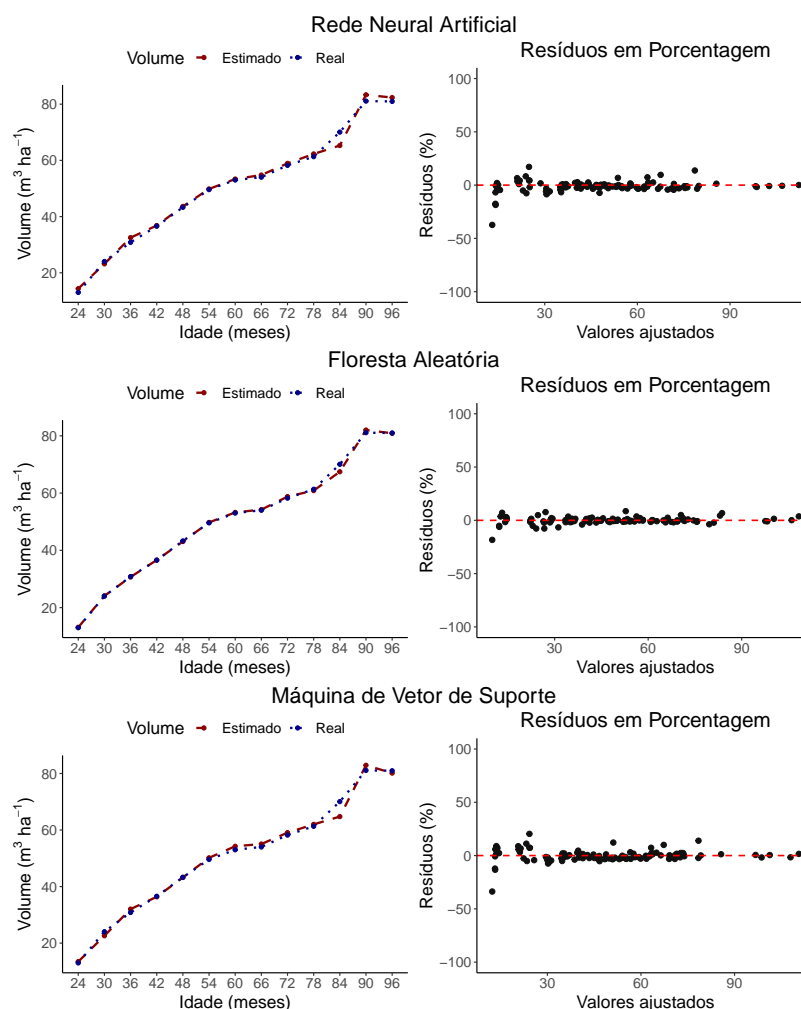
O algoritmo de aprendizagem de máquina que se mostrou com maior poder de generalização para o conjunto de teste foi a Floresta Aleatória (FA). Sua avaliação foi realizada com base no EAM e na REQM, pois esse algoritmo não é baseado no ajuste de um conjunto de parâmetros como acontece com as RNA's e MVS's, o mesmo é baseado em um conjunto de regras de decisões o que torna impossível o cálculo do IA<sub>c</sub> e S<sub>xy</sub>%. Esse método tem sido bastante empregado em problemas florestais ligados a área do sensoriamento remoto que geralmente envolve uma grande quantidade de variáveis preditoras, nessa área as FA's tem ganhado destaque em comparação com os demais métodos de aprendizagem de máquinas (LATIFI; NOTHDURFT; KOCH, 2010;

GÖRGENS; MONTAGHI; RODRIGUEZ, 2015; REIS et al., 2019).

Reis et al. (2018) trabalhando com a estimativa da área basal por meio de índices de vegetação obtidos com a metodologia LIDAR (Light Detection And Ranging) também chegaram a resultados superiores com as FA's em detrimento de outras técnicas como RNA, MSV e análise de regressão. Nesta pesquisa, para prognose do volume os modelos biológicos clássicos apresentaram resultados superiores quando comparados aos algoritmos de RNA e MVS e estatísticas de qualidade semelhantes com a FA. Entretanto, esses métodos não possuem pressupostos sobre os dados o que permite trabalhar com uma gama maior de variáveis explanatórias sem perda de informação do fenômeno (GÖRGENS; MONTAGHI; RODRIGUEZ, 2015).

O gráfico com a análise dos resíduos é apresentado na Figura 33. Como se pode ver a RNA e a MVS possuem uma tendência maior em subestimar o volume na fase inicial do crescimento.

Figura 33 – Análise dos resíduos para os algoritmos de aprendizagem de máquinas treinados para prognose do volume de clones de *Eucalyptus*.





Após adição das variáveis climáticas foram treinadas 300 RNA's. O melhor desempenho foi obtido com o algoritmo de RNA com os seguintes hiperparâmetros: neurônios (6:8:1), com as funções de ativação *Maxout* e linear na camada oculta e de saída, respectivamente; 500 épocas ou ciclos de treinamentos foram utilizados; as taxas de regularização foram L1 e L2 igual 0,001; a taxa de aprendizagem foi 0,14 e o *momentum* de 0,1. Quando foi adicionado as variáveis climáticas no algoritmo de MVS as melhores configurações foram: função de *kernel radial*, parâmetro custo de 5,6568, *gamma* de 0,1 e *epsilon* de 0,08. Foram treinados 38 algoritmos de florestas aleatórias sendo selecionado o que apresentou os seguintes hiperparâmetros: número de árvores = 1000, número de folhas médio de 58,84 e profundidade média de 9,06. As análises estatísticas empregadas na comparação dos algoritmos são mostradas na Tabela 31.

Tabela 31 – Resumo estatístico dos algoritmos de aprendizagem de máquinas com adição de variáveis climáticas para prognose do volume de povoamentos de clones de *Eucalyptus* spp.

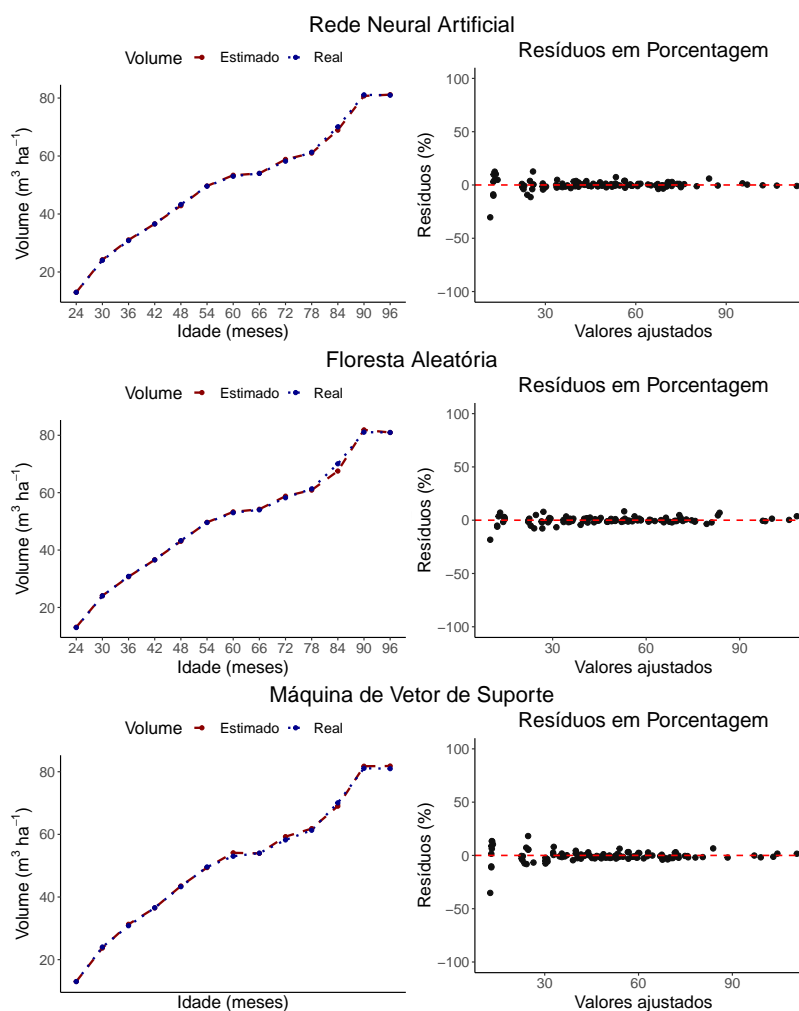
Algoritmos de aprendizagem de máquinas	Validação				Teste	
	IA <sub>c</sub>	S <sub>xy</sub> %	EAM	REQM	EAM	REQM
Redes Neurais Artificiais	0,9939	3,4884	0,8649	1,2268	3,5682	4,6677
Floresta Aleatória	-	-	1,8988	2,9184	2,5590	3,3850
Máquinas de Vetores de Suporte	0,9936	3,5771	1,2007	1,5117	3,6384	5,4381

REQM e EMA em m<sup>3</sup> ha<sup>-1</sup>

A adição de variáveis climáticas não proporcionou melhoras significativas na prognose do volume. Em relação aos algoritmos RNA's, MVS's e FA's o que pode acontecer é que algumas das variáveis não adicionam ganhos de informação significativos as predições e conseqüentemente os algoritmos dão menos preferências para essa variáveis, ou seja, eles dão preferência a variável mais importante. Como para a prognose do volume final o volume inicial entra como variável preditora e é altamente correlacionada com a variável que se pretende estimar os algoritmos tendem focar o aprendizado na variável com esta elevada correlação.

O mesmo comportamento de subestimava na fase inicial do crescimento é observado com os algoritmos de aprendizagem de máquinas com adição das variáveis climáticas (Figura 34).

Figura 34 – Análise dos resíduos para os algoritmos de aprendizagem de máquinas com adição de variáveis climáticas treinados prognose do volume de clones de *Eucalyptus*.



## 8.4 Conclusão

As melhores equações para prognose do volume de clones de *Eucalyptus* foram obtidas dos modelos de Mitscherlich, Brito-Silva e Clutter-Jones. A adição de variáveis climáticas no modelos biológicos proporcionou ganhos na qualidade das estimativas. Os algoritmos de aprendizagem de máquinas RNA, MSV e FA podem ser empregados com eficiência na prognose do volume. A FA apresentou estatísticas de qualidade semelhantes aos modelos clássicos. Quando foi adicionado as variáveis climáticas na RNA, MVS e FA os resultados não foram superiores.

## Referências Bibliográficas

- ABREU, J. C. d. **Alternativas para estimar o volume de árvores individuais em formações florestais no estado de Minas Gerais**. Tese (Ciência Florestal) — Universidade Federal de Viçosa, Viçosa, MG, 2019.
- BINOTI, D. H. B. et al. Support vector machine to estimate volume of eucalypt trees. **Revista Árvore**, v. 40, n. 4, p. 689–693, 2016.
- BINOTI, M. L. M. d. S. et al. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, n. 1, p. 97–105, 2015.
- BRITO, C. C. R. d. et al. Modelos de crescimento resultantes da combinação e variações dos modelos de chapman-richards e silva-bailey aplicados em leucaena leucocephala (lam.) de wit. **Ciência Florestal**, v. 17, n. 2, p. 175–185, 2007.
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 5. ed. Viçosa, MG: Editora UFV, 2017. 636 p.
- CASTRO, R. V. O. et al. Weibull function adjustment procedures for prognosing diameter distribution of Eucalyptus plantations. **Scientia Agraria Paranaensis**, v. 16, n. 4, p. 528–534, 2017.
- DIAMANTOPOULOU, M. J. et al. Estimation of weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. **Biosystems Engineering**, v. 133, p. 33–45, 2015.
- ELLI, E. F. et al. Climatic factors defining the height growth curve of forest species. **iForest-Biogeosciences and Forestry**, v. 10, n. 3, p. 547, 2017.
- FERRAZ-FILHO, A. C. et al. Dominant height projection model with the addition of environmental variables. **Cerne**, v. 17, n. 3, p. 427–433, 2011.
- GÖRGENS, E. B.; MONTAGHI, A.; RODRIGUEZ, L. C. E. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. **Computers and Electronics in Agriculture**, v. 116, p. 221–227, 2015.
- GUERA, O. G. M. et al. Alternatives to growth and yield prognosis for *Pinus caribaea* var. *caribaea* barrett & golfari. **Floresta e Ambiente**, v. 26, n. 4, 2019b.

- LATIFI, H.; NOTHDURFT, A.; KOCH, B. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/lidar-derived predictors. **Forestry**, v. 83, n. 4, p. 395–407, 2010.
- MAESTRI, R. **Modelo de Crescimento e Produção para Povoamentos Clonais de *Eucalyptus grandis* Considerando Variáveis Ambientais**. 158 p. Tese (Doutor em Ciências Florestais) — Universidade Federal do Paraná, Curitiba, 2003.
- MARTINS, F. B.; SOARES, C. P. B.; SILVA, G. F. d. Individual tree growth models for eucalyptus in northern brazil. **Scientia Agricola**, v. 71, n. 3, p. 212–225, 2014.
- ÖZÇELİK, R. et al. Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. **Forest Science**, v. 63, n. 3, p. 291–302, 2017.
- PINO, F. A. A questão da não normalidade: Uma revisão. **Revista de economia agrícola**, v. 61, n. 2, p. 17–33, 2014.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Disponível em: <<https://www.R-project.org/>>.
- REIS, A. A. et al. Spatial prediction of basal area and volume in *Eucalyptus* stands using landsat tm data: an assessment of prediction methods. **New Zealand Journal of Forestry Science**, v. 48, n. 1, p. 1, 2018.
- REIS, A. A. et al. Volume estimation in a *Eucalyptus* plantation using multi-source remote sensing and digital terrain data: a case study in minas gerais state, brazil. **International journal of remote sensing**, Taylor & Francis, v. 40, n. 7, p. 2683–2702, 2019.
- SILVA, J. A. A. **Dynamics of stand structure in fertilized slash pine plantations**. 143 p. Tese (Biometrics and Forest Management) — University of Georgia, Athens, Georgia, 1986.
- SILVA, J. A. A. Conceitos e princípios básicos de modelagem matemática em ciências florestais. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 11/12, p. 195–215, 2014/2015.
- SILVA, J. A. A. et al. Modelagem do crescimento volumétrico de clones de eucalipto (*Eucalyptus urophylla*) no Polo Gesseiro do Araripe-PE. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 13/14, p. 173–190, 2016/2017.

SILVA, J. W. L. **Modelagem da biomassa e da quantidade de carbono de clones de *Eucalyptus* da Chapada do Araripe-PE**. 105 p. Dissertação (Biometria e Estatística Aplicada) — Universidade Federal Rural de Pernambuco, Recife, PE, 2016.

SOUZA, S. R. R. et al. Redes neurais para estimativa volumétrica de clones de *Eucalyptus* spp. no Polo Gesseiro do Araripe. **Revista Brasileira de Biometria**, v. 36, n. 3, p. 715–729, 2018.

## CAPÍTULO 9

---

### Conclusão Geral

Os clones de *Eucalyptus* spp. possuem um grande potencial para serem empregados como fonte de recursos energéticos para as indústrias do Polo Gesseiro do Araripe-PE. Estes, possuem taxa de sobrevivência acima de 94% na região. Sua idade de rotação técnica é de aproximadamente 50 meses para os povoamentos mais adensados e para os arranjos 3 m x 3 m e 3 m x 2 m fica em torno de 60 meses. Desta forma, o povoamento de clones de *Eucalyptus* possui idade de rotação técnica inferior à de planos de manejo florestal (PMF) da Caatinga. Da mesma forma, o incremento médio anual (IMA) de PMF's da Caatinga é inferior ao IMA registrado nessa pesquisa.

Em relação as mudanças na condições climáticas da região a análise da série de precipitação anual revelou que existe uma tendência de queda nos níveis de chuva da ordem de 7,81 mm por ano. Enquanto que a série de temperatura média do ar possui uma tendência positiva. Pela análise descritiva ficou evidente que existe diferenças nos níveis das variáveis climáticas entre os dois cenários climáticos estudados. A produtividade dos clones de *Eucalyptus* foi maior para o cenário com melhor condição de clima.

Considerando o emprego de modelos mais robustos quanto a falta de aderência aos pressupostos da regressão, foi possível obter equações para estimar a produção volumétrica com boa acurácia por meio da metodologia de modelos com resíduos simétricos e não lineares generalizados. No entanto, apenas as equações com erros simétricos obtidas com a transformação logarítmica dos modelos de Schumacher e Hall e Spurr atenderam a todos os requisitos estatísticos. Já com os modelos não lineares generalizados foi possível gerar equações com baixo erro padrão de estimativa e permitir a atribuição de outras distribuições de probabilidade aos modelos clássicos da área florestal.

As técnicas de Inteligência Artificial, dentre elas as: redes neurais artificiais (RNA), máquinas de vetores de suporte (MVS) e florestas aleatórias (FA) se mostram uma boa alternativa para estimar a produção de povoamentos florestais. E ainda, são robustas a falta de normalidade e heterocedasticidade. As redes neurais artificiais e as máquinas de vetores de suporte se mostraram superiores ao algoritmo de florestas

aleatórias.

A inclusão de variáveis climáticas nos modelos de crescimento e prognose do volume em povoamentos de clones de *Eucalyptus* spp. plantados na região semiárida pernambucana permitiu a obtenção de equações com maior acurácia e sensíveis a variações do clima local. Bem como, os algoritmos de inteligência artificial RNA, MVS e FA possuem resultados semelhantes aos modelos biológicos de crescimento e prognose. Os algoritmos inteligentes são robustos a falta de normalidade, heterogeneidade e autocorrelação dos resíduos se mostrando apropriados para problemas florestais onde estas situações são comuns.

## Referências Bibliográficas

- ABREU, J. C. d. **Alternativas para estimar o volume de árvores individuais em formações florestais no estado de Minas Gerais**. Tese (Ciência Florestal) — Universidade Federal de Viçosa, Viçosa, MG, 2019.
- ABREU, Y. K. L. **Emprego de funções de densidade de probabilidade na modelagem da distribuição diamétrica de clones de *Eucalyptus* spp. no Polo Gesseiro do Araripe**. 113 p. Dissertação (Biometria e Estatística Aplicada) — Universidade Federal Rural de Pernambuco, Departamento de Estatística e Informática, Recife, 2018.
- ADAM, E. et al. Estimating standing biomass in papyrus (*Cyperus papyrus* L.) swamp: exploratory of in situ hyperspectral indices and random forest regression. **International Journal of Remote Sensing**, v. 35, n. 2, p. 693–714, 2014.
- ALVES, A. M. C. et al. Quantificação da produção de biomassa em clones de eucaliptos com 4, 5 anos, no Polo Gesseiro do Araripe-Pe. **Revista de Ciências Agrárias Amazonian Journal of Agricultural and Environmental Sciences**, v. 48, n. 1, p. 161–174, 2007.
- BACK, Á. J. Aplicação de análise estatística para identificação de tendências climáticas. **Pesquisa agropecuária brasileira**, v. 36, n. 5, p. 717–726, 2001.
- BAGUE, I. P. M. **Modelización del crecimiento diametral del encinar (*Quercus ilex* subsp. *ilex*) en el Parque Natural del Montnegre y el Corredor**. 68 p. Dissertação (Máster en Ingeniería de Montes) — Universidad de Lleida—Escuela Técnica Superior de Ingeniería Agraria, 2014.
- BAILEY, R. L.; CLUTTER, J. L. Base-age invariant polymorphic site curves. **Forest Science**, v. 20, n. 2, p. 155–159, 1974.
- BARRAZA, G. Q. N. **Sistema de crecimiento y rendimiento para las principales especies comerciales de *Pinus* en Durango, México**. 194 p. Tese (Doctor en Ciencias) — Institución de Enseñanza e Investigación en Ciencias Agrícolas, Montecillo, Texcoco, Edo. de Mexico, 2014.
- BARRIO-ANTA, M. et al. Predicción del índice de sitio a partir de variables ambientales y de concentración foliar de nutrientes en masas regulares de *Pinus pinaster* ait. en asturias. **Cuadernos de la Sociedad Española de Ciencias Forestales**, n. 34, p. 39–46, 2012.
- BARROS, B. C. et al. Volumetria e sobrevivência de espécies nativas e exóticas no Polo Gesseiro do Araripe, PE. **Ciência Florestal**, v. 20, n. 4, p. 641–647, 2010.
- BATES, D. M.; WATTS, D. G. **Nonlinear regression analysis and its applications**. New York, NY: Wiley Online Library, 1988. 365 p.



BERNIER, J.; FENG, Y.; ASAKAWA, K. Strategies for handling normality assumptions in multi-level modeling: a case study estimating trajectories of health utilities index mark 3 scores. **Health Rep**, v. 22, p. 45–51, 2011.

BHERING, L. L. et al. Application of neural networks to predict volume in eucalyptus. **Crop Breeding and Applied Biotechnology**, v. 15, n. 3, p. 125–131, 2015.

BINOTI, D. H. B. et al. Support vector machine to estimate volume of eucalypt trees. **Revista Árvore**, v. 40, n. 4, p. 689–693, 2016.

BINOTI, M. L. M. d. S. et al. Redes neurais artificiais para estimação do volume de árvores. **Revista Árvore**, v. 38, n. 2, p. 283–288, 2014.

BINOTI, M. L. M. d. S. et al. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, n. 1, p. 97–105, 2015.

BLAIN, G. C. Detecção de tendências monótonas em séries mensais de precipitação pluvial do estado de são paulo. **Bragantia**, v. 69, n. 4, p. 1027–1033, 2010.

BONAN, G. B.; DONEY, S. C. Climate, ecosystems, and planetary futures: The challenge to predict life in earth system models. **Science**, v. 359, p. 533–543, 2018.

BONETE, I. P. et al. Artificial neural network and mixed-effects modeling to describe the stem profile of *Pinus taeda* L. v. 50, n. 1, p. 1123–1132, 2020.

BRASIL, M. M. E. **Pérfil da Gipsita**. 34 p. (Relatório Técnico n. 34), Brasília, 2009.

BREIMAN, L. Random forests. **Machine learning**, v. 45, n. 1, p. 5–32, 2001.

BREUSCH, T. S.; PAGAN, A. R. A simple test for heteroscedasticity and random coefficient variation. **Econometrica: Journal of the Econometric Society**, p. 1287–1294, 1979.

BRITO, C. C. R. d. et al. Modelos de crescimento resultantes da combinação e variações dos modelos de chapman-richards e silva-bailey aplicados em leucaena leucocephala (lam.) de wit. **Ciência Florestal**, v. 17, n. 2, p. 175–185, 2007.

BURKHART, H. E.; TOMÉ, M. **Modeling Forest Trees and Stands**. New York: Springer, 2012. 460 p.

CALEGARIO, N. et al. Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não lineares generalizados. **Scientia Forestalis, Piracicaba**, v. 69, p. 38–50, 2005.

CAMPOS, A. R. d. et al. Tratamento de minérios. In: CETEM/MCTIC (Ed.). 6. ed. Rio de Janeiro: CETEM/MCTIC, 2018. cap. Calcinação, p. 725–752.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 5. ed. Viçosa, MG: Editora UFV, 2017. 636 p.

- CARON, B. O. et al. Quantificação da biomassa florestal em plantios de curta rotação com diferentes espaçamentos/quantification of forest biomass in short rotation plantations with different spacings. **Comunicata Scientiae**, v. 6, n. 1, p. 106–112, 2015.
- CARVALHO, A. C. et al. Produção de energia da madeira de espécies da Caatinga aliada ao manejo florestal sustentável. v. 48, n. 126, p. 1–13, 2020.
- CASTRO, C. A. d. O. et al. Brief history of *Eucalyptus* breeding in brazil under perspective of biometric advances. **Ciência Rural**, v. 46, n. 9, p. 1585–1593, 2016.
- CASTRO, R. V. O. et al. Weibull function adjustment procedures for prognosing diameter distribution of Eucalyptus plantations. **Scientia Agraria Paranaensis**, v. 16, n. 4, p. 528–534, 2017.
- CASTRO, R. V. O. et al. Projeção do diâmetro e altura de árvores em uma floresta estacional semidecidual por meio de redes neurais artificiais. v. 48, n. 125, p. 1–12, 2020.
- CEPAGRI/UNICAMP; EMBRAPA, I. A. **Agritempo**. 2019. Disponível em: <[www.agritempo.gov.br](http://www.agritempo.gov.br)>.
- CHEN, J.; GUPTA, A. K. **Parametric Statistical Change Point Analysis**. Kansas City, MO: Springer Science Business Media LLC, 2000. 190 p.
- CHEN, J.; GUPTA, A. K. **Parametric statistical change point analysis: with applications to genetics, medicine, and finance**. 2. ed. New York: Springer Science & Business Media, 2012. 281 p.
- CHIARELLO, F. et al. Artificial neural networks applied in forest biometrics and modeling: State of art (january/2007 to july/2018). **Cerne**, v. 25, n. 2, p. 140–155, 2019.
- CIESZEWSKI, C.; BAILEY, R. Generalized algebraic difference approach: Theory based derivation of dynamic site equations with polymorphism and variable asymptotes. **Forest Science**, v. 46, n. 1, p. 116–126, 2000.
- CIESZEWSKI, C. J. Three methods of deriving advanced dynamic site equations demonstrated on inland douglas-fir site curves. **Canadian Journal of Forest Research**, v. 31, n. 1, p. 165–173, 2001.
- CORDEIRO, M. A. et al. Estimativa do volume de acacia mangium utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. **Pesquisa Florestal Brasileira**, v. 35, n. 83, p. 255–261, 2015.
- COSTA, V. J.; FERREIRA, M.; CORDEIRO, M. T. A. Análise de séries temporais climáticas. **Revista de Ciências Agroveterinárias**, v. 14, n. 2, p. 169–177, 2015.
- CURRIE, I. D. On fitting generalized linear and non-linear models of mortality. **Scandinavian Actuarial Journal**, v. 2016, n. 4, p. 356–383, 2016.
- CYSNEIROS, F. J. A.; PAULA, G. A.; GALEA, M. **Modelos simétricos aplicados**. São Paulo: Associação Brasileira de Estatística, 2005. v. 1. 88 p.

DIAMANTOPOULOU, M. J. et al. Estimation of weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. **Biosystems Engineering**, v. 133, p. 33–45, 2015.

DRUCKER, H. et al. Support vector regression machines. In: JORDAN, M. I.; KEARNS, M. J.; SOLLA, S. A. (Ed.). **Advances in neural information processing systems**. Denver, Colorado, USA: MIT Press, 1997. p. 155–161.

DRUMOND, M.; LIMA, P.; SANTOS, R. A. V. Comportamento de algumas espécies/procedências de *Eucalyptus* no município de Lagoa Grande-PE. **Brasil Florestal**, n. 78, p. 75–80, 2003.

DRUMOND, M. A. et al. Performance of two hybrid clones of *Eucalyptus* planted under five spacings in the Araripe plateau, Pernambuco, Brazil. **Acta Horticulturae**, n. 959, p. 167–171, 2012.

DUMAS, M. D. Changes in temperature and temperature gradients in the french northern alps during the last century. **Theoretical and applied climatology**, v. 111, n. 1-2, p. 223–233, 2013.

ELLI, E. F. et al. Climatic factors defining the height growth curve of forest species. **iForest-Biogeosciences and Forestry**, v. 10, n. 3, p. 547, 2017.

FABRIKA, M.; VALENT, P.; MERGANIČOVÁ, K. Forest modelling and visualisation—state of the art and perspectives. **Central European Forestry Journal**, v. 66, n. 3-4, p. 147–165, 2019.

FERRAZ-FILHO, A. C. et al. Dominant height projection model with the addition of environmental variables. **Cerne**, v. 17, n. 3, p. 427–433, 2011.

FERREIRA, A. D. et al. Influence of spatial arrangements on silvicultural characteristics of three *Eucalyptus* clones at integrated crop-livestock-forest system. **African Journal of Agricultural Research**, v. 11, n. 19, p. 1734–1742, 2016.

FERREIRA, M. Z. **Modelagem da Influencia de Variáveis Ambientais no Crescimento e na Produção de *Eucalyptus* sp.** 112 p. Tese (Doutorado em Ciências Florestais) — Universidade Federal de Lavras, Lavras, 2009.

FILHO, J. I. P. et al. Avaliação econômica de projetos de florestamento com eucalyptus grandis hill ex maiden implantado sob diferentes espaçamentos. v. 28, n. 1, p. 240–248, 2018.

FILHO, L. M. A. L. et al. Modelagem do crescimento de clones de eucalyptus usando o modelo de chapman-richards com diferentes distribuições simétricas dos erros. **Ciência Florestal**, v. 22, n. 4, p. 777–785, 2012.

FONTENELE, N. d. M. et al. Volumetric and economic evaluation of *Eucalyptus* spp. clones in high forest, coppice and native vegetation systems at the Gypsum Pole of Araripe-PE. **Scientia Forestalis**, v. 46, n. 117, p. 77–86, 2018.

- GADELHA, F. H. L. **Desempenho silvicultural e avaliação econômica de clones híbridos de eucaliptos plantados em diferentes regimes de manejo para fins energéticos**. 147 p. Tese (Ciências Florestais) — Universidade Federal Rural de Pernambuco, Departamento de Ciências Florestais, Recife, 2014.
- GADELHA, F. H. L.; SILVA, J. A. A.; FERREIRA, R. L. C. Viabilidade econômica do cultivo de clones de *Eucalyptus* em função do espaçamento e do sistema silvicultural, em Araripina-PE. **Anais da Academia Pernambucana de Ciência Agrônoma**, v. 15, n. 2, p. 105–123, 2018.
- GADELHA, F. H. L. et al. Produtividade de clones de eucaliptos em diferentes sistemas de manejo para fins energéticos. **Pesquisa Florestal Brasileira**, v. 35, n. 83, p. 263–270, 2015.
- GARCÍA-GUTIÉRREZ, J. et al. A comparison of machine learning regression techniques for lidar-derived estimation of forest variables. **Neurocomputing**, v. 167, p. 24–31, 2015.
- GEDEON, T. D. Data mining of inputs: analysing magnitude and functional measures. **International Journal of Neural Systems**, v. 8, n. 02, p. 209–218, 1997.
- GENUER, R. Variance reduction in purely random forests. **Journal of Nonparametric Statistics**, v. 24, n. 3, p. 543–562, 2012.
- GENUER, R.; POGGI, J.-M.; TULEAU-MALOT, C. Vsurf: An r package for variable selection using random forests. **The R Journal**, v. 7, n. 2, p. 19–33, 2015.
- GONÇALVES, F. N.; BACK, Á. J. Análise da variação espacial e sazonal e de tendências na precipitação da região sul do Brasil. **Revista de Ciências Agrárias**, v. 41, n. 3, p. 11–20, 2018.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. London, England: MIT press, 2016. 800 p.
- GÖRGENS, E. B.; MONTAGHI, A.; RODRIGUEZ, L. C. E. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. **Computers and Electronics in Agriculture**, v. 116, p. 221–227, 2015.
- GUERA, O. G. M. et al. Modelos matemáticos para la determinación del turno óptimo en plantaciones forestales. **Madera y Bosques**, v. 25, n. 1, p. 1–19, 2019a.
- GUERA, O. G. M. et al. Alternatives to growth and yield prognosis for *Pinus caribaea* var. *caribaea* barrett & golfari. **Floresta e Ambiente**, v. 26, n. 4, 2019b.
- GUERRA, G. d. S. **Avaliação do desempenho e adequação de diferentes metodologias aplicadas para modelagem e inferência de emissão de gases do efeito estufa**. 76 p. Dissertação (Biometria e Estatística Aplicada) — Universidade Federal Rural de Pernambuco, Recife, PE, 2016.

- HALL, K. B. et al. A growth and yield model for *Eucalyptus benthamii* in the southeastern United States. **Forest Science**, v. 66, n. 1, p. 1–13, 2019.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning**. 2. ed. New York, NY: Springer series in statistics New York, 2009. (Springer Series in Statistics).
- HAYKIN, S. **Neural Network A comprehensive foundation**. [S.l.]: Tsinghua University Press, 2004. v. 2. 842 p.
- HESS, A. F. et al. Aplicação dos modelos lineares generalizados para estimativa do crescimento em altura. **Pesquisa Florestal Brasileira**, v. 35, n. 84, p. 427–433, 2015.
- HSU, J. **Multiple comparisons: theory and methods**. [S.l.]: Chapman and Hall/CRC, 1996. 234 p.
- HUTCHINS, D. A. et al. Climate change microbiology—problems and perspectives. **Nature Reviews Microbiology**, v. 17, n. 6, p. 391–396, 2019.
- IBÁ. **Indústria brasbrasileira de árvores**. São Paulo, 2019. 80 p.
- IBGE - Instituto Brasileiro de Geografia e Estatística. Produção da extração vegetal e da silvicultura 2017. **Prod. Extr. veg. e Silvíc**, v. 32, p. 1–8, 2017.
- IBGE - Intituto Brasileiro de Geografia e Estatística. **Produto Interno Bruto dos Municípios**. 2019. Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9088-produto-interno-bruto-dos-municipios.html>>.
- JAMES, G. et al. **Machine Learning**. New York: Springer, 2015. 440 p.
- JIMÉNEZ-RUANO, A.; MIMBRERO, M. R.; FERNÁNDEZ, J. de la R. Exploring spatial–temporal dynamics of fire regime features in mainland Spain. **Natural Hazards & Earth System Sciences**, v. 17, n. 10, p. 1697–1711, 2017.
- JIN, W.; HE, H. S.; THOMPSON, F. R. Are more complex physiological models of forest ecosystems better choices for plot and regional predictions? **Environmental Modelling & Software**, v. 75, p. 1–14, 2016.
- KEENAN, R. J. Climate change impacts and adaptation in forest management: a review. **Annals of Forest Science**, v. 72, n. 2, p. 145–167, 2015.
- KELLOMÄKI, S. et al. Sensitivity of managed boreal forests in Finland to climate change, with implications for adaptive management. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 363, n. 1501, p. 2339–2349, 2008.
- KILLICK, R.; ECKLEY, I. A. changepoint: An R package for changepoint analysis. **Journal of Statistical Software**, v. 58, n. 3, p. 1–19, 2014.
- KIRSCHBAUM, M.; FISCHIN, A. Climate change impacts on forests. In: WATSON, R.; ZINYOWERA, M. C.; MOSS, R. H. (Ed.). **Climate change 1995: impacts, adaptation and mitigation of climate change: scientific-technical analysis**. England: Cambridge: Cambridge University Press, 1996. p. 95–129.

- KRUG, T. Impacto, vulnerabilidade e adaptação das florestas à mudança do clima. **Parcerias estratégicas**, v. 13, n. 27, p. 43–72, 2010.
- LACERDA, F. F.; LOPES, G. M. B. Extremos de chuva e seca em Pernambuco: Uma realidade que veio para ficar. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 13/14, p. 42–46, 2016/2017.
- LACERDA, F. F. et al. Alterações climáticas globais; uma realidade em pernambuco. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 11/12, p. 121–154, 2014/2015.
- LATIFI, H.; NOTHDURFT, A.; KOCH, B. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/lidar-derived predictors. **Forestry**, v. 83, n. 4, p. 395–407, 2010.
- LEDELL, E. et al. **h2o: R Interface for 'H2O'**. [S.l.], 2019. R package version 3.24.0.5. Disponível em: <<https://CRAN.R-project.org/package=h2o>>.
- LEITE, H. G.; OLIVEIRA, F. H. Tavares de. Statistical procedure to test identity between analytical methods. **Communications in soil science and plant analysis**, v. 33, n. 7-8, p. 1105–1118, 2002.
- LINS, C. A. C. **Informe Geoquímico da Bacia do Araripe: estados de Pernambuco, Piauí e Ceará**. Recife, 2018. 31 p.
- LOBELL, D. B.; SCHLENKER, W.; COSTA-ROBERTS, J. Climate trends and global crop production since 1980. **Science**, v. 333, n. 6042, p. 616–620, 2011.
- LORENA, A. C.; CARVALHO, A. C. de. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 2007.
- LOTZE, H. K. et al. Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. **Proceedings of the National Academy of Sciences**, v. 116, n. 26, p. 12907–12912, 2019.
- MACHADO, S. d. A. et al. Efeito de variáveis climáticas no crescimento mensal de *Pinus taeda* e *Araucaria angustifolia* em fase juvenil. **Floresta e ambiente**, v. 21, n. 2, p. 170–181, 2014.
- MAESTRI, R. **Modelo de Crescimento e Produção para Povoamentos Clonais de *Eucalyptus grandis* Considerando Variáveis Ambientais**. 158 p. Tese (Doutor em Ciências Florestais) — Universidade Federal do Paraná, Curitiba, 2003.
- MAESTRI, R. et al. Modelagem do crescimento florestal considerando variáveis do ambiente: Revisão. **Scientia Agraria**, v. 14, n. 3, p. 103–110, 2013.
- MAGALHÃES, W. M. et al. Desempenho silvicultural de clones e espécies/procedências de *Eucalyptus* na região noroeste de Minas Gerais. **Cerne**, v. 13, n. 4, 2007.
- MARENGO, J. A. Vulnerabilidade, impactos e adaptação à mudança do clima no semi-árido do Brasil. **Parcerias estratégicas**, v. 13, n. 27, p. 149–176, 2010.

- MAROHASY, J.; ABBOT, J. Assessing the quality of eight different maximum temperature time series as inputs when using artificial neural networks to forecast monthly rainfall at cape otway, australia. **Atmospheric Research**, v. 166, p. 141–149, 2015.
- MARTINS, F. B.; SOARES, C. P. B.; SILVA, G. F. d. Individual tree growth models for eucalyptus in northern brazil. **Scientia Agricola**, v. 71, n. 3, p. 212–225, 2014.
- MENDIBURU, F. d. **agricolae: Statistical Procedures for Agricultural Research**. [S.I.], 2017. R package version 1.2-8. Disponível em: <<https://CRAN.R-project.org/package=agricolae>>.
- MENDONÇA, A. R. d. et al. Growth and yield models for eucalyptus stands obtained by differential equations. **Scientia Agricola**, v. 74, n. 5, p. 364–370, 2017.
- MEUNIER, I. M. J.; FERREIRA, R. L. C.; SILVA, J. A. A. O licenciamento de Planos de Manejo Florestal da Caatinga assegura sua sustentabilidade? **Pesquisa Florestal Brasileira**, v. 38, p. 1–7, 2018.
- MEYER, D. et al. **e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien**. [S.I.], 2017. R package version 1.6-8. Disponível em: <<https://CRAN.R-project.org/package=e1071>>.
- MIGUEL, E. P. et al. 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.
- MONTAÑO, R. A. N. R. **Aplicação de técnicas de aprendizado de máquina na mensuração florestal**. 102 p. Tese (Informática) — Universidade Federal do Paraná, Curitiba, PR, 2016.
- MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**. 3. ed. New Delhi: John Wiley e Sons, 2013. 641 p.
- MOSS, R. H. et al. The next generation of scenarios for climate change research and assessment. **Nature**, v. 463, n. 7282, p. 747–756, 2010.
- MOULIN, J. C. et al. Efeito do espaçamento, idade e irrigação no volume e densidade básica do eucalipto. **Floresta e Ambiente**, v. 24, p. 1–10, 2017.
- NIETO, P. G. et al. Support vector machines and neural networks used to evaluate paper manufactured using *Eucalyptus globulus*. **Applied Mathematical Modelling**, v. 36, n. 12, p. 6137–6145, 2012.
- NIETO, P. G. et al. Using multivariate adaptive regression splines and multilayer perceptron networks to evaluate paper manufactured using *Eucalyptus globulus*. **Applied Mathematics and Computation**, v. 219, n. 2, p. 755–763, 2012.
- OLANIYI, O.; OLUTIMEHIN, I.; FUNMILAYO, O. Review of climate change and its effect on nigeria ecosystem. **International journal of Rural Development, Environment and Health Research**, v. 3, n. 3, 2019.

OLIVEIRA, D.; BORROZZINO, E. Sazonalidade das tendências de temperatura do ar e precipitação pluvial no estado do paran . **Agrometeoros**, v. 26, n. 1, p. 139–149, 2018.

 Z ELIK, R. et al. Estimating tree bole volume using artificial neural network models for four species in Turkey. **Journal of environmental management**, v. 91, n. 3, p. 742–753, 2010.

 Z ELIK, R. et al. Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. **Forest Science**, v. 63, n. 3, p. 291–302, 2017.

PBMC. **Contribui o do Grupo de Trabalho 1 ao Primeiro Relat rio de Avalia o Nacional do Painel Brasileiro de Mudan as Clim ticas**. Rio de Janeiro, Brasil: Sum rio Executivo GT1, 2013. 24 p.

PENG, C. et al. Triplex: a generic hybrid model for predicting forest growth and carbon and nitrogen dynamics. **Ecological Modelling**, Elsevier, v. 153, n. 1, p. 109–130, 2002.

PETTITT, A. N. A non-parametric approach to the change-point problem. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, v. 28, n. 2, p. 126–135, 1979.

PEZZATTI, G. B. et al. Fire regime shifts as a consequence of fire policy and socio-economic development: an analysis based on the change point approach. **Forest Policy and Economics**, v. 29, p. 7–18, 2013.

PIMENTEL, D. J. O. et al. An lise de rentabilidade de manejo florestal sustent vel no assentamento Catol , em Serra Talhada-PE. **South American Journal of Basic Education, Technical and Technological**, v. 3, n. 1, p. 42–49, 2016.

PINO, F. A. A quest o da n o normalidade: Uma revis o. **Revista de economia agr cola**, v. 61, n. 2, p. 17–33, 2014.

POHLERT, T. **trend: Non-Parametric Trend Tests and Change-Point Detection**. [S.l.], 2018. R package version 1.1.1. Dispon vel em: <<https://CRAN.R-project.org/package=trend>>.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Dispon vel em: <<https://www.R-project.org/>>.

REIS, A. A. et al. Spatial prediction of basal area and volume in *Eucalyptus* stands using landsat tm data: an assessment of prediction methods. **New Zealand Journal of Forestry Science**, v. 48, n. 1, p. 1, 2018.

REIS, A. A. et al. Volume estimation in a *Eucalyptus* plantation using multi-source remote sensing and digital terrain data: a case study in minas gerais state, brazil. **International journal of remote sensing**, Taylor & Francis, v. 40, n. 7, p. 2683–2702, 2019.

REYER, C. Forest productivity under environmental change—a review of stand-scale modeling studies. **Current Forestry Reports**, v. 1, n. 2, p. 53–68, 2015.



REYER, C. P. et al. Are forest disturbances amplifying or canceling out climate change-induced productivity changes in european forests? **Environmental Research Letters**, v. 12, n. 3, p. 034027, 2017.

RIEGELHAUPT, E.; PAREYN, F. G. C. A questão energética e o manejo florestal da Caatinga. In: GARIGLIO, M. A. et al. (Ed.). **Uso sustentável e conservação dos recursos florestais da Caatinga**. Brasília, DF: Ministério do Meio Ambiente, Serviço Florestal Brasileiro, 2010. p. 256–275.

RIEGELHAUPT, E.; PAREYN, F. G. C.; BACALINI, P. O manejo florestal na caatinga: resultados da experimentação. In: GARIGLIO, M. A. et al. (Ed.). **Uso sustentável e conservação dos recursos florestais da Caatinga**. Brasília, DF: Ministério do Meio Ambiente, Serviço Florestal Brasileiro, 2010. p. 256–275.

ROCHA, K. D. **Produtividade volumétrica de clones de *Eucalyptus* spp. na região do Pólo gesseiro do Araripe**. 111 p. Tese (Ciências Florestais) — Universidade Federal Rural de Pernambuco, Departamento de Ciências Florestais, Recife, 2012.

RODRIGUES, M.; JIMÉNEZ, A.; RIVA, J. d. I. Analysis of recent spatial–temporal evolution of human driving factors of wildfires in spain. **Natural Hazards**, Springer, v. 84, n. 3, p. 2049–2070, 2016.

RYAN, M. G. et al. Factors controlling eucalyptus productivity: how water availability and stand structure alter production and carbon allocation. **Forest Ecology and Management**, v. 259, n. 9, p. 1695–1703, 2010.

SALEKIN, S. et al. A preliminary growth and yield model for *Eucalyptus globoidea* Blakely plantations in New Zealand. **New Zealand Journal of Forestry Science**, v. 50, p. 1–17, 2020.

SALES, F. d. C. V. et al. Ajustes de modelos volumétricos para o clone eucalyptus grandis x e. urophylla cultivados no agreste de pernambuco. **Floresta**, v. 45, n. 4, p. 663–670, 2015.

SANQUETTA, C. R. et al. Volume estimation of *Cryptomeria japonica* logs in southern brazil using artificial intelligence models. **Southern Forests: a Journal of Forest Science**, v. 80, n. 1, p. 29–36, 2018.

SANTANA, O. A. Minimum age for clear-cutting native species with energetic potential in the Brazilian semi-arid region. **Canadian Journal of Forest Research**, v. 47, n. 3, p. 411–417, 2016.

SANTOS, A. C. A. et al. Influência da variabilidade edafoclimática no crescimento de clones de eucalipto no nordeste baiano. **Pesquisa Florestal Brasileira**, v. 37, n. 91, p. 259–268, 2017a.

SANTOS, C. et al. Análise estatística da não estacionariedade de séries temporais de vazão máxima anual diária na bacia hidrográfica do rio pardo. **Holos**, v. 7, n. 32, p. 179–193, 2016.

- SANTOS, H. et al. O novo mapa de solos do Brasil: legenda atualizada. **Embrapa Solos-Documentos**, Rio de Janeiro: Embrapa Solos, p. 67, 2011.
- SANTOS, W. S. et al. Análise florística-fitosociológica e potencial madeireiro em área de caatinga submetida a manejo florestal. **Agropecuária Científica no Semiárido**, v. 13, n. 3, p. 203–211, 2017b.
- SCHIKOWSKI, A. B. et al. Modeling of stem form and volume through machine learning. **Anais da Academia Brasileira de Ciências**, v. 90, n. 4, p. 3389–3401, 2018.
- SCHLAEGEL, B. E. Green ash volume and weight tables. **Res. Pap.**, SO–206, p. 1–14, 1984.
- SCHNEIDER, R. et al. Growth and yield models for predicting tree and stand productivity. In: **Ecological Forest Management Handbook**. [S.l.]: ROUTLEDGE in association with GSE Research, 2016. v. 141, n. 178, p. 141–178.
- SCOLFORO, H. F. et al. Modeling dominant height growth of *Eucalyptus* plantations with parameters conditioned to climatic variations. **Forest Ecology and Management**, Elsevier, v. 380, p. 182–195, 2016.
- SCOLFORO, J. R. S. et al. Dominant height model for site classification of eucalyptus grandis incorporating climatic variables. **International Journal of Forestry Research**, v. 2013, 2013.
- SEREGHETTI, G. C. et al. Efeito do espaçamento no crescimento e na densidade básica da madeira de *Eucalyptus urophylla* x *Eucalyptus grandis* em florestas de ciclo curto. **Energia na Agricultura**, v. 30, n. 3, p. 257–262, 2015.
- SHIFLEY, S. R. et al. The past and future of modeling forest dynamics: from growth and yield curves to forest landscape models. **Landscape Ecology**, v. 32, n. 7, p. 1307–1325, 2017.
- SILVA, J. A. A. **Dynamics of stand structure in fertilized slash pine plantations**. 143 p. Tese (Biometrics and Forest Management) — University of Georgia, Athens, Georgia, 1986.
- SILVA, J. A. A. Potencialidades de florestas energéticas de *Eucalyptus* no Pólo Gesseiro do Araripe-Pernambuco. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 5 e 6, p. 301–319, 2008/2009.
- SILVA, J. A. A. Conceitos e princípios básicos de modelagem matemática em ciências florestais. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 11/12, p. 195–215, 2014/2015.
- SILVA, J. A. A.; NETO, F. d. P. **Princípios básicos de dendrometria**. [S.l.]: Imprensa Universitária da UFRPE, 1979. 198 p.
- SILVA, J. A. A. et al. Produtividade volumétrica de clones de *Eucalyptus* spp. no Polo Gesseiro do Araripe, Pernambuco. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 10, p. 240–260, 2013.

SILVA, J. A. A. et al. Modelagem do crescimento volumétrico de clones de eucalipto (*Eucalyptus urophylla*) no Polo Gesseiro do Araripe-PE. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 13/14, p. 173–190, 2016/2017.

SILVA, J. W. L. **Modelagem da biomassa e da quantidade de carbono de clones de *Eucalyptus* da Chapada do Araripe-PE**. 105 p. Dissertação (Biometria e Estatística Aplicada) — Universidade Federal Rural de Pernambuco, Recife, PE, 2016.

SILVA, J. W. L. et al. Quantificação do acúmulo de biomassa e de estoque de carbono de clones de *Eucalyptus* spp. na Chapada do Araripe-PE. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 15, n. 1, p. 147–158, 2018.

SILVA, R. S. et al. Desempenho silvicultural de *Tectona grandis* L. f. em diferentes espaçamentos em Cáceres, MT. **Floresta e Ambiente**, v. 23, n. 3, p. 397–405, 2016.

SILVEIRA, E. R.; REINER, D. A.; SMANIOTTO, J. R. Efeito do espaçamento de plantio na produção de madeira e serapilheira de *Eucalyptus dunnii* na região sudoeste do Paraná. **Revista Técnico-Científica**, v. 1, n. 2, p. 1–9, 2014.

SOARES, F. A. A. et al. Recursive diameter prediction and volume calculation of *Eucalyptus* trees using multilayer perceptron networks. **Computers and electronics in agriculture**, v. 78, n. 1, p. 19–27, 2011.

SOUZA, S. R. R. et al. Redes neurais para estimativa volumétrica de clones de *Eucalyptus* spp. no Polo Gesseiro do Araripe. **Revista Brasileira de Biometria**, v. 36, n. 3, p. 715–729, 2018.

SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. **The journal of machine learning research**, v. 15, n. 1, p. 1929–1958, 2014.

SUBRAMANIAN, N. **Impacts of climate change on forest management and implications for Swedish forestry**. Tese (Doctoral Thesis) — Swedish University of Agricultural Sciences, Alnarp, 2016.

SURATGAR, A. A.; TAVAKOLI, M. B.; HOSEINABADI, A. Modified levenberg-marquardt method for neural networks training. **World Acad Sci Eng Technol**, v. 6, p. 46–48, 2005.

TCHORBADJIEFF, A.; ANGELOV, I. Change point analysis as a tool to detect abrupt cosmic ray muons variations. In: GEORGIEV, K.; TODOROV, M.; GEORGIEV, I. (Ed.). **Advanced Computing in Industrial Mathematics**. Sofia, Bulgaria: Springer, 2017. p. 395–406.

TURNER, H.; FIRTH, D. **Generalized nonlinear models in R: An overview of the gnm package**. [S.l.], 2015. R package version 1.0-8. Disponível em: <<http://CRAN.R-project.org/package=gnm>>.

ULIANA, E. M. et al. Análise de tendência em séries históricas de vazão e precipitação: uso de teste estatístico não paramétrico. **Ambiente & Água-An Interdisciplinary Journal of Applied Science**, v. 10, n. 1, p. 82–88, 2015.

VALENÇA, M. J. **Aplicando redes neurais: Um guia completo**. Recife: Livro Rápido, 2005. 284 p.

VALENÇA, M. J. **Fundamentos das redes neurais: exemplos em Java**. Recife: Livro Rápido, 2010. 386 p.

VAPNIK, V. N. **Statistical learning theory**. New York: Wiley, 1998. v. 1. 736 p.

WAMELINK, G. et al. Modelling impacts of changes in carbon dioxide concentration, climate and nitrogen deposition on carbon sequestration by european forests and forest soils. **Forest Ecology and Management**, v. 258, n. 8, p. 1794–1805, 2009.

WEBER, V. P. et al. Modelagem linear generalizada para descrever o incremento em área transversal de árvores individuais de imbuia. **FLORESTA**, v. 48, n. 1, p. 123–132, 2018.

WHITE, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. **Econometrica: Journal of the Econometric Society**, p. 817–838, 1980.

YANG, K.; TU, J.; CHEN, T. Homoscedasticity: an overlooked critical assumption for linear regression. **General psychiatry**, v. 32, n. 5, 2019.

ZAMIN, N. T. et al. Effect of climate variables on monthly growth in modeling biological yield of *araucaria angustifolia* and *pinus taeda* in the juvenile phase. **International Journal of Forestry Research**, v. 2013, p. 1–9, 2013.